VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ
FACULTY OF INFORMATION TECHNOLOGY

ÚSTAV INFORMAČNÍCH SYSTÉMŮ
DEPARTMENT OF INFORMATION SYSTEMS

AUTOMATICKÁ DETEKCE AKTIVITY MALWARU V LOKÁLNÍCH SÍTÍCH

AUTOMATED DETECTION OF MALWARE ACTIVITY ON LOCAL NETWORKS

BAKALÁŘSKÁ PRÁCE
BACHELOR'S THESIS

AUTOR PRÁCE **ADAM PAP**
AUTHOR

VEDOUCÍ PRÁCE **doc. Ing. ONDŘEJ RYŠAVÝ, Ph.D.**
SUPERVISOR

BRNO 2024

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
FAKULTA INFORMAČNÍCH TECHNOLOGIÍ

Zadání bakalářské práce

157014

Ústav: Ústav informačních systémů (UIFS)
Student: **Pap Adam**
Program: Informační technologie
Název: **Automatická detekce aktivity malwaru v lokálních sítích**
Kategorie: Bezpečnost
Akademický rok: 2023/24

Zadání:

1. Seznamte se s problematikou komunikace malware.
2. Vytvořte datové sady pro účely návrhu a testování nástroje pro detekci malware ze síťové komunikace.
3. Proveďte analýzu dat a zaměřte se na extrakci významných vlastností.
4. Identifikujte další datové zdroje a informace použitelné pro identifikaci příznaků malware v síťové komunikaci.
5. Navrhněte metodu pro identifikaci malware komunikace kombinující různé techniky.
6. Implementujte metodu detekce malware v síťové komunikaci jako samostatný nástroj.
7. Vyhodnoťte nástroj pomocí vytvořených datasetů.
8. Integrujte vytvořený nástroj do vhodného prostředí pro monitorování síťového provozu.

Literatura:

- Pastor A, Mozo A, Vakaruk S, et al. Detection of encrypted cryptomining malware connections with machine and deep learning. *IEEE Access*. 2020;8:158036-158055. doi:10.1109/ACCESS.2020.3019658
- Oh C, Ha J, Roh H. A survey on tls-encrypted malware network traffic analysis applicable to security operations centers. *Appl Sci*. 2022;12(1). doi:10.3390/app12010155
- Roques O, Maffeis S, Cova M. Detecting Malware in TLS Traffic. 2019;(September).
- Anderson B, Chi A, Dunlop S, McGrew D. Limitless HTTP in an HTTPS World: Inferring the Semantics of the HTTPS Protocol without Decryption. Published online May 29, 2018. http://arxiv.org/abs/1805.11544

Při obhajobě semestrální části projektu je požadováno:
Body 1-4.

Podrobné závazné pokyny pro vypracování práce viz https://www.fit.vut.cz/study/theses/

Vedoucí práce: **Ryšavý Ondřej, doc. Ing., Ph.D.**
Vedoucí ústavu: Kolář Dušan, doc. Dr. Ing.
Datum zadání: 1.11.2023
Termín pro odevzdání: 9.5.2024
Datum schválení: 20.10.2023

Abstrakt

Cieľom tejto práce je analyzovať sieťovú komunikáciu malware a následne identifikovať vhodné významné vlastnosti, ktoré by umožnili vytvoriť vhodnú metódu na jeho detekciu. Súčasťou riešenia práce bolo aj vytvorenie dátovej sady z ktorej sa extrahovali IoC jednotlivých malware rodín. Tieto IoC boli následne overené cez platformu AlienVault OTX, z dôvodu overenia ich relevantnosti. Pre vyhodnotenie boli použité metriky ako miera falošnej pozitivity, presnosť a senzitivita. Na testovacích dátach oba IoC modely, vytvorených z dátových sád, dosiahli presnosť 99.337% a 94.732% pre dátovú sadu č. 2. IoC modely sady č. 1 v reálnej prevádzke falošne klasifikovali 3.03% komunikačných okien ako škodlivých. IoC modely sady č. 2 klasifikovali 5.66% okien škodlivými. Následne boli v testovacom prostredí spustené vzorky rôznych malware rodín, kde IoC modely sady č. 1 klasifikovali 7.14% okien škodlivými. Modely sady č. 2 klasifikovali 15.79%

Abstract

The aim of this work is to analyze the network communication of malware and then identify suitable significant features that would allow to develop a suitable method for its detection. As part of the solution of the thesis, a dataset was created from which an IoC for each malware family were extracted. These IoCs were then validated through the AlienVault OTX platform, in order to verify their relevance. Metrics such as false positive rate, accuracy and sensitivity were used for evaluation. On the test data, the two IoC models created from the datasets achieved an accuracy of 99.337% and 94.732% for dataset 1 and 2, respectively. The IoC models of dataset No. 1 falsely classified 3.03% of communication windows as malicious in real communication. IoC models of set No. 2 classified 5.66% as malicious. After the samples of different malware families were run on the machine, the IoC models of set No. 1 classified 7.14% of the windows as malicious. Set No. 2 models classified 15.79%.

Kľúčové slová

škodlivý softvér, dátová sada, rodiny škodlivého softvéru, lokálna sieť, indikátor kompromitácie, sieťová komunikácia, fuzzy množiny

Keywords

malware, dataset, malware family, local network, indicator of compromise, network communication, fuzzy set

Citácia

PAP, Adam. *Automatická detekce aktivity malwaru v lokálních sítích.* Brno, 2024. Bakalářská práce. Vysoké učení technické v Brně, Fakulta informačních technologií. Vedoucí práce doc. Ing. Ondřej Ryšavý, Ph.D.

Automatická detekce aktivity malwaru v lokálních sítích

Prehlásenie

Prehlasujem, že som túto bakalársku prácu vypracoval samostatne pod vedením pána doc. Ing. Ondřeja Ryšavého. Uviedol som všetky literárne pramene, publikácie a ďalšie zdroje, z ktorých som čerpal.

.....
Adam Pap
7. mája 2024

Poďakovanie

Týmto by som sa poďakovať pánovi docentovi Ryšavému za jeho trpezlivosť, odborné rady a pozitívny prístup počas konzultácií. A tiež by som sa chcel poďakovať svojej rodine za konštantnú podporu počas štúdia.

Obsah

Kapitola 1

Úvod

V súčasnej dobe plnej informácií, internetových technológií a ďalších vymožeností technologického pokroku, ktoré značne uľahčujú život, prichádza aj mnoho negatívnych aspektov s týmto spojené. Neustále sa vyskytujúce útoky na internete s cieľom získať alebo iným spôsobom zneužiť dáta jednotlivca prípadne skupiny ľudí prostredníctvom škodlivého softvéru viedlo k vytvoreniu rôznych monitorovacích nástrojov.

S ich pomocou je možné chovanie takéhoto škodlivého softvéru vo väčšine prípadov včas odhaliť v rámci lokálnej siete a patrične naň zareagovať. Na to aby takýto dômyselný monitorovací systém mohol vzniknúť je potrebné navrhnúť vhodnú metódu, ktorá by takúto analýzu bola schopná vykonávať v reálnom čase reálnej sieťovej prevádzky. Nakoľko objem dát, ktoré sa v súčastnosti prenášajú po sieti je obrovský a podrobne analyzovať každý tok po sieti by bolo príliš časovo náročné. Preto sa v rámci detekcie škodlivej komunikácie, zvyčajne spojenej s nejakým škodlivým softvérom, zameriava na určité významné vlastnosti, ktorými sa daný škodlivý softvér vyznačuje napr.: prístup k podozrivým doménam, IP adresám a podobne.

Avšak predtým než vôbec dôjde k návrhu metódy je nutné vedieť na aké vlastnosti sa zamerať pri sieťovej komunikácií škodlivého softvéru a na to poslúžia dátové sady, ktoré obsahujú jak škodlivú tak normálnu komunikáciu. Ďalšou časťou tejto práce je teda už spomenutá analýza dátových sád a extrakcia významných vlastností, ktoré by poskytli cenné informácie vhodné pre identifikáciu nežiaducej komunikácie v sieti. Následne vytvorené dátové sady poslúžili na otestovanie úspešnosti vytvorenej metódy a v poslednom rade došlo k nasadeniu metódy do virtuálneho prostredia vytvoreného prostredníctvom virtualizovaných klientov, ktorý simulovali jednoduchú lokálnu sieť.

V nasledujúcej kapitole 2 je popísané čo je to malware, ako sa malware chová v sieti a v neposlednom rade čo je to indikátor kompromitácie (anglicky Indicator of Compromise - IoC), na aké typy sa delí a aký je rozdiel medzi indikátorom útoku (anglicky Indicators of Attack) a indikátorom kompromitácie (anglicky Indicator of Compromise). V kapitole 3 sa rieši ako bola vytvorená dátová sada a akými spôsobmi sa zo spleti dát získali dáta, ktoré poskytujú kritické informácie pre vytvorenie spoľahlivej detekčnej metódy. A v neposlednej rade sa v tejto kapitole rieši aj akým spôsobom sa pristúpilo k overeniu relevantnosti získaných indikátorov kompromitácie. V kapitole 4 sa preberá zvolená metóda detekcie, ako sa spracúvajú jednotlivé indikátory kompromitácie, a na akom základe prebieha detekcia malwaru či už online alebo offline (analýza dopredu zachytenej sieťovej komunikácie nástrojom *Suricata*). Kapitola 5 sa zaoberá experimentami, ktoré boli vykonané jak nad dopredu zachytenou komunikáciou tak aj v reálnej sieťovej prevádzke.

Kapitola 2

Malware, komunikácia malwaru v lokálnej sieti

Dnešný digitálny svet je plný nástrah a nebezpečenstiev. Medzi ne môžeme uvažovať práve škodlivý softvér respektíve malware. Mnoho dnešných malwerov dokáže komunikovať po sieti, odosielať citlivé údaje obetí útočníkom alebo autorom daného škodlivého softvéru. Táto kapitola sa zameriava práve na vysvetlenie pojmov ako čo je to malware, aké typy malwaru existujú, čo to je lokálna sieť, ako sa prejavuje malware v rámci lokálnej siete a čo sú to indikátory kompromitácie a ako súvisia s identifikáciou malwaru.

2.1 Malware

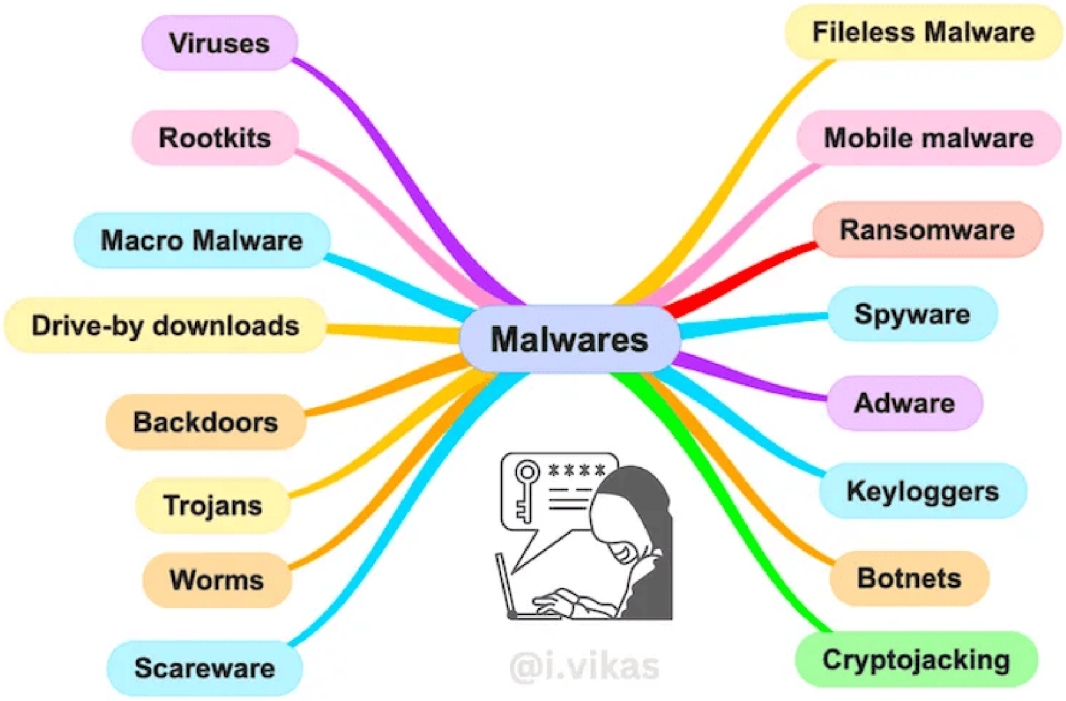
Slovo malware respektíve (slovensky škodlivý softvér) je zloženie dvoch slov ***mal****icious* zlomyseľný a *soft****ware*** softvér, tiež ho môžeme označiť ako zlomyseľný kód či softvér, ktorý sa zvyčajne snaží poškodiť, zablokovať alebo inak zmocniť informácií v napadnutom systéme [2]. Toto je samozrejme len jedna z mnohých definícií, ktorá bola ponúknutá. Existujú aj iné definície, ako napríklad, že malware je softvér, ktorý zámerne vykonáva zámer útočníka respektíve autora daného malwaru, ktorý je zvyčajne škodlivého charakteru [15].

Existuje mnoho spôsob ako sa malware dokáže dostať do počítačového systému a infikovať ho. Jedny z najčastejších spôsobov ako sa to môže stať je Internet a email. Inými slovami to znamená, že k napadnutiu môže dôjsť hocikedy počas toho ako je zariadenie pripojené k Internetu. Avšak nie všetky webové stránky a iné zdroje, ktoré sa nachádzajú na Internete sú škodlivé. Zariadenie sa môže infikovať ak na ňom užívateľ prezerá obsah infikovanej webovej stránky alebo keď navštevuje legitímnu stránku, ktorá zobrazuje škodlivé reklamy, sťahuje infikované súbory, inštaluje programy alebo aplikácie od neznámych prípadne neoverených poskytovateľov, otvára prílohy od neznámych potencionálne nebezpečných emailových adries a podobne. V skutočnosti je omnoho viacej spôsobov ako sa zariadenie môže nakaziť malwarom.

Motívom autorov malwaru sú rôzne. Cieľom môže byť zarábať peniaze na obeti, sabotovať schopnosť obete pracovať respektíve dokončiť svoju prácu alebo šíriť nejaký politický názor. Malware ako taký nemá možnosť fyzicky poškodiť hardware napadnutého zariadenia avšak môže odcudziť informácie uložené v zariadení ktoré napadne, zašifrovať alebo odstrániť dáta, zmeniť alebo ukradnúť základné funkcionality zariadenia, špehovať užívateľa bez jeho vedomia alebo súhlasu [5].

2.2 Typy malwaru

Malware ako taký je široký pojem, ktorý zahŕňa množstvo rôznych typov škodlivého softvéru. Na základe typu je možné malware rozdeliť do niekoľkých skupín [20]. Takéto rozdelenie je možné vidieť na obrázku 2.1.



Obr. 2.1: Obrázok zobrazujúci rôzne typy malwaru, prebraný z [10].

- **Vírus** je jedným z najjednoduchších typov softvéru. Jedná sa o kus kódu ktorý je súčasťou spustiteľnej aplikácie bez vedomia užívateľa. Akonáhle je spustený, ako súčasť inej legitímnej aplikácie, dokáže ukradnúť citlivé dáta prípadne spustiť iný útok [20] [13].
- **Červ** je veľmi podobný vírusu avšak hlavným rozdielom medzi nimi je fakt, že červ sa dokáže rozšíriť na ostatné zariadenia prostredníctvom siete a k jeho spusteniu nie je potrebný zásah užívateľa [20].
- **Trójsky kôň** je trieda malwaru, ktorá sa javí ako legitímna aplikácia a to spôsobuje, že je ťažko odhaliteľný. Preto sa táto trieda malwaru spolieha na sociálne inžinierstvo ako jeden z hlavných spôsobov svojho rozširovania sa na ďalšie zariadenia [20].
- **Adware**, jeho jediným cieľom je ukazovať užívateľovi reklamy a tým potencionálne zarábať peniaze jeho tvorcom. Na tento typ malwaru sa dá pozerať aj ako na podmnožinu spywaru (špehovací softvér) [20].

- **Spyware**, ako už z mena vyplýva jedná sa o typ malwaru, ktorý špehuje napadnuté zariadenie. Medzi jeho typické praktiky patrí napríklad, monitorovanie histórie vyhľadávania, ktorú následne posiela respektíve predáva tretím stranám pre upravenie zobrazovaných reklám [20].

- **Rootkit** je ďalšou triedou malwaru, ktorá poskytuje prístup k zdrojom užívateľovi (útočníkovi), ktorý by za normálnych okolností nemal mať k nim prístup. Tento malware je veľmi dobre schovaný a preto je náročné ho nájsť a odstrániť z napadnutého systému [20].

- **Keylogger** je typ malwaru, ktorý cieli na logovanie všetkých užívateľom stlačených kláves a následné uloženie týchto dát. Medzi týmito dátami sa môžu vyskytovať aj citlivé dáta ktoré keylogger odchytil z klávesnice a to napríklad: heslá, čísla platobných kariet, PIN kódy a iné citlivé dáta [20].

- **Ransomware** je jedným asi z najnepríjemnejších typov malwaru. Jeho cieľom je zašifrovať všetky dáta napadnutého zariadenia a požadovať od obete isté množstvo peňazí (avšak v súčastnosti sa vyžadujú platby aj v kryptomenách) aby dostala dešifrovací kľúč. Typicky zariadenia napadnuté týmto typom malwaru sú plne funkčné, avšak všetky dáta na danom zariadení sú zašifrované a zvyčajne sa na obrazovke nachádzajú informácie o požiadavku útočníka [20].

