VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ
FACULTY OF INFORMATION TECHNOLOGY

ÚSTAV INFORMAČNÍCH SYSTÉMŮ
DEPARTMENT OF INFORMATION SYSTEMS

DETEKCE MALIGNÍCH DOMÉN S VYUŽITÍM REPUTAČNÍCH SYSTÉMŮ V NÁSTROJI DOMAINRADAR

MALICIOUS DOMAIN DETECTION WITH REPUTATION SYSTEMS IN DOMAINRADAR

BAKALÁŘSKÁ PRÁCE
BACHELOR'S THESIS

AUTOR PRÁCE **MATĚJ ČECH**
AUTHOR

VEDOUCÍ PRÁCE **Ing. RADEK HRANICKÝ, Ph.D.**
SUPERVISOR

BRNO 2025

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
FAKULTA INFORMAČNÍCH TECHNOLOGIÍ

Zadání bakalářské práce

162578

Ústav:	