- **Botnet** je sieť počítačov, ktoré sú infikované malwarom a sú pod kontrolou jedného útočníka tiež označovaného ako *bot-herder*. Každé infikované zariadenie, nazývaného tiež aj ako bot, pracuje spoločne s ostatnými zariadeniami v botnete. Útočník (*bot-herder*) riadi vzájomnú prepojenosť medzi týmito napadnutými zariadeniami a využíva ich na vykonávanie rôznych kybernetických aktivít ako napríklad spúšťanie automatických skriptov v sieti [6].

Malware rodiny

V rámci pojmu malware je ešte ďalší dôležitý pojem a tým sú malware rodiny (anglicky malware families). Rodina malwaru je skupina malwaru, ktorá zdieľa spoločné charakteristiky a zvyčajne aj autora. Každý poskytovateľ bezpečnostného softvéru obvykle používa vlastné názvy pre jednotlivé malware rodiny [7]. Mnohé z týchto rodín sa dajú nájsť na internetových stránkach napríklad na Malpedia[1] a jednotlivé vzroky sa dajú zohnať napríklad na stránke Abuse.ch MalwareBazaar[2].

Takýto malware má podobné vlastnosti, ktoré môžu byť užitočné pri vytváraní signatúr na detekciu a klasifikáciu. Tieto signatúry môžu byť buď statické alebo dynamické v závislosti od toho ako boli získané. Statická signatúra môže byť založená na sekvencií bajt kódu, inštrukcií binárneho asembleru alebo importovanej knižnici DLL (anglicky Dynamic Link Library). Dynamické signatúry sú zvyčajne založené na terminálových príkazoch sieťovej komunikácie, sekvencií funkcií a systémových volaní alebo na aktivitách súborového systému [18].

[1] https://malpedia.caad.fkie.fraunhofer.de/families
[2] https://bazaar.abuse.ch/browse/

2.3 Lokálna sieť

Lokálne siete sa začali masívne rozširovať po nástupe osobných počítačov (anglicky *Personal Computers* (PC)), zvyšovaním ich výkonnosti zmenšovaním sa. Toto malo za následok obrovský nárast zariadení vo firmách, továrňach a v neposlednom rade aj v domácnostiach, čo viedlo k myšlienke prepojiť tieto zariadenia medzi sebou. Medzi dôvody prečo sa tak deje sa môžu uviesť napríklad zdieľanie zdrojov a výmena dát medzi zariadeniami. Možnosť výmeny údajov medzi zariadeniami je tiež presvedčivým dôvodom pre prepojenie. Jednotlivý užívatelia počítačov nepracujú samostatne ale skôr pracujú ako súčasť lokálnej siete, čo poskytuje výmenu správ, príloh s inými užívateľmi a prístup k údajom a programom z viacerých zdrojov.

Tieto požiadavky na komunikáciu medzi viacerými počítačovými systémami v rámci organizácie a medzi počítačmi a zdieľanými zdrojmi spĺňa definíciu lokálnej siete, ktorá je definovaná takto: Lokálna sieť je komunikačná sieť, ktorá poskytuje prepojenie rôznych komunikujúcich zariadení v rámci malej oblasti. Pojmom malá oblasť sa myslí jedna budova, prípadne skupina viacerých navzájom susediacich budov [21].

Malware v rámci siete

Moderné typy malwaru využívajú sofistikované spôsoby ako sa zamaskovať nielen pred modernými anti-vírusovými programami, ale aj pred skúsenými sieťovými administrátormi. Mnohé z malware programov používajú Internet na to aby mohli komunikovať s riadiacim serverom (anglicky *Command and Control*, (C&C)) a prijímať od neho ďalšie príkazy, úlohy prípadne softvérové aktualizácie. Avšak keď sa takýto malware pokúša komunikovať cez sieť, zvyčajne k tomu využíva dobre známe protokoly (napr. HTTP, HTTPS) aby sa nepozorovane mohol dostať cez bránu firewall. V niektorých prípadoch práve populárne webové stránky môžu byť zapojené v týchto zlomyseľných aktivitách ako proxy [14].

2.4 Indikátory kompromitácie (IoC)

Indikátor kompromitácie (anglicky Indicator of Compromise) sú forenzné dáta, ktoré sú využívané v rámci kybernetickej bezpečnosti pre potvrdenie alebo vyvrátenie výskytu kybernetických útokov. Mimo iné sa používajú aj v rámci vytvárania rôznych obranných stratégií proti spomínaným útokom. Indikátory kompromitácie sú tiež použiteľné na zistenie bezpečnostných slabín daného systému a tiež na to akým spôsobom bol kybernetický útok vykonaný. I keď významnosť IoC nemožno v oblasti bezpečnosti bagatelizovať, nie sú všetkým čo je potrebné na vytvorenie účinnej obrannej stratégie. Skôr su brané ako istý predpoklad respektíve pravdepodobnosť že daný systém je pod útokom malwaru.

Čiže v rámci definície IoC je možné povedať, že indikátory kompromitácie (ďalej ako IoC) definujú správanie alebo dáta ktoré poukazujú na to, že došlo k prieniku do systému, narušeniu jeho dát alebo k inému útoku. Ich výskyt jasne poukazuje na to, že daný systém, doména, sieť má isté bezpečnostné slabiny/trhliny, na ktoré by sa v budúcnosti malo zamerať a pracovať na ich odstránení. Dáta získané z IoC nachádzajú uplatnenie primárne počas vyhľadávania hrozieb nakoľko tieto dáta sú zvyčajne k dispozícií až potom čo došlo k prieniku do systému.

Mnoho odborníkov z oblasti kybernetickej bezpečnosti sa zhodli na troch základných znakoch, ktoré musí dané IoC spĺňať aby mohlo byť vôbec považované za IoC. Jedná sa o tieto tri podmienky:

- IoC musí vykazovať príznaky toho, že došlo k nejakej neoprávnenej prípadne inej škodlivej udalosti v systéme. Jedná sa o podmienku pozorovateľnosti.
- Musia existovať nejaké doplniteľné informácie, ktoré by mohli byť nápomocnými bezpečnostným tímom. Medzi takéto metadáta sa môžu radiť napríklad zdroje IoC, čas výskytu a iné artefakty, ktoré sú nejakým spôsobom prepojené s útokom. Tu sa jedná o podmienku metadát.
- Artefakt by mal zapadať do kontextu útoku, ktorý sa odohral. Napríklad v prípade ak sa jedná o *phishing* útok, IoC by mohli byť napríklad podozrivé emailové prílohy, URL ktoré sú typickými prvkami prezrádzajúcimi spomenutý typ útoku. Jedná sa o podmienku kontextu.

Práve vďaka IoC vieme dohľadať odkiaľ prípadne aj z akého zariadenia bol útok vykonaný a aj aké nástroje, a aplikácie boli použité [19].

Artefakty a indikátory kompromitácie

Tu by bolo vhodné sa pozastaviť nad tým čo to vlastne je artefakt a IoC a aký je medzi nimi rozdiel, nakoľko sa mnohokrát pletie význam týchto dvoch termínov. Je to aj z dôvodu toho, že mnoho aplikácií, ktoré sa používajú na boj proti malwaru a kybernetickým útokom moc nerozlišujú tieto dva pojmy.

- **Artefakty** sú časti forenzných dát pozorovaných v priebehu vykonávania útoku, ktoré boli získané počas dynamickej analýzy malwaru alebo hrozby v prostredí *sandbox*-u. Tieto získané dáta sa označujú aj ako artefakty analýzy (anglicky *analysis artifacts*), ktoré zvyčajne obsahujú URL adresy, súbory, IP adresy, záznamy o zmenách položiek v registroch, ktoré boli použité, vytvorené alebo upravené malwarom počas vykonávania útoku [11] .
- **Indikátor kompromitácie** je časť forenzných dát priamo súvisiacich s daným malwarom. Tieto dáta môžu byť použité na identifikáciu budúcich útokov daného malwaru na počítačový systém alebo sieť. Platí teda tvrdenie, že IoC môžu byť kombináciou určitých artefaktov alebo sa môže jednať len o jeden konkrétny artefakt [11].

Indikátory útoku a aký je rozdiel oproti Indikátorom kompromitácie

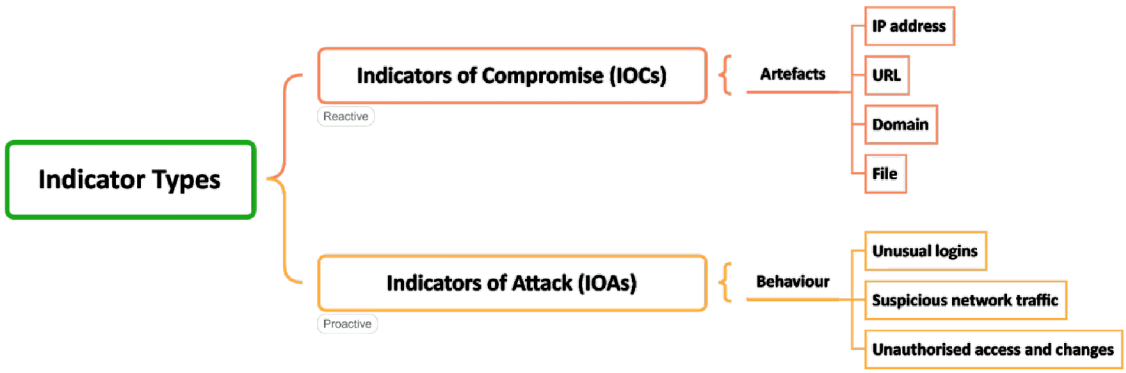
Indikátory útoku (anglicky Indicators of Attack) je možné definovať ako správanie alebo vzorce (anglicky patterns) používané pre identifikáciu prebiehajúceho útoku. Hlavným rozdielom medzi IoA a IoC je stav útoku, inak povedané či je útok vykonávaný práve teraz alebo už bol vykonaný. IoA identifikuje zámer malwaru a techniky použité počas útoku na počítačový systém alebo sieť. Ak IoA odhalí práve potencionálne prebiehajúci útok tak IoCs sú použité pre dôkladnejšiu analýzu daného útoku, ktorý už prebehol. Pomocou IoAs vieme zamedziť danému útoku počas jeho vykonávania.

Oba spomenuté indikátory sú dôležité, avšak existujú ešte ďalšie významné rozdiely na základe ktorých ich vieme identifikovať.

Prvým z týchto rozdielov je, že IoA poskytujú včasné informácie, ktoré sá zídu vhod pri riešení prebiehajúcich útokov. Včasná identifikácia IoA zvyčajne znamená, že situáciu alebo dáta je možné ešte zachrániť pred tým, než dôjde k jej eskalácií a ďalšiemu zhoršeniu. IoC na druhej strane môžu naznačiť kto je za útokom a ako k nemu došlo, za použitia akých nástrojov.

Ďalším z významných rozdielov je, že v prípade odhalenia útoku počas jeho vykonávania je možné tento útok zastaviť pred tým než sa situácia zvrhne v katastrofu a tým predísť zbytočne veľkým finančným stratám prípadne riziku straty všetkých dát. Povaha analýzy za použitia IoC umožňuje reagovať na útok až po jeho skončení.

Čiže hlavným rozdielom je, že IoC sú založené na známej škodlivej činnosti určitej malware rodiny, zatiaľ čo IoA sú založené na konkrétnych taktikách, technikách a postupoch, využívaných útočníkmi viď. obrázok 2.2. Takisto IoC sú zvyčajne reaktívne, zatiaľ čo IoA sú proaktívne a môžu pomôcť pri odhalení potencionálnych hrozieb, skôr než spôsobia škody [4].



Obr. 2.2: Rozdiel medzi indikátorom kompromitácie IoC a indikátorom útoku IoA z pohľadu ich použiteľnosti pri odhalení potencionálnych útokov, prebraný z [4].

Posledným z významných rozdielov je fakt, že IoA sa nepopisujú ako forenzné dáta, ale jedná sa skôr o vzory a techniky, ktoré naznačujú prebiehajúci útok. Z toho vyplýva, že tieto dáta sú dosť nepredvídateľné a sú predmetom zmeny na základe zámerov malwaru alebo cieľov útočníka. Naopak IoC sú zvyčajne potvrdené dáta, ktoré majú istý formát na základe ktorého ich možno klasifikovať a porovnávať s minulými informáciami. Inými slovami sú to statické dáta s ktorými sa vo všeobecnosti ľahšie pracuje [19].

Typy Indikátorov kompromitácie

V rámci kybernetickej bezpečnosti existujú tri typy IoCs a to:

- **Sieťové IoCs** sú zvyčajne detegované analyzovaním sieťovej prevádzky. Medzi tieto IoCs sa zaraďujú podozrivé IP adresy, názvy domén, URL a podobne [19].
- **Súborové IoCs** šu súčasťou súborov, ktoré sa nachádzajú na hostiteľskom systéme. Môže sa jednať o súbory s hašovaním, názvy súborov alebo cesty k súborom [19].
- **Behaviorálne IoCs** sa zvyčajne zisťujú pozorovaním rôznych vzorcov (anglicky patterns) správania v systéme alebo sieti, ktoré môžu naznačovať škodlivú činnosť. Medzi takéto typy správania sa môžu radiť napríklad vysoká návštevnosť webovej lokality (DDoS útok), opakované zadávanie nesprávneho hesla pri prihlásení a iné [19].

Formy výskytu IoC

IoC sa zvyčajne vyskytujú v niekoľkých formách. Medzi najbežnejšie z nich patrí napríklad podozrivá komunikácia, ktorá vychádza z konkrétneho počítaču, veľký počet neúspešných prihlásení, komunikácia s nejakou exotickou IP adresou a ďalšie [19].

Význam jednotlivých spomenutých príkladov resp. foriem je následovný:

- **Podozrivá odchádzajúca sieťová komunikácia**.

  Zvyčajne to je známka útoku typu C&C (anglicky Command and Control), kedy útočník môže prevziať kontrolu nad daným kompromitovaným zariadením a využiť ho pre svoje potreby. Tiež to môže naznačovať únik resp. stratu dát. Čiže je možné povedať, že hlavným dôsledkom tohoto IoC je strata dát [19].

- **Opakované pokusy o neúspešné prihlasovanie**.

  Neúspešné prihlásenie je u mnohých užívateľov pravidelným a bežným javom. Avšak v niektorých prípadoch to môže naznačovať, že útočník používa na prihlásenie, či už do systému alebo služby, falošné alebo nepresné prihlasovacie údaje. V tomto prípade sa môže jednať o snahu o ukradnutie účtu alebo kompromitovať daný systém/službu vo všeobecnosti [19].

- **Komunikácia s exotickou IP adresou**.

  Ak je IP adresa s ktorou zariadenie komunikuje z oblasti, ktorá nie je typickou, je vhodné zvýšiť pozornosť. Nakoľko útočníci zvyčajne pracujú z neznámych lokalít, menia si často IP adresu zariadenia z ktorého vykonávajú útok. To všetko z dôvodu toho aby získali ľahšie prístup do zariadenia obete, ktorá si myslí, že komunikuje s dôverným zariadením. Prípadne aby bolo ťažšie vypátrať identitu útočníka [19].

Kapitola 3

Dátová sada, analýza a extrakcia významných vlastností malwaru

Hlavným zámerom tejto práce je detegovať malware v rámci lokálnej siete avšak na to aby bolo toto možné je nutné navrhnúť vhodnú metódu. Preto je potrebné ako prvé určiť, aké vlastnosti malware počas komunikácie v rámci siete zanechával respektíve zanecháva. Toto vedie na zistenie všetkých dostupných informácií o jednotlivých malware vzorkoch. Táto práca sa zameriava výhradne na indikátory kompromitácie (anglicky Indicator of Compromise - IoC) a ich analýzou.

V rámci tejto kapitoly bude reprezentovaná dátová sada, za pomoci čoho bola vytvorená, aké informácie sa zistili z analýzy a aké vlastnosti malwaru sa podarilo identifikovať. Ďalej sa bude riešiť extrakcia indikátorov kompromitácie (IoC) z dátovej sady a tiež aký nástroj bol použitý pre overenie validity extrahovaných IoC.

3.1 Vytvorenie dátovej sady

Dátová sada bola vytvorená za pomoci nástroja vytvoreného v rámci bakalárskej práce Detekce komunikace malware v síťových tocích [17]. Dátová sada, ktorá bola vytvorená za účelom návrhu a testovania nástroja pre detekciu malwaru zo sieťovej komunikácie poslúžila tiež aj na pre analýzu, prostredníctvom ktorej sa identifikovali významné vlastnosti jednotlivých malware rodín. Dôvodom prečo sa dátová sada vytvárala miesto toho aby sa použila existujúca dátová sada je fakt, že mnohé z nich nemajú riadnu dokumentáciu a tiež sú mnohokrát zastaralé alebo irelevantné. Ďalším z dôvodov je tiež to, že mnoho týchto dátových sád nemá takú úroveň detailov zachytených počas analýzy vzorku malwaru akú má vytvorená sada. Nástroj, pomocou ktorého bola daná dátová sada zhotovená, používa sandbox Triage[3]. Triage je sandboxová webová aplikácia, do ktorej je prostredníctvom webového prehliadača možné vkladať rôzne malware vzorky na analýzu. Analýza takéhoto vzorku pozostáva zo statickej a dynamickej analýzy. Výsledky oboch analýz sa dajú stiahnuť prostredníctvom webového prehliadača alebo aplikácie s prístupom na API, vo formáte PDF alebo JSON (anglicky Javascript object notation)[4]. Sada vytvorená v rámci tejto práce pozostáva zo 17 malware rodín (viď. podkapitola 2.2) a každá rodina má po 5 vzorkov malwaru, rodiny boli vybraté na základe webovej stránky *any-run*[5].

[3] https://tria.ge/

[4] https://www.json.org/json-en.html

[5] https://any.run/cybersecurity-blog/malware-trends-q3-2023/

V rámci analýzy sa zameriavalo primárne na indikátory kompromitácie (viď. podkapitola 2.4), ďalej len ako IoC, jednotlivých vzorkov malwaru. Analýza týchto forenzných dát bola vykonaná s pomocou postupov z knihy *Data driven Security*[16] a jazyku python, konkrétne za použitia *IPython notebook*[6].

Analýza dátovej sady

Ako už bolo spomenuté je nutné najprv získať dáta na osnove ktorých bude možné zostrojiť vhodný klasifikátor. Avšak získanie takýchto dát nie je nutne jednoduchá záležitosť, nakoľko mnoho ukazovateľov, ktoré sú poskytnuté výsledkami dynamickej alebo statickej analýzy sú nepoužiteľné, irelevantné alebo jednoducho zanedbateľné a preto sa neoplatí nimi ďalej zaoberať. Výsledky statickej analýzy sú ignorované nakoľko nemajú žiadny vplyv v rámci tejto práce.

Tiež je nutno poznamenať, že hlásenia (anglicky reports) pre jednotlivé malware rodiny, ktoré boli poskytnuté zo sandbox prostredia Triage sú vo formáte JSON. Tento formát je vhodnejším pre spracúvanie štruktúrovaných dát, nakoľko umožňuje omnoho bohatšiu ale aj komplexnejšiu reprezentáciu získaných dát z analýz než napríklad CSV[7] formát. I keď XML (anglicky eXtensible Markup Language)[8] je tiež vcelku vhodným formátom pre reprezentáciu alebo prenos dát, tak oproti JSON formátu je menej čitateľný a náročnejší na spracovanie[16].

Ako prvé je vhodné sa najprv oboznámiť s dátami, ktoré sú k dispozícií, akého dátového typu sú (*integer*, *float*, *string*), aké hodnoty môžu nadobúdať, čo za hodnoty sa najčastejšie v dátovej sade vyskytujú alebo opakujú. V tabuľke 3.1 je možné vidieť úryvok dát z dátovej sady. Je možné vidieť, že niektoré záznamy v stĺpci *Application protocol* majú znak pomlčky, to je z dôvodu, že nástroj pre vytváranie dátových sád pracuje len s protokolmi *DNS* (Domain Name System), *HTTPS* (Hypertext Transfer Protocol Secure), *HTTP* (Hypertext Transfer Protocol) a *TLS* (Transport Layer Security), ak protokol nebol uvedený, je na jeho mieste uvedená pomlčka. To isté platí aj v prípade ak je daný tok označený ako *normal*, v tom prípade rodina malwaru nie je uvedená z pochopiteľných dôvodov.

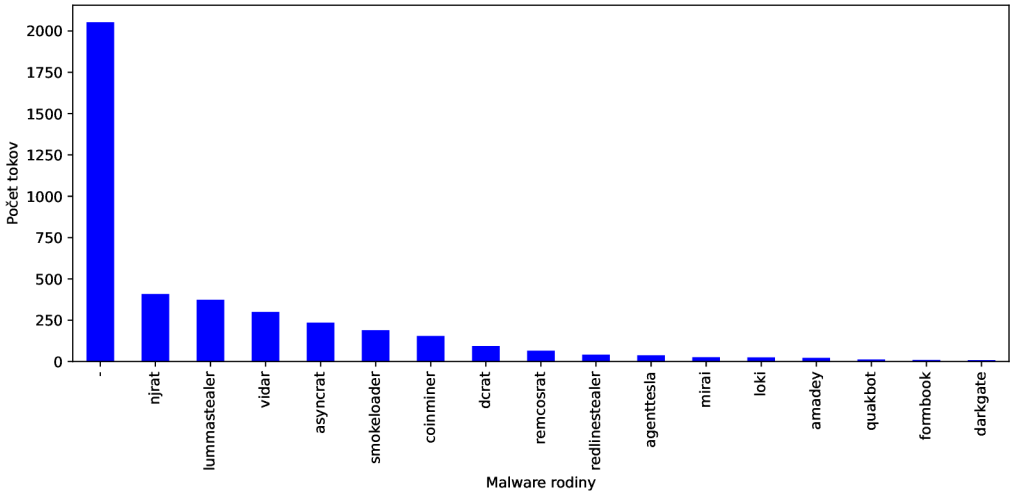
Src IP	Dst IP	Dst port	Protocol	Application protocol	Duration	Total bytes	Total packets	label	family
10.127.0.29	64.185.227.156	443	tcp	tls	367	596	10	malware	agenttesla
10.127.0.201	104.237.62.212	443	tcp	http	100757	7983	23	malware	agenttesla
10.127.0.188	63.250.35.178	587	tcp	-	100146	9620	44	malware	agenttesla
10.127.0.177	158.160.82.150	443	tcp	https	583	937	11	malware	coinminer
10.127.0.201	8.8.8.8	53	udp	dns	20	230	2	normal	-
10.127.0.61	67.195.204.77	25	tcp	-	15009	260	5	malware	coinminer
10.127.0.13	224.0.0.251	5353	udp	-	128045	146	2	malware	mirai
10.127.0.224	142.251.39.100	80	tcp	http	218	3987	13	normal	-
10.127.0.6	193.37.197.23	80	tcp	http	864	1170	10	malware	smokeloader
10.127.0.69	104.21.35.235	443	tcp	tls	38	5181	9	malware	vidar

Tabuľka 3.1: Tabuľka zobrazujúca časť dát z dátovej sady.

Z pohľadu dát, sa tu vyskytujú dva typy a to kvalitatívne a kvantitatívne dáta. Kvantitatívne dáta popisujú množstvo niečoho, v prípade tejto práce sa jedná napríklad o *Total packets*, naopak kvalitatívne prvky dát majú skôr opisný charakter. Napríklad čísla portov sú reprezentované číslicami ale nereprezentujú množstvo, port 443 nie je väčší ako port 80 preto čísla portov patria do kvalitatívnej skupiny [16]. Pre lepšiu interpretáciu toho aké

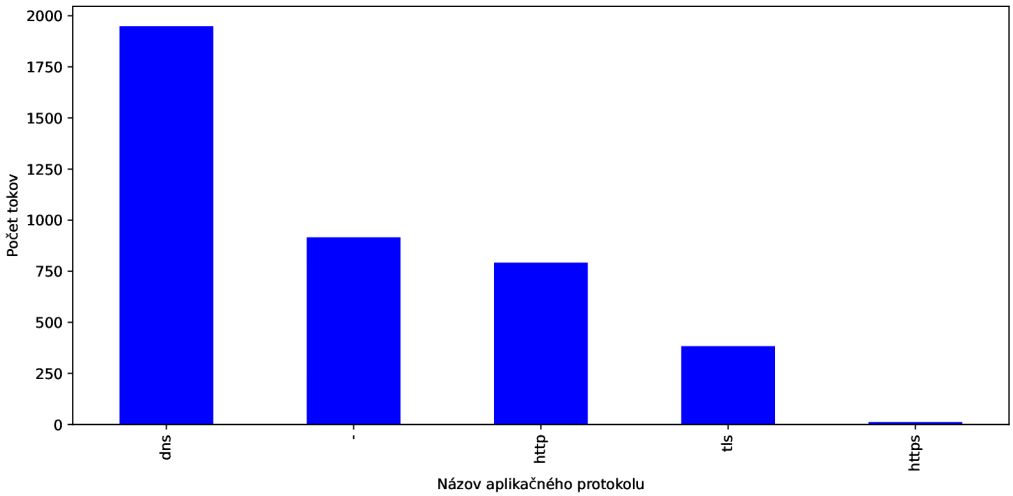
[6] https://code.visualstudio.com/docs/datascience/jupyter-notebooks
[7] https://en.wikipedia.org/wiki/Comma-separated_values
[8] https://developer.mozilla.org/en-US/docs/Web/XML/XML_introduction

zastúpenie majú v rámci dátovej sady jednotlivé rodiny je vhodné to zobraziť v podobe grafu 3.1.



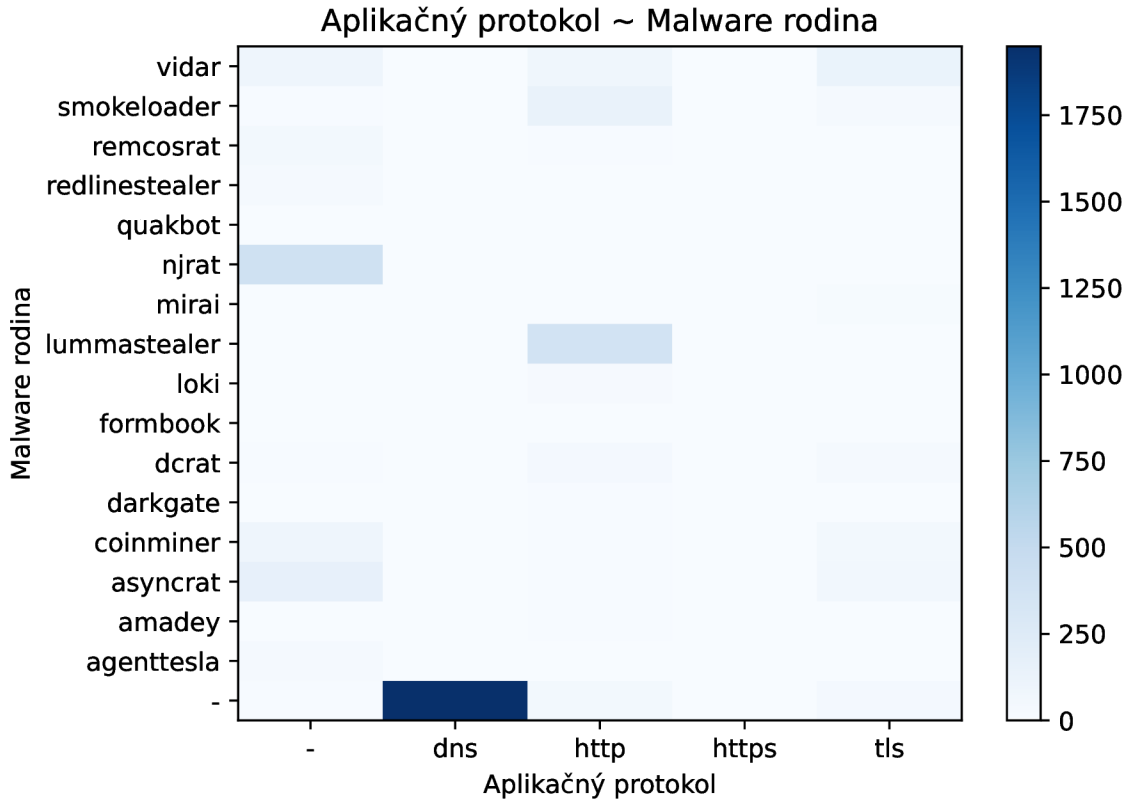
Obr. 3.1: Graf zobrazujúci počet sieťových tokov malware rodín a aj normálnej komunikácie.

V grafe je možné vidieť aj toky, ktoré sú označené pomlčkou. Tieto toky reprezentujú normálnu komunikáciu, ktorá nástrojom pre vytváranie dátových sád nebola označená ako škodlivou. To z dôvodu toho, že toky, ktoré boli zachytené pre jednotlivé vzorky malwaru sa nevyskytovali v komunikácií zachytenej Triage (priečinok *network*). Do týchto tokov nástroj pre vytváranie dátových sad vložil aj všetky DNS dotazy. V nasledujúcom grafe 3.2 je vidieť aké aplikačné protokoly sú najviac používané v rámci celej dátovej sady.



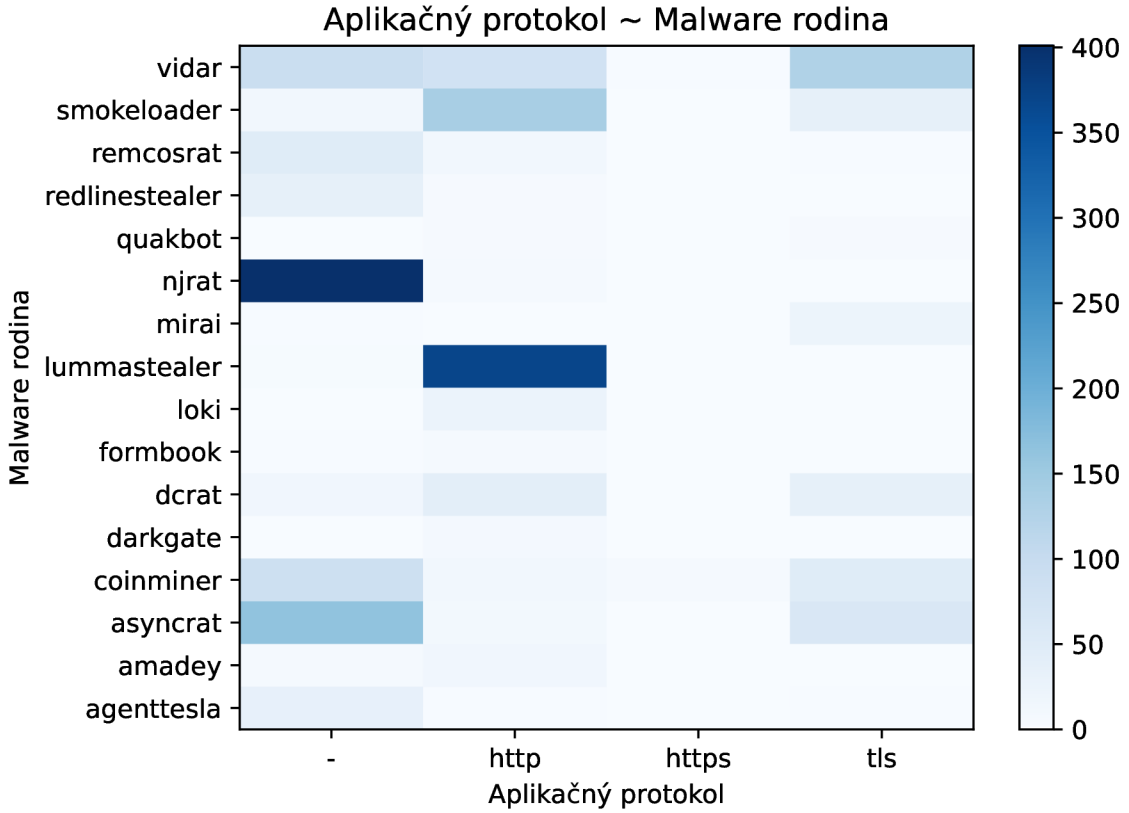
Obr. 3.2: Graf zobrazujúci výskyt aplikačných protokolov v rámci dátovej sady, za povšimnutie stojí tiež fakt že nemalá časť dát nemá definovaný aplikačný protokol z dôvodu toho, že nástroj pre vytváranie dátových sád pracuje len s DNS, HTTP, HTTPS a TLS protokolmi.

Avšak bolo by dobré vidieť vzťah medzi aplikačnými protokolmi a komunikáciou (normálnou aj malware) pre uvedomenie si toho, čo sa oplatí nechať a čo naopak treba odfiltrovať. Na tento účel poslúži kontingenčná tabuľka, čo je prakticky tabuľkové zobrazenie viacrozmerného rozdelenia početností konkrétnych premenných. Jazyk python v tomto prípade využíva teplotnú mapu (anglicky heatmap), ktorá prostredníctvom farieb a ich sýtosti informuje o početnosti jednotlivých položiek [16]. Túto tabuľku je možné vidieť na obrázku 3.3. Veľká časť komunikácie je sústredená pri DNS a nedefinovanej malware rodine, pravdepodobne sa jedná o normálnu komunikáciu, preto by bolo dobré sa jej zbaviť z dôvodov uvedených vyššie v tejto podkapitole.



Obr. 3.3: Kontingenčná tabuľka teplotnej mapy zobrazujúca vzťah medzi aplikačnými protokolmi a komunikáciou (normálnou aj malware).

Na to poslúži obrázok 3.4, kde je už vidieť, že po odfiltrovaní normálnej komunikácie je jasnejšie poznať aký malware využíval ktorý aplikačný protokol. Týmto sa zredukoval nepotrebný počet záznamov v dátovej sade.



Obr. 3.4: Kontingenčná tabuľka teplotnej mapy zobrazujúca vzťah medzi aplikačnými protokolmi a vyfiltrovanými malware rodinami.

3.2 Extrakcia indikátorov kompromitácie malware rodín dátovej sady a overenie ich relevantnosti

Analýzou dát dátovej sady bolo zistené, že malware rodiny *njrat* a *lummastealer* vykazujú nadpriemerne veľkú komunikáciu na sieti (viď. obrázok 3.4), čo bude znamenať samozrejme viac indikátorov kompromitácie (IoC) pre spomenuté rodiny. Je jasné, že aj u ostatných malware rodín boli v rámci možností získané tieto indikátory avšak nie v tak veľkom počte.

Na to aby sa bolo možné dostať k indikátorom kompromitácie bolo potrebné si ich vytiahnuť z jednotlivých hlásení (anglicky reports) pre dané malware rodiny. Hlásenia boli vo formáte JSON. Daná štruktúra je zobrazená na výpise 3.1. Tento úryvok výpisu je konkrétne z hlásenia pre malware vzorku z rodiny *njrat*. Na danom výpise je možné vidieť všetky tri typy IoC (viď. podkapitola 2.4) a to konkrétne URL adresy, domény a IP adresy, ktoré malware použil počas dynamickej analýzy v sandbox prostredí Triage. Samozrejme, nie všetky zachytené IoC musia nutne obsahovať škodlivé URL adresy domény alebo IP adresy. Niektoré z nich môžu patriť normálnej komunikácií. Napr vo výpise 3.1 je možné vidieť IoC, ktoré obsahuje skupinu IP adries medzi ktorými je IP adresa Google DNS serveru. To, že nejaká aplikácia komunikuje s DNS serverom, a k tomu ešte verejným serverom DNS, nutne neznamená, že sa jedná o malware respektíve inak škodlivú komunikáciu.

```
    "iocs": {
            "urls": [
            "https://tse1.mm.bing.net/th?id=OADD2.10239317301500
                ↪ _1UAMZFMFEP1QV3EDL&pid=21.2&w=1080&h=1920&c=4"
            ],
            "domains": [
            "146.78.124.51.in-addr.arpa",
            "67.254.221.88.in-addr.arpa",
            "tse1.mm.bing.net",
            "50.23.12.20.in-addr.arpa",
            "95.221.229.192.in-addr.arpa",
            "198.187.3.20.in-addr.arpa",
            "73.159.190.20.in-addr.arpa",
            "218.240.110.104.in-addr.arpa",
            "208.194.73.20.in-addr.arpa",
            "23.236.111.52.in-addr.arpa",
            "240.221.184.93.in-addr.arpa",
            "88.16.208.104.in-addr.arpa"
            ],
            "ips": [
            "8.8.8.8",
            "204.79.197.200",
            "52.111.229.43"
            ]
    }
```

Výpis 3.1: Štruktúra zachytených IoC malwaru vzorku z rodiny *njrat*.

Overenie relevantnosti získaných indikátorov kompromitácie

Na to aby bolo možné pracovať so spoľahlivými dátami je nutné použiť ešte nejaký iný referenčný zdroj informácií na základe ktorého sa bude môcť presnejšie rozhodnúť, či je dané IoC relevantné, to znamená či je považované za škodlivé alebo nie. Samozrejme 100 % istotu, že dané IoC je skutočne škodlivé sa dosiahnuť nedá, avšak týmto spôsobom sa aspoň zníži počet falošných poplachov pri detekcií malwaru. Práve na tento účel existuje mnoho nástrojov ako napríklad ThreatStop a ich Check IoC[9], nástroje od Abuse.ch[10] a to URLhaus[11], ThreatFox[12] alebo nástroj, ktorý bol použitý v rámci tejto práce a to AlienVault[13]. Ďalšie podobné nástroje mimo vrátane spomenutých sú prehľadne zobrazené v tabuľke 3.2. Mnohé z týchto nástrojov poskytujú viac než len overovanie IoC. Napríklad len taký ThreatStop je komplexný cloudový automatizovaný nástroj pre sledovanie hrozieb. Tento nástroj premieňa najnovšie údaje o hrozbách do politík vynucovania (anglicky enforcement policies) na základe ktorých, je následne automaticky aktualizovaný firewall, servery DNS, smerovače

[9] https://www.threatstop.com/check-ioc
[10] https://abuse.ch
[11] https://urlhaus.abuse.ch/
[12] https://threatfox.abuse.ch/
[13] https://otx.alienvault.com/browse/global/pulses?include_inactive=0&sort=-modified&page=1&limit=10

a koncové stanice s jediným cieľom, a to zastaviť kybernetické útoky než sa z nich stanú prieniky (anglicky breaches). Tento nástroj rieši problémy s bezpečnosťou jak prichádzajúcej tak odchádzajúcej sieťovej prevádzky prostredníctvom automatizovaného spravodajstva o hrozbách založeného na cloude a ochranného systému DNS [9].

Služba	**Typ informácií**	**API dostupné**	**Platené/Otvorené**
AlienVault OTX	IP IoC, IoC, URL IoC	Áno	Otvorené
ThreatFox	IP IoC, Domain IoC, URL IoC	Áno	Otvorené
VirusTotal	IP IoC, URL IoC	Áno	Platené/Otvorené
IBM X-Force	IP IoC, URL IoC	Áno	Platené
URLhaus	URL IoC	Áno	Otvorené
GreyNoise	IP IoC	Áno	Platené/Otvorené
ThreatSTOP IoC-Checker	IP IoC, Domain IoC, URL IoC	Áno	Platené

Tabuľka 3.2: Prehľad služieb poskytujúcich dodatočné informácie o IoC.

V rámci tejto práce bude využitá možnosť OTX DirectConnect API spomenutého nástroja AlienVault pre automatické overenie IoC, ktoré boli získané z dátovej sady. AlienVault je open-source nástroj pre management bezpečnostných udalostí a informácií (anglicky SIEM - Security Information and Event Management) s funkciami ktoré zahrňujú zber, normalizáciu a koreláciu udalostí, mimo toho táto platforma poskytuje aj správu protokolov detekciu narušenia bezpečnosti (anglicky intrusion detection), hlásenie incidentov alebo iných bezpečnostných anomálií [1]. Platforma OTX (Open Threat Exchange) umožňuje organizáciám prípadne jednotlivcom skúmať jednotlivé hrozby, porovnávať výsledné údaje medzi sebou a integrovať informácie o hrozbách do svojich bezpečnostných systémov [12]. OTX sa skladá z dvoch hlavných komponentov:

- **Pulzy** reprezentujú zbierky indikátorov kompromitácie (IoC) nahlásených komunitou platformy OTX. Tieto nahlásené IoC sú následne preskúmané ostatnými členmi komunity. Pulzy poskytujú prehľadné zhrnutie hrozby v ktorom je možné nájsť informácie o tom aké typy softvéru bývajú terčom útoku, ďalšie súvisiace IoC nahlásené inými užívateľmi platformy [8].
- **Reputácia IP** poskytuje oznámenia o škodlivej komunikácií medzi známymi hostiteľmi, u ktorých sa vie, že sú infikovaný a zariadeniami, ktoré používa užívateľ respektíve organizácia [8].

Ako už bolo spomenuté, komunita v rámci platformy OTX zdieľa medzi sebou informácie o hrozbách vo forme tzv. pulzov. Tieto pulzy pozostávajú minimálne z jedného alebo viacerých IoC. IoC je teda časť artefaktu alebo artefakt samotný pozorovaný v rámci siete alebo v koncovom bode, o ktorom sa s vysokou mierou spoľahlivosti usudzuje, že je vektorom hrozby. Medzi takéto vektory hrozby môžu patriť napríklad rôzne zariadenia, ktoré útočník využil alebo infraštruktúra ktorá bola použitá útočníkom na vykonanie útoku. Tabuľka 3.3 obsahuje zoznam typov IoC, ktoré platforma OTX rozlišuje. Čo sa reputácií IP adries týka, tak táto komponenta OTX identifikuje nielen IP adresy ale aj domény z celého sveta, ktoré boli pridané komunitou OTX. Táto komponenta ich označuje ako potencionálne škodlivými alebo podozrivými, kým sa nezískajú ďalšie, aktuálnejšie údaje, ktoré zvýšia ich hodnotenie v rebríčku hrozieb. Táto komponenta IP reputácie dopĺňa informácie OTX o cenné dáta o aktívnych, alebo potencionálne škodlivých aktivitách objavujúcich sa po celom svete, ktoré sú viazané na domény alebo IP adresy [8].

Typ IoC	**Popis**
CIDR	Beztriedne medzidoménové smerovanie. Špecifikuje rozsah IP adries v sieti, ktorý je podozrivý z malware aktivít alebo útoku.
CVE	Identifikácia spoločných zraniteľností a rizík (anglicky Common Vulnerabilities and Exposures).
doména	Názov domény pre webovú stránku alebo server podozrivý z hosťovania alebo zapojenia do malware aktivít. Domény môžu tiež zahŕňať rad hostiteľských názvov.
email	Emailová adresa spojená s aktivitami malwaru.
FileHash (MD5, SHA1, SHA256, PEHASH)	Výpočet hash pre súbor, na základe ktorého sa vie určiť, či bol obsah súboru pozmenený alebo poškodený.
filepath	Jedinečná poloha v súborovom systéme zdroja podozrivého z malware aktivít.
hostname	Hostiteľský názov serveru umiestneného v rámci domény podozrivej z malware aktivít.
IPv4/IPv6	IP adresa používaná ako zdrojová/cieľová pre server alebo iné zariadenie podozrivé z malware aktivít.
Mutex	Objekt vzájomného vylúčenia umožňujúci viacerým programovým vláknam zdieľať rovnaký zdroj. Zámky (anglicky Mutex) sú často používané malwarom ako mechanizmus na detekciu, či systém už bol nakazený.
FileHash-SHA256	SHA256-formát hash, ktorý sumarizuje architektúru a obsah súboru považovaného za podozrivý.
URI	Jednotný identifikátor zdroja (anglicky URI - Uniform resource identifier), ktorý popisuje explicitnú cestu k súboru dostupného online, ktorý je podozrivý z malware aktivít.
URL	Jednotné umiestnenie zdroja (anglicky URL - Uniform resource locations), ktoré sumarizuje online umiestnenie súboru alebo zdroja spojeného s malware aktivitou.

Tabuľka 3.3: Prehľad typov indikátorov kompromitácie (IoC) identifikované platformou OTX. Prevzaté z [8].

V tejto práci je už spomenutý AlienVault OTX a jeho API DirectConnect použitý pre overenie relevantnosti získaných IoC. Výsledkom dotazu na API pre IP adresu, doménu a URL je JSON výpis s niekoľkými sekciami, tieto výpisy sú nespracované a obsahujú mnoho dát, ktoré sú pre túto prácu nepodstatnými. U niektorých IoC nemusia byť prítomné všetky sekcie ak nie sú k dispozícií relevantné dáta. Výpisy sa skladajú z viacerých sekcií, ktorých význam je následovný [3]:

- **IP**
  - general - všeobecné informácie o IPv4 adrese a iné sekcie ktoré sú momentálne k dispozícií pre zadanú IPv4 adresu.
  - reputation - dáta OTX o malware aktivitách zaznamenaných laboratóriami AlienVault (komponenta Reputácia IP).
  - geo - podrobnejší zoznam geografických údajov (kód krajiny, mesto apod.).
  - malware - vzorky malwaru, ktoré boli analyzované laboratóriami AlienVault a u ktorých bolo zistené, že sa pripájajú na danú adresu IP.
  - url_list - zoznam URL, ktoré súvisia s danou IP adresou.
  - passive_dns - informácie o pasívnych DNS záznamoch (hostiteľské mená, domény) u ktorých bolo zistené, že súvisia so zadanou IP adresou.
  - http_scans - meta-dáta pre http(s) pripojenia pre túto IP adresu.
- **Domény**
  - general - všeobecné informácie o doméne a iné sekcie ktoré sú momentálne k dispozícií pre zadanú doménu adresu.
  - geo - podrobnejší zoznam geografických údajov (kód krajiny, mesto apod.).
  - malware - vzorky malwaru, ktoré boli analyzované laboratóriami AlienVault a u ktorých bolo zistené, že sa pripájajú na danú adresu doménu.
  - url_list - zoznam URL, ktoré súvisia s danou doménou.
  - passive_dns - informácie o pasívnych DNS záznamoch (hostiteľské mená, domény) u ktorých bolo zistené, že súvisia so zadanou doménou.
  - whois - záznamy Whois pre zadanú doménu. item http_scans - meta-dáta pre http(s) pripojenia pre túto doménu.
- **URL**
  - general - geografické informácie ukazujúce do histórie a iné sekcie ktoré sú momentálne k dispozícií pre zadanú URL.
  - url_list - kompletné informácie z analýz zadanej URL z laboratórií AlienVault.

Spracované záznamy je možné vidieť vo výpise 3.2 kde tieto výpisy sú vo formáte ktorý zodpovedá:

- **IP záznam**
  - IPv4 adresa: počet súvisiacich pulzov - pasívne DNS záznamy
- **Záznam domény**
  - Názov domény: počet súvisiacich pulzov
- **URL záznam**
  - URL adresa:počet súvisiacich pulzov

```
------------IP IOC------------
148.72.177.212: 1 - [
        'bomnegocio.jogarxadrez.net', 'update.textbin.net',
        'www.orkuty.net', 'www.bomnegocio.net', 'bomnegocio.net',
        'orkuty.net', 'www.criarenquete.com', 'criarenquete.com',
        'www.freelas.net', 'freelas.net',
        'www.chat.batepapoonline.net', 'chat.batepapoonline.net'
]
18.228.115.60: 5 - [
        'campaign-ui.guilherme.cityshoppe.sa.ngrok.io',
        'backend.scopo.online',
        'free.ratchet-bodegadigitaldigital.app'
]
18.229.146.63: 3 - [
        'imperiossword.quest',
        'campaign.guilherme.cityshoppe.sa.ngrok.io',
        'bitbucket-entrypoint.local.playing-possum.com'
]
35.157.111.131: 4 - [
'7.tcp.eu.ngrok.kiyoon.kim', '7.tcp.eu.ngrok.io'
]

----------Domain IOC----------
ip-api.com: 41

-----------URL IOC------------
https://pt.textbin.net/download/rcd5ihynxw: 5
http://apps.identrust.com/roots/dstrootcax3.p7c: 50
https://api.ipify.org/: 6
http://ip-api.com/json: 16
}
```

Výpis 3.2: Vyfiltrované relevantné IoC malware rodiny *njrat*

Počet pulzov sa nachádza v objekte nazvanom ako *pulse_info*. Je to objekt, ktorý obsahuje informácie o pulzoch, ktoré súvisia so zadaným IoC. Pulz je zbierka IoC nahlásených komunitou používateľov platformy OTX. Tento objekt zvyčajne obsahuje tieto kľúčové prvky:

- **count** - Počet pulzov spojených so zadaným IoC.
- **pulses** - Pole objektov, pričom každý objekt reprezentuje jeden pulz. Každý pulz môže obsahovať ďalšie informácie, ako sú id, name (názov pulzu), description (popis pulzu), created (dátum a čas vytvorenia pulzu), modified (dátum a čas poslednej úpravy pulzu), tags (značky spojené s pulzom), a ďalšie dáta ak sú k dispozícii.

Tieto informácie budú následne použité v ďalšom kroku pri návrhu metódy pre detekciu malware komunikácie v rámci zachytenej sieťovej komunikácie nástrojom *Suricata*[14].

[14] https://suricata.io/

Kapitola 4

Metóda identifikácie malwaru v sieťovej komunikácií

Nasledujúca kapitola sa zaoberá metódou pre detekciu malwaru v rámci sieťovej komunikácie. Ako prvé je v danej kapitole riešený problém toho, ako získané údaje z dátovej sady (viď. podkapitola 3) vhodne spracovať a použiť pri detekcií. Následne sa rieši aj ako je využitý, už spomínaný AlineVault OTX (viď. sekcia 3.2) pri vytváraní modelov. Modelom je v tomto prípade označovaná fuzzy množina, ktorá reprezentuje skóre IoC (IP adresy, domény, URL adresy) jednotlivých malware rodín. V tejto kapitole sa tiež rieši ako je dané skóre (anglicky možno nazvať ako membership value) vypočítané a prečo metóda používa šesť takýchto množín, dve pre IP IoC, dve pre doménové IoC a dve pre URL IoC.

Ďalej sa bude riešiť ako pracuje daná metóda na detekciu malwaru v sieťovej komunikácií, ako sa počítajú prahové hodnoty (anglicky treshold values) malware rodín. Tiež bude spomenuté v akých dvoch módoch daný nástroj pracuje a čo majú tieto dva módy spoločné a čo a aké sú medzi nimi odlišnosti.

4.1 Modely indikátorov kompromitácie

Ako už bolo spomenuté, tieto modely reprezentujú fuzzy množiny, ktoré obsahujú skóre jednotlivých IoC, tým sa myslí IP, domén a URL adries jednotlivých malware rodín. Skóre týchto IoC reprezentuje mieru príslušnosti do akej sú tieto IoC spojené s konkrétnou malware rodinou. Dáta spomínaných IoC boli získané z JSON hlásení (anglicky reports) prostredníctvom už spomenutého nástroja vytvoreného v rámci bakalárskej práce Detekce komunikace malware v síťových tocích [17]. Štruktúru jedného z hlásení je možné vidieť vo výpise 3.1.

Každý typ IoC (IP, doména, URL) je reprezentovaný dvoma modelmi (ďalej ako fuzzy množiny). Dvomi množinami z dôvodu toho aby sa vedelo presnejšie rozhodnúť o tom či je dané IoC relevantným alebo nie. To znamená, že síce dané IoC získa v prvej fuzzy množine isté skóre, ktorého výpočet je vysvetlený v nasledujúcom odseku, ale skóre z AlienVault OTX môže byť nižším prípadne rovným 0 a tým pádom pri detekcií daný indikátor nebude mať takú váhu respektíve vážnosť ako indikátory ktoré získali omnoho väčšie skóre.

Výpočet hodnôt skóre jednotlivých IoC v prvej množine je počítané ako počet výskytov IoC v danej rodine vydelené celkovým počtom všetkých IoC v rodine, tým získame skóre priemerného výskytu daného indikátoru v rámci jednej rodiny.

V druhej fuzzy množine je skóre jednotlivých IoC počítané inak, a to za využitia AlienVault OTX, z ktorého sa vyberie počet pulzov, čo predtsavuje zbierku IoC nahlásených komunitou platformy AlienVault, a spočíta sa počet pasívnych DNS záznamov (hostiteľské mená, domény), ktoré sa viažu k danému indikátoru, viď. pseudokód 1. Následne sa to podelí hodnotu 2 čím sa získa surová (anglicky raw) hodnota pre daný indikátor. Avšak je potrebné aby hodnoty v týchto fuzzy množinách boli normalizované nakoľko sa nachádzajú v rôznych rozmedziach. Normalizácia číselných hodnôt je proces pri ktorom dochádza ku škálovaniu hodnôt do známeho rozsahu. V prípade tejto práce sa jedná o rozsah 0-1. Škálovanie je vykonané prostredníctvom metódy *Min-Max*, ktorej vzorec vypadá následovne $X_{norm} = (X - X_{min})/(X_{max} - X_{min})$.

Algorithm 1 Pseudokód popisujúci výpočet skóre IoC za pomoci hodnôt pulzov z AlienVault OTX.

1: $pulse\_count = get\_pulses\_OTX(ioc)$
2: $passive\_dns\_count = get\_passive\_DNS\_records(ioc)$
3: $raw\_membership\_val = (pulse\_count + passive\_dns\_count)/2.0$
4: $max\_val = get\_max\_raw\_memebership$
5: $min\_val = get\_min\_raw\_memebership$
6: $normalized\_val = (raw\_membership\_val - min\_val)/(max\_val - min\_val)$

Po tom ako sú hodnoty správne normalizované sú zapísané do fuzzy množiny s ktorou sa ďalej pracuje pri detekcií. Časť tejto fuzzy množiny je možné vidieť vo výpise 4.1. Je možné vidieť, že niektoré IP adresy dosiahli pomerne vysoké skóre výskytu 1.0, čo znamená, že daný IoC, v tomto prípade IP adresa, má vysoký počet pulzov prípadne pasívnych DNS záznamov čo len pomáha pri spresnení detekcie malwaru v sieťovej komunikácií.

Dáta, ktoré sú získané prostredníctvom AlienVault OTX API sú uložené do JSON súborov kde každý typ IoC má svoj vlastný súbor. Tým sa efektívne predišlo získavaniu tých istých dát pri opakovanom spúštaní nástroju. Ak by z nejakého dôvodu daný JSON súbor (ďalej ako medzipamäť (anglicky cache)) neobsahoval dáta o všetkých IoC, nástroj pre detekciu malwaru sa ich pokúsi získať opätovným dotazovaním sa na AlienVault OTX API. Vďaka tomu sa efektívne znížila čakacia doba potrebná pre vytvorenie daných fuzzy množín.

```
amadey: {
        '138.91.171.81': 0.007, '13.107.253.64': 0.016,
        '204.79.197.200': 1.0, '8.8.8.8': 0.909,
        '216.58.213.10': 0.024, '142.250.187.234': 0.064
}
avaddon: {
        '8.8.8.8': 0.909, '2.17.5.133': 0.002,
        '172.217.16.234': 0.078
}
backdoor.teamviewer: {
        '8.8.8.8': 0.909, '94.156.8.42': 0.004,
        '216.58.201.97': 0.042
}
```

Výpis 4.1: Úryvok IP fuzzy množiny vytvorenej za pomoci AlienVault OTX nástroja.

Ako je možné vidieť, vo výpise jednej z fuzzy množín 4.1 sa nachádza aj IP adresa ako 8.8.8.8, čo je IP adresa verejne dostupného DNS serveru firmy Google. Túto IP adresu majú takmer všetky vzorky každej malware rodiny, výnimkou sú vzorky kde sa nezachytila žiadna komunikácia. To, že sa nejaký počítač, prípadne iné zariadenie pripojené k internetu, pripája na túto IP adresu zvyčajne nemusí naznačovať škodlivú komunikáciu. Väčšina zariadení na internete komunikuje s týmito adresami z legitímnych dôvodov, ako napríklad rezolúcia DNS dotazov, prístup k službám veľkých spoločností ako Google, Microsoft a pod. Avšak môžu existovať prípady, kedy má zahrnutie aj takýchto IP adries do detekcie malwaru v sieťovej komunikácií svoje opodstatnenie a zmysel. Ignorovaním týchto typov verejne známych IP adries by mohlo dôjst k strate časti informácie, nakoľko mnoho z týchto vzoriek jednotlivých malware rodín vykonávajú napríklad už spomínanú DNS rezolúciu aby vedeli naviazať spojenie so serverom útočníka prípadne aby mohli odosielať ukradnuté dáta od napadnutého užívateľa k útočníkovi. Z tohto dôvodu sú verejne známe IP adresy, typu 8.8.8.8 ponechané vo fuzzy množinách u jednotlivých malware rodín. Príkladom toho, že tieto IP adresy môžu, i keď nepriamo, participovať v rámci škodlivej komunikácií je ukázaná vo výpise 4.2. V danom výpise je vidieť, že prostredníctvom tejto IP adresy došlo k požiadavku na prevod uvedených domén na IP adresy prostredníctvom A záznamov služby DNS.

```
{"timestamp": "2024-03-02T20:18:20.463728+0100",
       "flow_id": 1147271881042410, "pcap_cnt": 3519, "event_type": "dns",
       "src_ip": "10.127.1.2", "src_port": 60646, "dest_ip": "8.8.8.8",
       "dest_port": 53, "proto": "UDP", "pkt_src": "wire/pcap",
       "dns": {"type": "query", "id": 3681,
              "rrname": "host-file-host6.com", "rrtype": "A",
              "tx_id": 0, "opcode": 0}, "malware": true
       }

% {"timestamp": "2024-03-02T20:20:03.659490+0100",
       "flow_id": 862166955978827, "pcap_cnt": 32804, "event_type": "dns",
       "src_ip": "10.127.1.2", "src_port": 51498, "dest_ip": "8.8.8.8",
       "dest_port": 53, "proto": "UDP", "pkt_src": "wire/pcap",
       "dns": {"type": "query", "id": 45716,
                     "rrname": "clients2.google.com", "rrtype": "A",
                     "tx_id": 0, "opcode": 0}, "malware": true
              }

{"timestamp": "2024-03-02T20:20:06.897032+0100",
       "flow_id": 1882398686084854, "pcap_cnt": 32921, "event_type": "dns",
       "src_ip": "10.127.1.2", "src_port": 52004, "dest_ip": "8.8.8.8",
       "dest_port": 53, "proto": "UDP", "pkt_src": "wire/pcap",
       "dns": {"type": "query", "id": 57460,
              "rrname": "api5.check-data.xyz", "rrtype": "A",
              "tx_id": 0, "opcode": 0}, "malware": true
       }
```

Výpis 4.2: Časť zachytenej komunikácie ktorá bola prečítaná nástrojom *Suricata* z *pcap* súborov malware vzorku rodiny *backdoor.teamviewer* z dynamickej analýzy vykonanou nástrojom Triage.

```
backdoor.teamviewer: {
        clients2.google.com: 0.011,
        api4.check-data.xyz: 0.011,
        host-file-host6.com: 0.011
}
backdoor.teamviewer: {
        host-file-host6.com: 0.239,
        api5.check-data.xyz: 0.021,
        clients2.google.com: 0.514
}
```

Výpis 4.3: Skóre domén z fuzzy množín ktoré sú spomenuté vo výpise 4.2.

4.2 Detekcia malwaru v sieťovej komunikácií

Ako už bolo spomenuté na začiatku tejto kapitoly táto metóda respektíve nástroj sa používa v spojení so *Suricatou*. Tiež je vhodné si ujasniť pojmy nástroj a metóda aby sa nezamienali respektíve aby bolo jasné čo znamenajú v kontexte tejto práce. Pojmom metóda sa myslí postup, ktorý bol zvolený pre detekciu malwaru v sieťovej komunikácií a nástroj je už konkrétne integrovanie danej metódy s ostatnými časťami tohoto nástroju napríklad s už spomenutými fuzzy množinami.

Je vhodné začať tým že tento nástroj je založený na spracovávaní zachytávaných sieťových tokov zo *Suricaty*, ktorá dané toky (anglicky flows) zapisuje do súboru vo formáte JSON v reálnom čase. Následne nástroj pre detekciu posudzuje jednotlivé toky a zisťuje či indikátory útoku (anglicky Indicators of Attack) (IP, domény URL adresy) ktoré boli v nich nájdené sú uvedené vo fuzzy množinách, ktoré obsahujú indikátory kompromitácie. A ak áno aké skóre získali, k akým rodinám patria a porovná či dané skóre je väčšie alebo rovné prahovej hodnote jednotlivých daných malware rodín. Tento režim je označený ako režim online.

Druhým režimom ktorý nástroj pre detekciu podporuje je režim offline. Tento režim je založený na tom, že daný nástroj pre detekciu sa môže použiť na analýzu už vopred existujúceho JSON súboru, ktorý je vytvorený opäť *Suricatou*, ktorá len prečíta už existujúce *PCAP* súbory zachytenej sieťovej komunikácie jednotlivých vzorkov malware rodín.

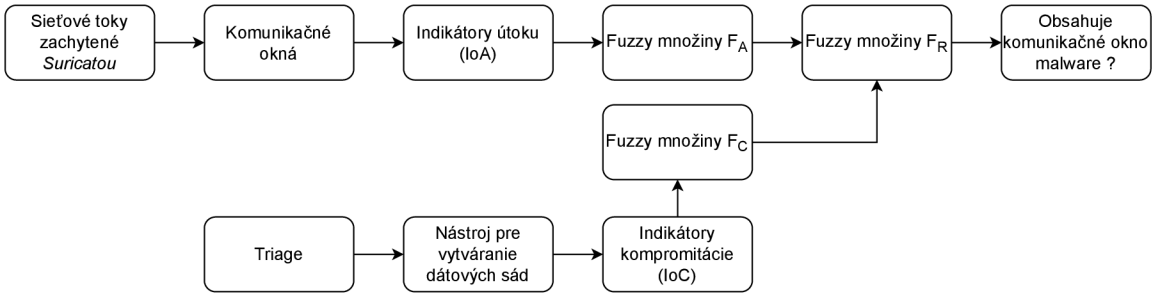
Tieto dva spomenuté módy sa líšia primárne v tom ako pristupujú k získaným dátam zo *Suricaty*. Ich následné spracovanie dát a ich použitie v rámci detekcie je rovnaké.

Metóda detekcie

Metóda ktorá je využitá v oboch spomínaných módoch detekčného nástroja využíva spomínané fuzzy množiny na analýzu údajov o sieťovej prevádzke a rozhodnutia či je daná komunikácia potencionálne škodlivá alebo nie. Po extrahovaní IoA z jednotlivých zachytených tokov sú vytvárané komunikačné okná. Pre každú zdrojovú IP adresu sa jednotlivé toky rozdelia do okien na základe časovej pečiatky (anglicky timestamp). Dôvodom prečo sú toky zoskupené do komunikačných okien je ten, že je žiadúce aby bolo známe akú komunikáciu vykonávali jednotlivé stanice, respektíve počítače, v jednotlivých časových úsekoch. Vďaka tomu je možné odhaliť kontext sieťových udalostí. To samo o sebe môže byť kľúčovým pre odhalenie vzorov (anglicky patterns) malware komunikácie alebo iných anomálií,

ktoré sa vyskytli v danom časovom úseku sieťovej komunikácie, ako napríklad prudký nárast prevádzky na podozrivú IP adresu alebo doménu, čo by mohlo znamenať, že daná stanica je pravdepodobne infikovaná malwarom. Ďalším dôvodom je aj fakt ze spracovávať toky generované nástrojom *Suricata* v reálnom čase je celkom dosť výpočtovo náročné oproti spracúvaniu komunikačných okien v prípade online módu.

Tieto komunikačné okná obsahujú už jednotlivé IoA (rozdiel medzi IoC a IoA viď. sekcia 2.4). Teraz je potrebné určiť ako moc sa zhodujú tieto IoAs s IoCs modelmi. Čiže IoA si označíme ako fuzzy množinu F_A. A fuzzy množiny jednotlivých modelov označíme ako F_C. Následne sa vykoná prienik medzi týmito dvomi fuzzy množinami $F_R = F_A \cap F_C$. Prienik medzi týmito množinami je vykonaný prostredníctvom funkcie *min*, ktorá porovnáva dve hodnoty a vráti menšiu z nich. Skóre jednotlivých IoA vo fuzzy množine F_A je rovné 1, to preto lebo tieto IoA sa vyskytli v zachytenej komunikácií, to znamená že, miera ich ich výskytu je istá čiže je rovná 1. Tým pádom vo výslednom prieniku F_R sa budú nachádzať len IoC danej malware rodiny, ktoré boli nájdené v rámci zachytenej komunikácie, respektíve zachyteného komunikačného okna. Popísaný princíp je možné vidieť na obrázku 4.1.



Obr. 4.1: Princíp fungovania detekčného algoritmu.

Ďalej je nutné zistiť či týchto zachytených IoC v prieniku F_R je dostatočné množstvo a či sú tieto IoC natoľko významnými, že dané okno bude možné prehlásiť za potencionálne škodlivé. Sčítaním skóre jednotlivých IoC konkrétnych malware rodín je možné síce získať nejaké kladné desatinné číslo avšak toto číslo samo o sebe nič moc nepovie.

Na to je nutné vypočítať tzv. prahovú hodnotu (anglicky treshold value) pre každú malware rodinu. Prahová hodnota je získaná ako suma skóre IoC, napríklad IP adries, ktoré boli nájdené v jednotlivých malware vzorkoch konkrétnych malware rodín. Prakticky sa dané modely použili na detekciu vzorkov malwaru z ktorých boli tieto modely vytvorené. Čiže sa postupne brali vzorky jednotlivých malware rodín ako komunikácia, ktorú je nutné analyzovať a prípadne v nej detegovať, škodlivú komunikáciu. Ak sa dané IoC z modelu vyskytuje v danom vzorku jeho dosiahnuté skóre je pripočítané k celkovej sume. A pre každé IoC vzniknú dve takéto sumy, nakoľko existujú pre každý typ IoC (Ip, domény, URL adresy) dva modely vytvorené každý iným spôsobom (viz. podkapitola 4.1). Nakoniec sa vykoná už len priemer týchto dvoch súm pre každú vzorku konkrétnej malware rodiny viď. výpis 4.4. Avšak je tu ešte problém toho, že jednotlivé malware vzorky sú od seba navzájom odlišné z pohľadu sieťovej komunikácie, niektoré vzorky sa vyznačujú bohatšou sieťovou komunikáciou iné zase nie. To sa samozrejme odrazí aj na množstve získateľných informácií potrebných pre danú detekciu. A preto je potrebné vytvoriť nejakú spoločnú charakteristiku pre vzorky konkrétnej malware rodiny. Toho sa dosiahne prostredníctvom priemeru a štandardnej odchýlky. Kde po spočítaní priemeru hodnôt napríklad rodiny *amadey* z vý-

pisu 4.4 a spočítaní štandardnej odchýlky sa vypočíta prahová hodnota pre danú rodinu ako $TRESHOLD_{FAMILY} = AVG - STDEV$.

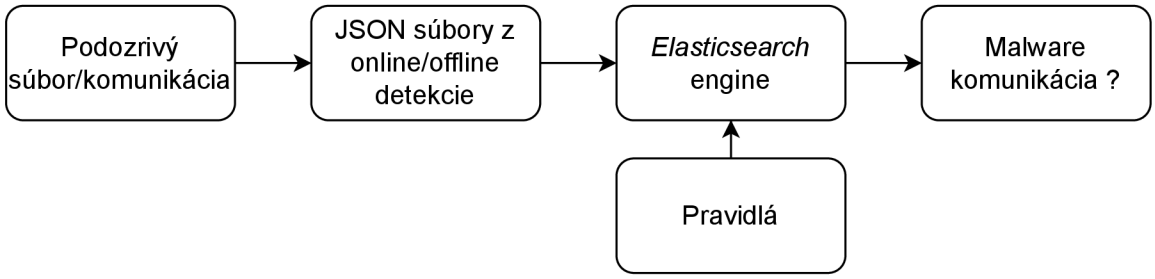
Prípadné lokálne extrémy (lokálne minimá, lokálne maximá) sú odstránené aby výsledná prahová hodnota nebola skreslená týmito odľahlými hodnotami. Na toto je použitá technika medzikvartilového rozpätia (anglicky interquartile range). Ktorej princíp je možné v skratke popísať následovne. Metóda vypočíta rozsah medzi prvým a tretím kvadrantom zadaných údajov a následne odstráni hodnoty, ktoré spadajú buď pod prvý kvadrant mínus určitý násobok vypočítaného medzikvartilového rozsahu alebo nad tretí kvadrant plus ten istý násobok rozsahu.

```
amadey: [1.285, 1.407, 1.33],
avaddon: [0.623, 0.892, 0.892, 0.873],
backdoor.teamviewer': [4.813, 1.082],
darkgate: [1.329, 1.313, 0.801, 1.327, 0.689, 1.313, 0.672, 0.672,
1.313, 1.38, 1.313, 1.313, 1.33, 0.83, 1.313, 1.641, 0.672],
loki: [1.329, 1.653, 0.623, 2.255, 1.728]
```

Výpis 4.4: Hodnoty spriemerovaných súm jednotlivých malware vzorkov konkrétnych rodín (každé číslo reprezentuje jeden vzorok).

4.3 Výsledky detekcie a možná integrácia do vhodného prostredia

Výsledky oboch týchto režimov sú zapísané do JSON súborov, viď. výpis 4.5. To z dôvodu ďalšieho spracovania týchto dát. Tieto dáta by mohli byť ďalej spracúvané napríklad analytickým nástrojom *Elasticsearch*[15]. *Elasticsearch* je platforma poskytujúca nielen analýzu dát a prácu s nimi ale poskytuje aj funkcionality ako SIEM (anglicky Security Event Management System). Na základe tohoto by vytvorený nástroj na detekciu malwaru v lokálnej sieti mohol byť integrovaný s týmto nástrojom. Výsledkom takejto integrácie by mohlo byť komplexné sandbox-ové prostredie, v ktorom by sa mohli spúšťať vzorky rôznych podozrivých súborov a sledovať ich chovanie. Prípadne by sa mohla vykonať analýza vopred zachytenej sieťovej komunikácie *Suricatou*. Dáta z JSON súborov by následne mohli byť poslané do *Elasticsearch*, ktorý by na osnove vopred definovaných pravidiel mohol upozorniť na potencionálne škodlivú komunikáciu, ktorú tieto súbory vykazujú. Táto myšlienka takéhoto nápadu je zobrazená na obrázku 4.2.



Obr. 4.2: Myšlienka integrácie vytvoreného nástroju na detekciu malwaru s *Elasticsearch*.

[15] https://www.elastic.co/

Skóre jednotlivých rodín v komunikačných oknách reprezentuje celkové skóre, ktoré daná rodina dosiahla v rámci istého časového úseku. Toto skóre je spočítané ako vážený priemer, kde váha pre skóre IP a domén je 1.0 a váha pre URL skóre je 0.5. Dôvodom prečo URL skóre má váhu 0.5 je ten že, URL IoC, ktoré boli získané v rámci vzorkov malwaru sa častokrát nezhodovali s nájdenými IoA v sieťovej komunikácií. Tiež je vhodné spomenúť, že IoC typu IP a domén sú v rámci tejto práce presnejšími indikátormi malware komunikácie nakoľko poskytujú najväčšie množstvo dát. Vo výpise si je tiež možné všimnúť, že niektoré komunikačné okná majú označenie *"malware": true*, čo označuje dané okno, že v ňom bola pravdepodobne detekovaná škodlivá komunikácia. Toto označenie dostávajú okná ak jedna z malware rodín má skóre vyššie ako 0.


```
{"timestamp": "2024-03-02T20:12:00_0", "pc": "10.127.0.70",
 "malware_families_scores": {"amadey": 0.704, "avaddon": 0.0,
       "backdoor.teamviewer": 0.0, "darkgate": 0.699,
       "dridex": 0.709, "formbook": 0.621, "gcleaner": 0.0,
       "healer": 0.0, "heodo": 0.727, "irata": 0.0, "loki": 0.0,
       "mirai": 0.0, "pikabot": 0.0, "quakbot": 0.683,
       "redlinestealer": 0.0, "remcosrat": 0.592,
       "riseprostealer": 0.477, "stealc": 0.0, "tofsee": 0.0,
       "trickbot": 0.695, "vidar": 0.578}, "malware": true}

{"timestamp": "2024-03-02T20:12:00_0", "pc": "10.127.0.66",
 "malware_families_scores": {"amadey": 0.665, "avaddon": 0.0,
       "backdoor.teamviewer": 0.0, "darkgate": 0.675,
       "dridex": 0.669, "formbook": 0.602, "gcleaner": 0.0,
       "healer": 0.0, "heodo": 0.727, "irata": 0.0, "loki": 0.0,
       "mirai": 0.0, "pikabot": 0.0, "quakbot": 0.641,
       "redlinestealer": 0.0, "remcosrat": 0.595,
       "riseprostealer": 0.477, "stealc": 0.0, "tofsee": 0.0,
       "trickbot": 0.66, "vidar": 0.565}, "malware": true}

{"timestamp": "2024-03-02T20:17:00_0", "pc": "88.221.135.217",
 "malware_families_scores": {"amadey": 0.0, "avaddon": 0.0,
       "backdoor.teamviewer": 0.0, "darkgate": 0.0,
       "dridex": 0.0, "formbook": 0.0, "gcleaner": 0.0,
       "healer": 0.0, "heodo": 0.0, "irata": 0.0, "loki": 0.0,
       "mirai": 0.0, "pikabot": 0.0, "quakbot": 0.0,
       "redlinestealer": 0.0, "remcosrat": 0.0,
       "riseprostealer": 0.0, "stealc": 0.0, "tofsee": 0.0,
       "trickbot": 0.0, "vidar": 0.0}}
```


Výpis 4.5: Výpis nástroju po vykonaní detekcie. Časová pečiatka (anglicky timestamp) definuje kedy bolo dané časové okno zachytené.

Kapitola 5

Experimenty a ich vyhodnotenie

Táto kapitola sa zaoberá experimentovaním a vyhodnotením dosiahnutých výsledkov, ktoré metóda pre detekciu malwaru opísaná v kapitole 4 dosiahla. Na začiatok je popísané s akými dátovými sadami bol nástroj otestovaný a aké metriky boli použité na jeho vyhodnotenie.

V ďalšej podkapitole 5.2 je prezentovaný výsledok detekcie už existujúcej komunikácie, kedy je nástroj spustený v režime offline. Ďalej je popísaný rozdiel medzi dvoma spôsobmi počítania prahovej hodnoty (anglicky treshold value) s ktorými bolo experimentované. Prvý spôsob, ktorý je aj použitý vo výslednej metóde, využíva medzikvartilové rozpätie na zbavenie sa lokálnych extrémov (lokálne minimá, maximá) a následne vypočítava priemer a štandardnú odchýlku. V druhom spôsobe sa medzikvartilové rozpätie nepoužíva a rovno sa počíta priemer a štandardná odchýlka. Dátová sada využitá na analýzu v kapitole 3 nebola použitá z dôvodu chyby v nástroji, ktorý bol použitý na vytvorenie danej dátovej sady. Problém spočíval v tom, že nástroj nezískal všetky *pcap* súbory, z ktorých by sa za pomoci *Suricaty* získali sieťové toky, ktoré by mohli byť analyzované.

V podkapitole 5.3 sa prezentujú výsledky dosiahnuté v reálnej prevádzke s detekčným nástrojom spusteným v online režime. A následne v závere je celkové zhrnutie výsledkov.

5.1 Testovacie dáta

Pre otestovanie funkčnosti detekčnej metódy bolo použitých hneď niekoľko dátových sád, ktoré boli vytvorené za pomoci už spomenutého nástroju [17] viď. podkapitola 3.1. Ďalej *pcap* súbory jednotlivých dátových sád boli prečítané nástrojom *Suricata*, pre vytvorenie JSON súboru (*eve.json*), ktorý bude obsahovať všetky zachytené komunikačné toky, ktoré charakterizujú komunikáciu vzorkov.

Avšak je nutné získať referenčné hodnoty voči ktorým budú môcť byť výsledky porovnané a následne vyhodnotené. Treba poznamenať, že nie všetky sieťové toky, ktoré sú v *pcap* súboroch sú nutne škodlivými. Na to aby sa získali referenčné hodnoty poslúžili sieťové hlásenia (anglicky network reports) poskytnutých nástrojom Triage z dynamickej analýzy. Sieťové hlásenia obsahujú okrem zachytených obojsmerných sieťových tokov aj spúšťané procesy a stavy registrov operačného systému. Tieto hlásenia sú vygenerované pre každý vzorok každej malware rodiny. Ich formát je možné vidieť vo výpise 5.1. Spomenutý skript v jazyku *python* vyhľadáva či toky zaznamenané v hláseniach sieťovej komunikácie vzorkov malwaru sa nachádzajú aj vo výslednom *eve.json*. Ak daný tok zo sieťového hlásenia nástroju Triage bol objavený v *eve.json* dostal označenie *"malware": true* a tým pádom je anotovaný ako potencionálne škodlivým. Ostatné toky, ktoré neboli nájdené v sieťových

hláseniach neboli v *eve.json* nijak označené ani menené. Takto sa dalo dostať k referenčnej hodnote sieťových tokov, ktoré sú potencionálne škodlivými a oddeliť ich od normálnych sieťových tokov.

```
    {
            "id": 24,
            "src": "10.127.0.66:49808",
            "dst": "13.107.253.64:443",
            "proto": "tcp",
            "pid": 4116,
            "procid": 81,
            "first_seen": 36735,
            "last_seen": 36794,
            "rx_bytes": 40,
            "rx_packets": 1,
            "tx_bytes": 46,
            "tx_packets": 1
    },
```

Výpis 5.1: Príklad zachyteného toku zo sieťového hlásenia vygenerovaného nástrojom Triage.

Konkrétne boli použité dve dátové sady na experimenty a na vytvorenie modelov IoC respektíve fuzzy množín. Všetky tri dátové sady majú dohromady 89 314 tokov. Z toho 41.16% bolo označených za škodlivé a zbytok 58.84% je braný ako normálna komunikácia viď. tabuľka 5.1.

Typ toku	**Dátová sada č. 1**	**Dátová sada č. 2**
Malware	20 973	15 786
Normálne	26 262	26 293
Malware + Normal	47 235	42 079

Tabuľka 5.1: Tabuľka zobrazujúca prehľad dátových sád použitých pri experimentoch.

Dátová sada č. 1 pozostáva z 12 malware rodín kde každá rodina pozostáva z 18 vzorkov, dátová sada č. 2 obsahuje 21 malware rodín počet vzorkov pre jednotlivé malware rodiny sa pohybuje v rozsahu od 1 vzorku až 19 vzorkov. Dôvodom takejto fluktuácie je pravdepodobne chyba v nástroji, ktorý bol použitý pre vytváranie dátových sád.

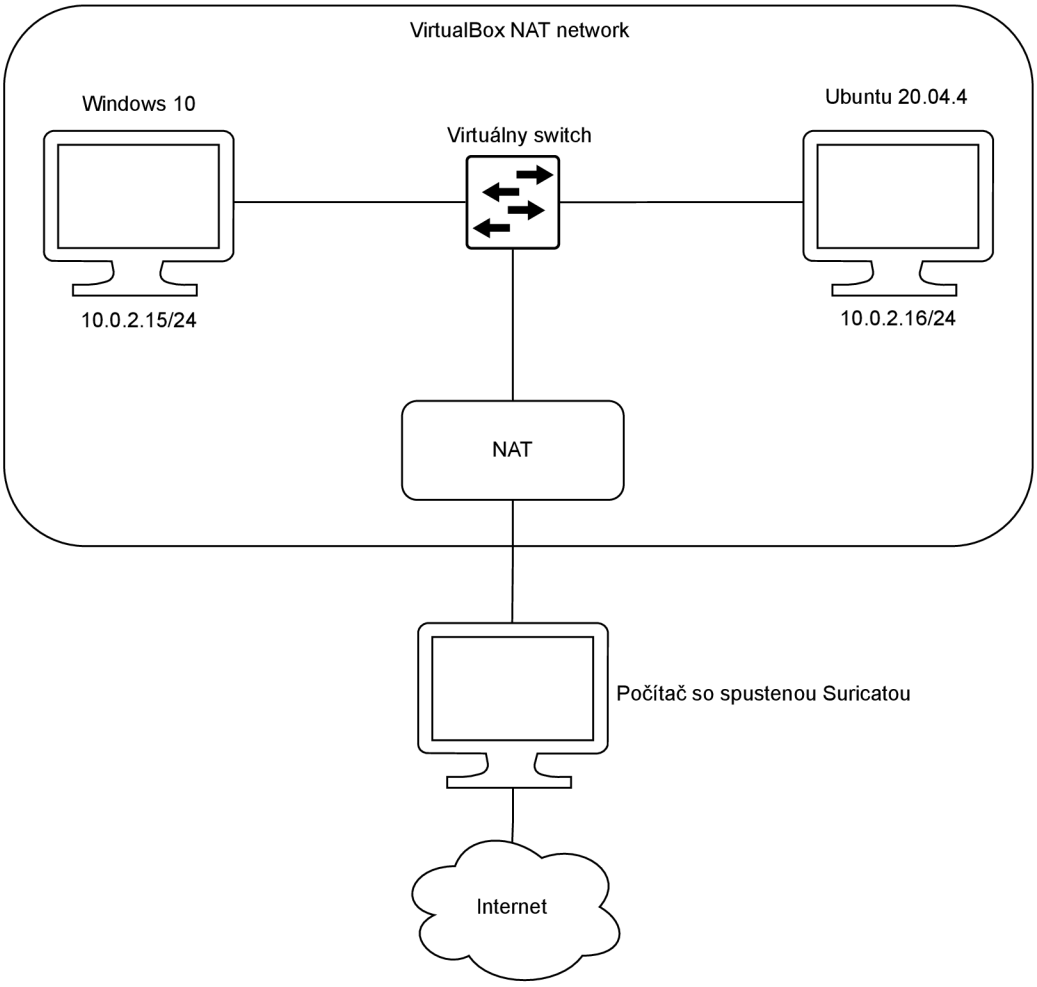
Zachytenie normálnej sieťovej komunikácie

Predošlé dátové sady obsahovali jak normálnu komunikáciu tak aj škodlivú. Z týchto dátových sad boli vytvorené príslušné modely (fuzzy množiny) jednotlivých typov IoC, ktoré boli použité pre detekciu.

Avšak pre otestovanie miery falošnej pozitivity (anglicky false positive) bolo potrebné vygenerovať sieťovú prevádzku, ktorá by odrážala typické chovanie užívateľa na sieti. Na to boli vytvorené dve virtuálne stanice, ktoré boli v jednej podsieti, a to konkrétne 10.0.2.0/24. Jedna stanica mala na sebe nainštalovaný operačný systém Windows 10, na ktorom bol nainštalovaný FlareVM sandbox[16]. Druhá stanica mala Ubuntu 20.04.4 LTS so základnými

[16] https://medium.com/@haroon00525/flare-vm-lab-setup-isolated-lab-environment-for-malware-analysis-6e7c23af875

aplikáciami. Následne na hlavnej fyzickej stanici bol spustený nástroj *Suricata*, ktorý počúval len na IP adresách, ktoré mali priradené jednotlivé stanice. To z dôvodu aby sa zamedzilo zachytávaniu ostatných nepotrebných sieťových tokov. Celkovú topológiu je možné vidieť na obrázku 5.1.



Obr. 5.1: Schéma zapojenia, ktorá bola použitá pre zachytenie normálnej sieťovej komunikácie v reálnom prostredí simulovaného za pomoci VirtualBoxu.

Komunikácia generovaná týmito stanicami pozostávala z navštevovania známych webových lokalít. tieto lokality boli vybrané na základe dát z webu Moz[17]. Aby množstvo tokov zachytenej normálnej komunikácie odpovedalo približne veľkosti predošlých dátových sád bolo zachytených 28 477 tokov.

[17] https://moz.com/top500

5.2 Experimenty s testovacími dátami

V tejto podkapitole budú riešené experimenty, ktoré boli vykonané nad dátovými sadami popísanými v 5.1. Na začiatok budú opísané metriky úspešnosti, ktoré boli použité na vyhodnotenie dosiahnutých výsledkov.

Samotné experimenty budú pozostávať z detekcie malwaru v zachytenej komunikácií vytvorenej zo získaných *pcap* súborov za pomoci IoC modelov jednotlivých dátových sád. Ďalej budú vyhodnotené ako dopadli modely IoC jednotlivých dátových sád pri detekcií dopredu zachytenej normálnej komunikácie za pomoci virtuálneho prostredia simulujúceho reálnu prevádzku. A ako posledné budú experimenty porovnania výpočtu prahovej hodnoty s a bez medzikvartilového rozpätia.

Metriky úspešnosti

Existuje mnoho rôznych metrík a spôsobov ako vyhodnotiť úspešnosť určitej metódy pre detekciu malwaru. Pre účely tejto práce boli použité základné metriky pre overenie toho, či navrhnutá metóda je vhodná na daný typ úlohy. Tieto metriky umožňujú pochopiť a zhodnotiť či je spomínaná metóda účinná a ak áno do akej miery prípadne v čom má daná metóda nedostatky a v čom je naopak dostačujúca.

Na takéto vyhodnotenie bola použitá tzv. (anglicky confusion matrix). Je to matica ktorá sa skladá zo štyroch hodnôt. Tieto hodnoty sú následovné:

- **Pravdivo negatívne** (anglicky True negative (TN)) - reprezentuje komunikačné okná, ktoré sú negatívnymi a boli označené ako negatívne. V kontexte tejto práce sú to komunikačné okná, ktoré neboli metódou označené ako škodlivé a škodlivými v skutočnosti aj neboli.
- **Falošne negatívne** (anglicky False negative (FN)) - reprezentuje komunikačné okná, ktoré sú pozitívnymi a boli označené ako negatívne. V kontexte tejto práce sú to komunikačné okná, ktoré boli metódou označené ako normálne ale v skutočnosti boli škodlivými.
- **Pravdivo pozitívne** (anglicky True positive (TP)) - reprezentuje komunikačné okná, ktoré sú pozitívnymi a boli označené ako pozitívne. V kontexte tejto práce sú to komunikačné okná, ktoré boli metódou označené ako škodlivé a škodlivými v skutočnosti aj boli.
- **Falošne pozitívne** (anglicky False positive (FP)) - reprezentuje komunikačné okná, ktoré sú negatívnymi a boli označené ako pozitívne. V kontexte tejto práce sú to komunikačné okná, ktoré boli metódou označené ako škodlivé ale v skutočnosti boli normálnymi. Niekedy je táto hodnota označovaná aj ako falošný poplach (anglicky False alarm).

U hodnôt **TN** a **TP** je žiaduce aby boli čo najvyššie, naopak u hodnôt **FP** a **FN** je vhodné aby boli čo najnižšie z očividných dôvodov. V kontexte tejto práce je dôležitou hodnotou **FP**. To z dôvodu toho, že ak by nástroj bol nasadený v reálnej prevádzke, nie je žiaduce aby označoval veľa tokov za škodlivé i keď nie sú, to by zbytočne zvyšovalo záťaž na sieťového administrátora, prípadne iného človeka alebo tím, ktorý má na starosti kyberbezpečnosť.

Tieto hodnoty sa ďalej môžu použiť na výpočet ostatných pokročilejších metrík pre dodatočné pochopenie presnosti danej metódy. Tieto ďalšie metriky zahŕňajú konkrétne:

- **Senzitivita** (anglicky Sensitivity (SEN)) - alebo tiež ako skutočná miera pozitívnosti. Je metrika, ktorá hovorí o tom ako dobre daná metóda deteguje malware. Čím je senzitivita vyššia tým má model menšiu pravdepodobnosť prehliadnutia skutočnej malware komunikácie. Cieľom je aby táto metrika bola čo najvyššie, ideálne rovno 1. Vzorec pre výpočet je následovný:

$$\mathrm{SEN} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}$$

- **Miera falošne pozitívnych výsledkov** (anglicky False positive rate (FPR)) - je to metrika, ktorá hovorí o tom ako často daný model resp. metóda nesprávne identifikuje okná, ktoré v skutočnosti neobsahujú žiadnu malware komunikáciu za škodlivé. Cieľom je aby táto metrika bola čo najnižšou, ideálne rovno 0. Vzorec pre výpočet je následovný:

$$\mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{FP} + \mathrm{TN}}$$

- **Presnosť** (anglicky Accuracy (ACC)) - jedná sa o metriku, ktorá hovorí o tom ako daný model resp. metóda detegovala v sieťovej komunikácií malware k pomeru celkovej zachytenej komunikácie. Cieľom je aby táto metrika bola čo najvyššie, ideálne rovno 1. Vzorec pre výpočet je následovný:

$$\mathrm{ACC} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}}$$

Základné experimenty s testovacími dátami

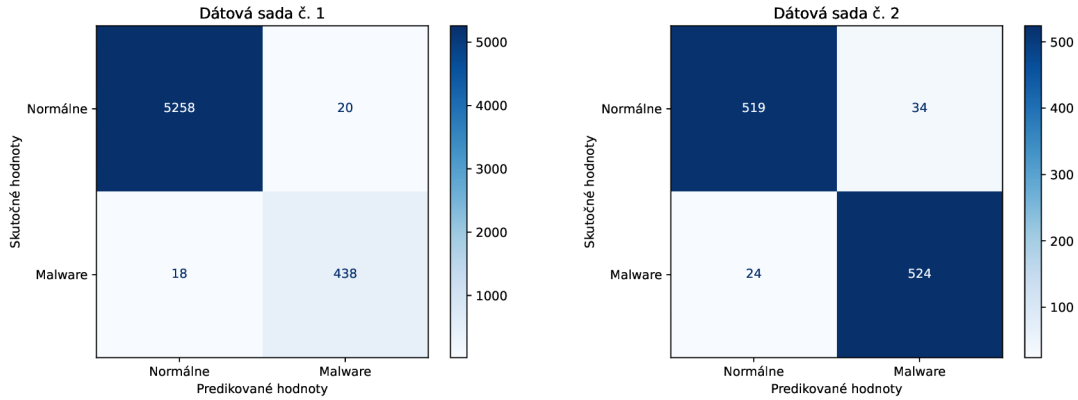
Ako prvé boli vykonané experimenty s testovacími dátami. Jedná sa o dátové sady, ktoré boli spomenuté v podkapitole 5.1. Následovné experimenty boli vykonávané s prahovou hodnotou počítanou s využitím medzikvartilového rozpätia. A veľkosťou komunikačných okien nastavených na 1 minútu.

Výsledky týchto experimentov je možné vidieť na v tabuľke 5.2. Podrobnejšie výsledky, poskytuje tabuľka 5.3, ktorá obsahuje metriky, ktoré boli dosiahnuté na jednotlivých dátových sadách.

Ďalej sa vypočítali *confusion matricies* pre každú dátovú sadu, kde je možné presne vidieť koľko tokov, ktoré metóda mala označiť za škodlivé označila a v koľkých sa naopak zmýlila. Tieto matice je možné vidieť na obrázku 5.2 pre obe dátové sady.

	Dátová sada č. 1	**Dátová sada č. 2**
Spracované okná	5734	1101
Referenčné malware okná	456	548
Detegované malware okná	458	558
Detegované normálne okná	5276	543

Tabuľka 5.2: Výsledky základných experimentov nad dátovými sadami č. 1 a č. 2.



Obr. 5.2: *Confusion matrix* dátových sád zobrazujúce počty TN, FN, TP a FP.

	Dátová sada č. 1	**Dátová sada č. 2**
Senzitivita	0.96053	0.9562
Miera falošnej pozitivity	0.00379	0.06148
Presnosť	0.99337	0.94732

Tabuľka 5.3: Metriky úspešnosti dosiahnuté v rámci jednotlivých dátových sád

Experimenty so zachytenou normálnou komunikáciou

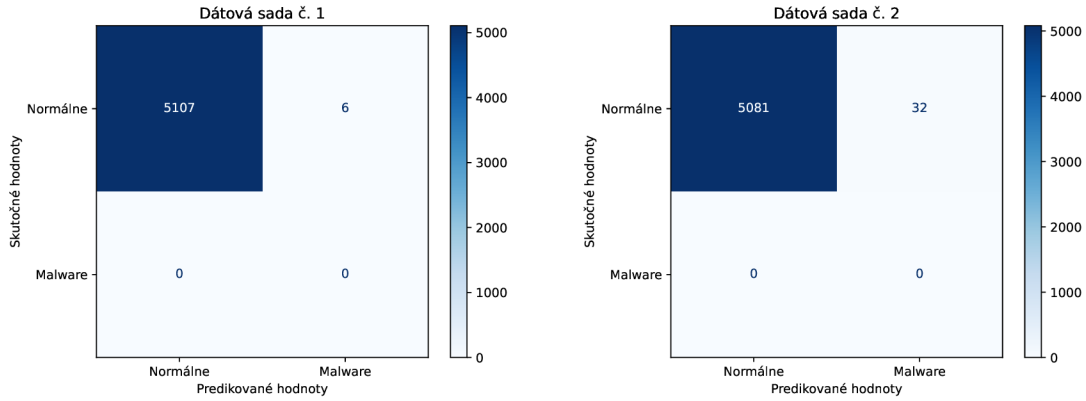
Tieto experimenty boli vykonané nad normálnou komunikáciou zachytenou nástrojom *Suricata* z vytvorenej virtuálnej podsiete viď. 5.1. Cieľom bolo otestovať koľko komunikačných okien metóda označí za potencionálne škodlivé na základe modelov vytvorených z dátových sád č. 1 a č. 2, i keď v tejto komunikácií by nemalo byť žiadne z vytvorených okien označené ako nositeľ malware komunikácie. Následovné experimenty boli vykonávané s prahovou hodnotou počítanou s využitím medzikvartilového rozpätia. A veľkosťou okien nastavených na 1 minútu.

V tabuľke 5.4 je možné vidieť celkový počet spracovaných okien a koľko z nich bolo detegovaných ako škodlivými na základe modelov vytvorených z dátových sád. Treba poznamenať, že v tomto prípade žiadna referenčná hodnota malware okien nie je uvedená z dôvodu toho, že táto komunikácia neobsahuje, resp. by nemala, obsahovať žiadnu malware komunikáciu. Podrobnejšie výsledky je možné vidieť v tabuľke 5.5 spolu s vypočítanými *confusion matricies* na obrázku 5.3.

Z výsledkov je možné vidieť, že dátová sada č. 1 poskytuje omnoho menší počet falošne označených okien ako tomu je v prípade dátovej sady č. 2. To môže byť spôsobené tým, obe dátové sady obsahujú iné malware rodiny a ich chovanie v rámci siete je odlišné, ďalším dôvodom je aj pravdepodobný nedostatok dát, čo vedie k nepresnostiam detekcie vytvorenej metódy.

	Dátová sada č. 1	Dátová sada č. 2
Spracované okná	5113	5113
Detegované malware okná	6	32
Detegované normálne okná	5107	5081

Tabuľka 5.4: Výsledky experimentov nad zachytenou normálnou komunikáciou.



Obr. 5.3: *Confusion matrix* z detekcie normálnej komunikácie na základe modelov vytvorených z dátových sád č. 1 a č. 2 zobrazujúce počty TN, FN, TP a FP.

	Dátová sada č. 1	Dátová sada č. 2
Miera falošnej pozitivity	0.00117	0.00626
Presnosť	0.99883	0.99374

Tabuľka 5.5: Metriky úspešnosti dosiahnuté pri experimentoch nad normálnou komunikáciou na základe modelov vytvorených z dátových sád č. 1 a č. 2.

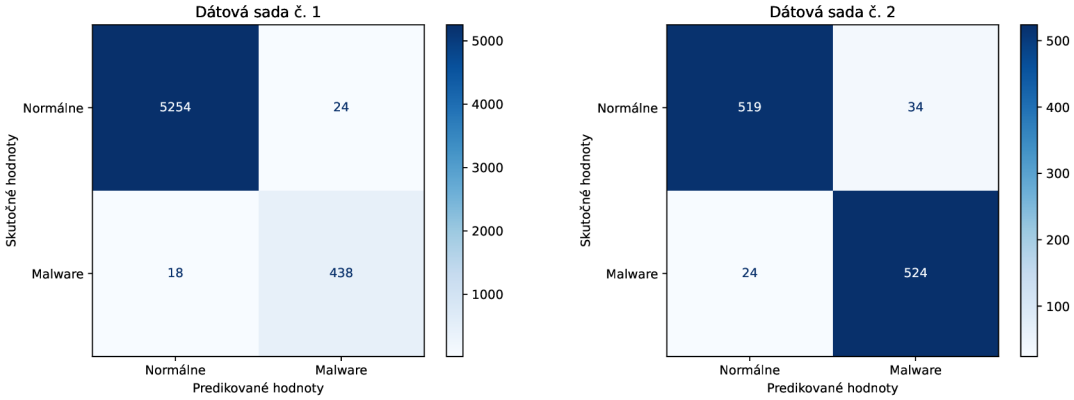
Porovnanie výpočtu prahových hodnôt

Ak by sa pre výpočet prahových hodnôt nepoužila metóda medzikvartilového rozpätia a lokálne extrémy by boli ponechané, došlo by k určitému skresleniu v rámci detekcie. A práve tieto experimenty majú poukázať na rozdiel medzi týmito dvomi, z logického hľadiska validnými, spôsobmi výpočtu prahových hodnôt.

Experimenty v tabuľke 5.6 boli vykonané nad zachytenou komunikáciou vytvorenou zo získaných *pcap* súborov jednotlivých dátových sád. Je vidieť že v dátovej sade č. 2 sa hodnoty nezmenil, to môže byť spôsobené tým, že medzikvartilové rozpätie nemalo taký drastický účinok na prahové hodnoty jednotlivých malware rodín. *Confusion matricies* viď. 5.4 boli vypočítané pre podrobnejšiu vizualizáciu dosiahnutých výsledkov. V tabuľke 5.7 je možné vidieť ostatné pokročilejšie metriky, ktoré boli použité pre vyhodnotenie.

	S medzikvartilovým rozpätím		Bez medzikvartilového rozpätia	
	Dátová sada č. 1	**Dátová sada č. 2**	**Dátová sada č. 1**	**Dátová sada č. 2**
Spracované okná	5734	1101	5734	1101
Referenčné malware okná	456	548	456	548
Detegované malware okná	458	558	462	558
Detegované normálne okná	5276	543	5272	543

Tabuľka 5.6: Výsledky základných experimentov nad dátovými sadami č. 1 a č. 2. s rozdielnym výpočtom prahových hodnôt.



Obr. 5.4: *Confusion matrix* dátových sád zobrazujúce počty TN, FN, TP a FP, pre dva rôzne spôsoby výpočtu prahových hodnôt

	S medzikvartilovým rozpätím		Bez medzikvartilového rozpätia	
	Dátová sada č. 1	**Dátová sada č. 2**	**Dátová sada č. 1**	**Dátová sada č. 2**
Senzitivita	0.96053	0.9562	0.96053	0.9562
Miera falošnej pozitivity	0.00379	0.06148	0.00455	0.06148
Presnosť	0.99337	0.94732	0.99268	0.94732

Tabuľka 5.7: Metriky úspešnosti dosiahnuté v rámci jednotlivých dátových sád s rozdielnym výpočtom prahových hodnôt.

Z daných experimentov je možné konštatovať, že odstránením lokálnych extrémov pri výpočte prahových hodnôt prináša istú mieru zlepšenia. V prvej dátovej sade je počet falošne pozitívnych okien 0.42% ((24/5734) ∗ 100) oproti 0.35%, pri druhej dátovej sade neboli namerané žiadne rozdiely. Pri väčších dátových sadách by rozdiel bol pravdepodobne markantnejší.

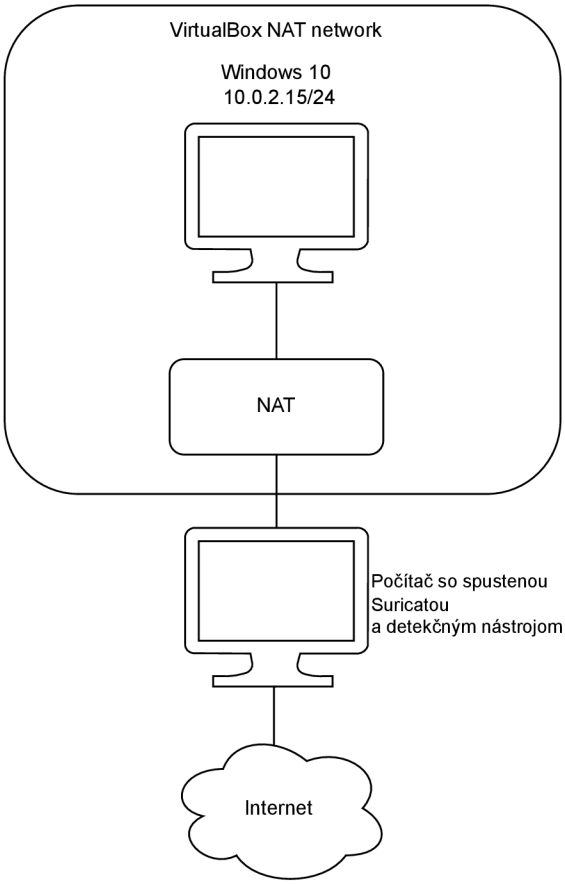
5.3 Experimenty v reálnej prevádzke

Posledné z rady experimentov boli vykonané v reálnej prevádzke, ktorá bola simulovaná za pomoci virtuálneho stroja. Preto bola vytvorená virtuálna stanica za pomoci ktorej bol nástroj na detekciu otestovaný. *Suricata* pre zachytávanie sieťovej komunikácie spolu s nástrojom pre detekciu malwaru bol spustený na fyzickom počítači na ktorom bežal Windows 10 a samotný nástroj bol spustený v prostredí WSL[18] . Nástroj pre detekciu malwaru mal nastavenú veľkosť komunikačných okien na 1 minútu. Konfigurácia nástroju *Suricata* je

[18] https://learn.microsoft.com/en-us/windows/wsl/install

takmer nezmenená až na sieť v ktorej sa nachádza daná monitorovacia stanica, aby sa zachytávala len komunikácia, ktorá pochádza z danej podsiete. V tomto prípade sa jednalo o podsieť 10.2.0.0/24.

Virtuálny stroj mal operačný systém Windows 10 s už spomínanou nadstavbou FlareVM, z dôvodu toho aby malware nemohol jdnoducho zistiť, či sa nachádza vo virtuálnom prostredí alebo nie. Jedná sa o to, že mnohé malwary v súčasnosti ako prvé zisťujú, či nie sú spustené vo virtuálnom prostredí aby nemohlo dôjsť k ich analýze prípadne tzv. (anglicky reverse engineering)-u. Následne ako bola príslušná virtuálna stanica vytvorená, bola prepojená prostredníctvom NAT k internetu. Prístup k internetu je tiež vážnym faktorom, nakoľko ak malware nemá k nemu prístup, nebude prejavovať žiadnu sieťovú komunikáciu. Boli skúšané mnohé nástroje ako je napríklad *FakeNet*[19], prípadne aj modifikovaný Ubuntu s nadstavbou nazývanou REMnux[20], ktorá obsahuje mnoho užitočných nástrojov pre analýzu malwaru z pohľadu kódu, pamätovej stopy (anglicky memory footprint) a tak ďalej. Tiež poskytuje aj nástroje na simuláciu sieťovej prevádzky, fake DNS a podobné nástroje. Avšak aj napriek tomuto pri testovaní so spomenutými nástrojmi malware neprejavoval známky sieťovej komunikácie, s ktorou by sa dalo pracovať. Výslednú topológiu je možné vidieť na obrázku 5.5.



Obr. 5.5: Schéma zapojenia, ktorá bola použitá jak pre vykonanie experimentov v reálnej prevádzke.

[19] https://sourceforge.net/projects/fakenet/
[20] https://docs.remnux.org/

Samotný experiment pozostával z dvoch častí. V prvej časti sa bol vo virtuálnom stroji s operačným systémom Windows 10 spustený webový prehliadač Google Chrome. V priebehu 3 minút bola generovaná sieťová prevádzka, ktorá pozostávala z navštevovania bežných webov, ako napríklad facebook.com, youtube.com a mnohé ďalšie viď. stránka Moz[21]. V ideálnom prípade by nástroj nemal označiť žiadne z vytvorených komunikačných okien ako škodlivé. Pre experiment boli použité obe dátové sady pre porovnanie IoC modelov z nich vytvorených.

Ako prvé boli otestované IoC modely dátovej sady č. 1. V prvej časti bolo zachytených 1160 tokov z ktorých bolo vytvorených 66 okien kde nástroj označil 2 okná za škodlivé. V tomto prípade sa jedná o falošne pozitívne detekcie. Čiže v prvej časti bolo chybne označených 3.03% okien. Výpis takéhoto okna, ktoré bolo označené škodlivým je možné vidieť vo výpise 5.2.

```
Window start: 2024-04-20 19:42:00
timestamp     src_ip dest_ip dns http
2024-04-20 19:42:27.282494 10.2.0.2 142.250.179.131 NaN NaN
2024-04-20 19:42:27.612301 10.2.0.2 10.2.0.1 update.googleapis.com NaN
2024-04-20 19:42:30.886768 10.2.0.2 10.2.0.1 cdn.alza.cz NaN
... ... ... ... ...
2024-04-20 19:42:51.601116 10.2.0.2 142.251.36.35 NaN NaN
2024-04-20 19:42:51.678163 10.2.0.2 195.181.172.26 NaN NaN
2024-04-20 19:42:51.705947 10.2.0.2 84.17.46.53 NaN NaN
```

Výpis 5.2: Komunikačné okno zachytené počas prvej časti experimentu. Jedná sa len o časť komunikačného okna nakoľko dané okno obsahovalo 56 tokov.

V druhej časti sa na virtuálnom stroji Windows 10 spustila vzorka malwaru z rodín *upatre*, *njrat* a *sodinobiki*. Všetky tieto rodiny sú súčasťou dátovej sady č. 1. Tieto vzorky sa prejavovali miernou sieťovou komunikáciou, ktorú by opäť v ideálnom prípade mal nástroj odhaliť. Spomínané malware vzorky, po dobe 3 minút od ich spustenia bolo zachytených 896 tokov z ktorých bolo vytvorených 28 okien. Celkovo tak nástroj označil 7.14% okien z 28 za škodlivé, príklad takéhoto okna, resp. jeho úryvok, je možné vidieť vo výpise 5.3.

```
Window start: 2024-04-20 19:45:00
timestamp     src_ip dest_ip dns http
2024-04-20 19:45:29.956057 10.2.0.2 157.240.247.8 NaN NaN
2024-04-20 19:45:29.956093 10.2.0.2 2.16.6.85 NaN NaN
2024-04-20 19:45:32.954306 10.2.0.2 142.250.179.198 NaN NaN
... ... ... ... ...
2024-04-20 19:45:44.953046 10.2.0.2 147.229.191.143 NaN NaN
2024-04-20 19:45:51.835656 10.2.0.2 185.159.159.1 NaN protonstatus.com/
    ↪ vpn_status
2024-04-20 19:45:53.951526 10.2.0.2 20.190.159.68 NaN NaN
```

Výpis 5.3: Komunikačné okno zachytené po spustení malware vzoriek. Jedná sa len o úryvok daného okna v skutočnosti malo dané okno 36 tokov.

Celkovo je možné vidieť, že IoC modely vytvorené za pomoci dátovej sady č. 1 detegovali 4 okná z celkových 96 ako škodlivé. Príklad zachyteného škodlivého okna s jednotlivými

[21] https://moz.com/top500

skóre malware rodín je možné vidieť vo výpise 5.4. Stojí za povšimnutie, že v danom okne boli detegované príznaky malware rodín, ktorých vzorky sa ale nespúštali počas testovania. To môže byť spôsobné tým, že daná rodina má podobne chovanie v rámci sieťovej komunikácie, to znamená, že boli nájdené podobné IoC, ktoré sa zhodujú s inými rodinami ktoré boli v danom okne označené.

```
{"timestamp": "2024-04-20T19:43:00_0", "pc": "10.2.0.2",
       "malware_families_scores": {"asyncrat": 0.0,
              "azorult": 0.181, "darkcomet": 0.0,
              "dcrat": 0.196, "emotet": 0.0, "hawkeye": 0.0,
              "icedid": 0.0, "metasploit": 0.0, "netwire": 0.0,
              "njrat": 0.0, "sodinokibi": 0.0, "upatre": 0.0},
              "malware": true}
```

Výpis 5.4: Komunikačné okno s dosiahnutým skóre jednotlivých malware rodín.

IoC modely dátovej sady č. 2 vykazovali následujúce výsledky. V prvej časti bolo zachytených 2118 tokov z ktorých bolo vytvorených 53 okien. Nástroj na základe IoC modelov z dátovej sady č. 2 označil 3 okná za potencionálne škodlivé. Príklad takéhoto okna je možné vidieť vo výpise 5.5. V prvej časti nástroj klasifikoval celkovo 5.66% okien z celkového počtu 53 okien ako škodlivé.

```
Window start: 2024-04-20 22:08:00
timestamp     src_ip dest_ip dns
2024-04-20 22:08:25.286256 10.2.0.2 13.107.213.67 NaN
2024-04-20 22:08:25.525662 10.2.0.2 142.251.39.99 NaN
2024-04-20 22:08:25.932133 10.2.0.2 10.2.0.1 pay.google.com
... ... ... ...
2024-04-20 22:08:52.359041 10.2.0.2 185.159.159.148 NaN
2024-04-20 22:08:56.394885 10.2.0.2 147.229.191.143 adservice.google.com
2024-04-20 22:08:56.897656 10.2.0.2 52.167.85.21 NaN
```

Výpis 5.5: Komunikačné okno zachytené počas prvej časti experimentu. Jedná sa len o časť okna nakoľko okno samotné obsahovalo 189 tokov.

Následne po spustení malware vzoriek z rodín *vidar*, *darkgate* a *avaddon* bolo zachytených 900 tokov z ktorých bolo vytvorených 19 komunikačných okien. Z toho tri okná boli označené za škodlivé, viď. výpis 5.6, čo predstavuje 15.79%.

```
Window start: 2024-04-20 22:13:00
timestamp     src_ip dest_ip dns  http
2024-04-20 22:13:26.241245 10.2.0.2 142.251.36.10 NaN NaN
2024-04-20 22:13:26.241279 10.2.0.2 10.2.0.1 NaN NaN
2024-04-20 22:13:26.241311 10.2.0.2 142.250.179.198 NaN NaN
... ... ... ... ...
2024-04-20 22:13:59.554890 10.2.0.2 10.2.0.1 google-ohttp-relay-
    ↪ safebrowsing.fastly-edge.com NaN
2024-04-20 22:13:59.597984 10.2.0.2 151.101.37.91 NaN NaN
2024-04-20 22:13:59.685968 10.2.0.2 10.2.0.1 selectwendormo9tres.com NaN
```

Výpis 5.6: Komunikačné okno zachytené počas druhej časti experimentu. Toto okno obsahovalo celkovo 35 tokov.

Vo výsledku dátová sada č. 2 a IoC modely vytvorené na základe IoC jednotlivých vzoriek malware rodín, ktoré táto dátová sada obsahovala, detegovala 6 okien z celkových 72 ako škodlivé. Príklad okien, ktoré boli označené za škodlivé sú vo výpisoch 5.5 a 5.6. Vo výpise 5.7 je príklad okna s dosiahnutým skóre jednotlivých malware rodín. Opäť stojí za povšimnutie fakt, že dané okno obsahuje pravdepodobne komunikáciu rodiny *backdoor.teamviewer*, i keď žiadny zo spustených vzoriek nebol z tejto rodiny. Toto môže byť spôsobné opäť tým, že daná rodina má podobne chovanie v rámci sieťovej komunikácie, s inými rodinami ktoré boli v danom okne označené.


```
{"timestamp": "2024-04-20T22:13:00_7", "pc": "10.2.0.2",
       "malware_families_scores": {"amadey": 0.0, "avaddon": 0.0,
              "backdoor.teamviewer": 0.277, "darkgate": 0.0,
              "dridex": 0.0, "formbook": 0.0, "gcleaner": 0.0,
              "healer": 0.0, "heodo": 0.0, "irata": 0.259,
              "loki": 0.0, "mirai": 0.0, "pikabot": 0.0,
              "quakbot": 0.0, "redlinestealer": 0.0,
              "remcosrat": 0.0, "riseprostealer": 0.0,
              "stealc": 0.0, "tofsee": 0.0,
              "trickbot": 0.0, "vidar": 0.0}, "malware": true}
```


Výpis 5.7: Komunikačné okno s dosiahnutým skóre jednotlivých malware rodín.

V uvedených experimentoch bolo možné sledovať mierne navýšenie detekcie malwaru v sieťovej komunikácií po tom ako boli spustené malware vzorky. Avšak stále je tu istý výskyt falošne pozitívnych detekcií, ktorý je vyšší ako u testovacích dát, kde miera falošnej pozitivity nepresiahla 1%. Nástroj do istej mieri je schopný detegovať ak je na počítači spustený nejaký malware, avšak občas nástroj nesprávne deteguje malware rodinu, ktorej vzorok nebol vôbec spustený na danom zariadení. To je spôsobené, ako už bolo spomínané, podobnosťou chovania jednotlivých malware rodín. Je takmer isté, že ak by nástroj mal k dispozícií väčšiu dátovú sadu, s väčším obsahom informácií, bola by detekcia presnejšia a určite aj citlivejšia na prejavy malware komunikácie.

Kapitola 6

Záver

Cieľom tejto práce bolo zoznámiť sa s problematikou malwaru a jeho detekcie v rámci sietí a následne analyzovať sieťovú komunikáciu malwaru a identifikovať významné vlastnosti, ktoré by boli vhodné pre vytvorenie nástroja na detekciu malwaru. Ako prvé bolo potrebné naštudovať problematiku malwaru, na základe čoho sa malware delí čo sú to malware rodiny a ako sa prejavuje v rámci sieťovej komunikácie. Následne bola vytvorená dátová sada, zložená zo 17 malware rodín po 5 vzoriek, za pomoci *sandbox* prostredia Triage. Na osnove tejto vytvorenej dátovej sady prebehla analýza sieťovej komunikácie malwaru a identifikácia vlastností, ktoré by mohli byť nápomocnými v rámci detekcie. Po nájdení týchto významných vlastností bolo potrebné ďalej overiť ich relevantnosť, to znamená či ich je možno považovať za indikátory malware komunikácie alebo nie. Následne bola metóda vo forme funkčného nástroja nasadená v simulovanom reálnom prostredí a otestovaná na vytvorených dátových sadách a aj v rámci reálnej sieťovej prevádzky.

Výsledky experimentov vykonaných nad dvomi dátovými sadami, dátová sada č. 1 dátová sada č. 2, ukázali nasledujúce výsledky. Dátová sada č. 1 dosiahla 99.337% presnosť a dátová sada č. 2 dosiahla 94.732% presnosť. Pri experimentoch nad zachytenou normálnou komunikáciou mala dátová sada č. 1 99.883% a č. 2 mala 99.374% a falošná pozitivita bola na úrovni 0.117% respektíve 0.626% u dátovej sady č.2. Tieto dosiahnuté výsledky je možné označiť vynikajúcimi. Následne experimenty v reálnej prevádzke u modelov IoC vytvorených na osnove dátovej sady č. 1 ukázali, že počet nesprávne detegovaných okien dosiahla 3.03% resp. 5.66% v rámci modelov dátovej sady č. 2. Celkovo je možné povedať, že výsledky prvej časti tohto experimentu, nie sú až tak vynikajúce ako v prípade experimentov nad dopredu zachytenou komunikáciou, kde miera falošnej pozitivity nepresiahla 1%.

Dosiahnuté výsledky sú vcelku skvelé ale stále je dosť priestoru na zlepšenie hlavne čo sa týka detekcie v reálnej prevádzke. Ako prvé by bolo určite vhodné metóde poskytnúť viac dát na základe ktorých by sa mohli vytvoriť silné IoC modely, ktoré by obsahovali dostatočné množstvo informácií pre metódu aby sa rozhodla, či je dané okno škodlivé alebo nie. Druhým zaujímavým rozšírením by mohlo byť využitie natrénovaného algoritmu na určenie prahových hodnôt pre každé okno zvlášť a nie globálne.

Celkovo práca dosiahla svojich cieľov a jej výsledkom je nástroj, ktorý je použiteľný na detekciu malwaru v sieťovej komunikácií. Mimo iné práca demonštrovala, kvalitu vytvorenej metódy a to jak v offline režime, kedy sieťová komunikácia bola dopredu zachytená tak aj v režime online, kedy metóda detegovala danú komunikáciu v reálnom čase a posudzovala, či sa v nej vyskytuje malware alebo nie.

Literatúra

[1] *AlienVault OSSIM is trusted by security professionals across the globe* [online]. ATT Cybersecurity [cit. 2024-01-11]. Dostupné z: https://cybersecurity.att.com/products/ossim.

[2] Malware. *Biblioteka.sk* [online]. biblioteka [cit. 2023-12-19]. Dostupné z: https://www.biblioteka.sk/encyklopedia/index.php?pojem=Malware.

[3] *DirectConnect API* [online]. ALIENVAULT [cit. 2024-01-11]. Dostupné z: https://otx.alienvault.com/api.

[4] *Importance of IOC Detection Rules* [online]. Talanos Cybersecurity [cit. 2024-04-12]. Dostupné z: https://www.talanoscybersecurity.com/blogs/news/importance-of-ioc-detection-rules.

[5] Malware or malicious software definition. *Malwarebytes* [online]. Malwarebytes [cit. 2023-12-19]. Dostupné z: https://www.malwarebytes.com/malware.

[6] What is a botnet. *Malwarebytes* [online]. Malwarebytes [cit. 2023-12-19]. Dostupné z: https://www.malwarebytes.com/botnet.

[7] Malware names. *Microsoft* [online]. Microsoft [cit. 2023-12-26]. Dostupné z: https://learn.microsoft.com/en-us/microsoft-365/security/intelligence/malware-naming?view=o365-worldwide.

[8] *Open Threat Exchange and USM Appliance* [online]. ATT Cybersecurity [cit. 2024-01-11]. Dostupné z: https://cybersecurity.att.com/documentation/usm-appliance/otx/about-otx.htm.

[9] DNS Security Solutions. *ThreatStop* [online]. ThreatStop [cit. 2024-01-05]. Dostupné z: https://www.threatstop.com/.

[10] *Understanding Malware: Exploring the World of Cyber Threats* [online]. Medium [cit. 2024-04-12]. Dostupné z: https://medium.com/@i.vikas/understanding-malware-exploring-the-world-of-cyber-threats-9ad2182d94b1.

[11] IOCs vs Artifacts: What is the difference. *VMRAY* [online]. VMRAY [cit. 2023-12-21]. Dostupné z: https://www.vmray.com/iocs-vs-artifacts-whats-the-difference/.

[12] *What Is Alienvault And How It Works* [online]. Mike Gingerich Global, LLC [cit. 2024-01-11]. Dostupné z: https://www.mikegingerich.com/blog/what-is-alienvault-and-how-it-works/.

[13] Baker, K. *What are the Types of Malware?* [online]. Crowdstrike, 28. februára 2023 [cit. 2023-12-19]. Dostupné z: https://www.crowdstrike.com/cybersecurity-101/malware/types-of-malware/.

[14] Bekerman, D., Shapira, B., Rokach, L. a Bar, A. Unknown malware detection using network traffic classification. [online]. Beer Sheva, Israel: IEEE. 2015, zv. 5, č. 2, s. 134–142, revidováno 21. 3. 2014, [cit. 2023-12-19]. DOI: 10.1109/CNS.2015.7346821. Dostupné z: https://ieeexplore.ieee.org/abstract/document/7346821?casa_token=EgY3Aibl1gMAAAAA:z8hypGq1VL5obc1Nh8IgSHPfmLfXRJKxsynfuIuOXRpYtEc4_IqPKvikWKrHy5Pbt8jkIvNN7A.

[15] Ekta Gandotra, S. S. Malware Analysis and Classification: A Survey. [online]. verze 1.0. Chandigarh: PEC University of Technology, Chandigarh, India. Február 2014, zv. 5, č. 2, s. 9, revidováno 21. 3. 2014, [cit. 2023-12-19]. DOI: 10.4236/jis.2014.52006. Dostupné z: https://www.scirp.org/pdf/JIS_2014040110394271.pdf.

[16] Jay, J. a Bob, R. Data-driven security: analysis, visualization and dashboards. In: [book]. Indianapolis, Indiana: John Wiley & Sons, 2014, kap. 2, 3 [cit. 2023-12-22]. ISBN 978-1-118-79372-5.

[17] Korvas, V. *DETEKCE KOMUNIKACE MALWARE V SÍŤOVÝCH TOCÍCH.* Brno, CZ, 2023. Bakalářská práce. Vysoké učení technické v Brně, Fakulta informačních technologií. Dostupné z: https://www.vut.cz/studenti/zav-prace/detail/146955.

[18] Malicious, A. 13th International Conference on a (MALWARE), U. S. Behavioral Malware Classification using Convolutional Recurrent Neural Networks. [online]. Nantucket: MA, USA. Október 2018, s. 103–111, [cit. 2023-12-26]. DOI: 10.1109/MALWARE.2018.8659358. Dostupné z: https://ieeexplore.ieee.org/abstract/document/8659358?casa_token=S_XP1oZOSwEAAAAA:xloqgEx0RXFeJykL5PqMHU8l-1xqFykmqKF4yv4QVH-PMbWTTskgRheLJNmWdXJ7Mf_YhK1IiA.

[19] Onyegbula, B. Indicators of Compromise (IoCs): What Are They and How Do They Strengthen Cyber Defense? *Splunk>* [online]. splunk>, 31. mája 2023 [cit. 2023-12-21]. Dostupné z: https://www.splunk.com/en_us/blog/learn/ioc-indicators-of-compromise.html.

[20] Pachhala, N., Jothilakshmi, S. a Battula, B. P. A Comprehensive Survey on Identification of Malware Types and Malware Classification Using Machine Learning Techniques. [online]. IEEE. 2021, s. 1207–1214, [cit. 2023-12-19]. DOI: 10.1109/ICOSEC51865.2021.9591763. Dostupné z: https://ieeexplore.ieee.org/document/9591763.

[21] Stallings, W. Local Networks. *ACM Comput. Surv.* New York, NY, USA: Association for Computing Machinery. Marec 1984, zv. 16, č. 1, s. 3–41, [cit. 2023-12-21]. DOI: 10.1145/861.871. ISSN 0360-0300. Dostupné z: https://doi.org/10.1145/861.871.

Príloha A

Obsah priloženého pamäťového média

Adresárová štruktúra uložená na pamäťovom médiu je nasledujúca:

- **01-BP_text** - priečinok obsahujúci textový výstup tejto bakalárskej práce vo formáte PDF, spolu so zdrojovými súbormi LaTeX a obrázkami potrebných na vygenerovanie spomenutého PDF.
- **02-Malware_detector** - priečinok obsahujúci programovú časť bakalárskej práce. Súčasťou tohto priečinku sú aj dátové sady, ktoré boli použité na vykonanie experimentov.

Výsledná štruktúra je zobrazená v nasledujúcom strome:

```
Pamäťové médium
├── 01-BP_text/
│   ├── obrazky-figures/
│   ├── template-fig/
│   ├── bib-styles/
│   ├── "xpapad11.pdf"
│   ├── "xpapad11-print.pdf"
│   ├── "zadani.pdf"
│   ├── "xpapad11.tex"
│   ├── "xpapad11-01-kapitoly-chapters.tex"
│   ├── "xpapad11-20-literatura-bibliography.bib"
│   ├── "xpapad11-30-prilohy-appendices.tex"
│   ├── "Makefile"
│   └── "fitthesis.cls"
└── 02-Malware_detector
    ├── _analysis of the dataset/
    │   └── "systematic_analysis.ipynb"
    ├── Cache backups/
    │   ├── dataset1/
    │   │   ├── "checked_families_domains"
    │   │   ├── "checked_families_ips"
    │   │   └── "checked_families_url"
    │   └── dataset2/
    │       ├── "checked_families_domains"
    │       ├── "checked_families_ips"
    │       └── "checked_families_url"
    ├── Example datasets/
    │   ├── dataset1/
    │   └── dataset2/
    ├── models_cache/
    │   ├── "checked_families_domains"
    │   ├── "checked_families_ips"
    │   └── "checked_families_url"
    ├── src/
    │   ├── "detection_functions.py"
    │   ├── "domain_model.py"
    │   ├── "ip_model.py"
    │   ├── "offline_detection.py"
    │   ├── "online_detection.py"
    │   ├── "stats.py"
    │   └── "url_model.py"
    ├── "dataset_checker.py"
    ├── "main_malware_detection.py"
    ├── "README.md"
    └── "requirements.txt"
```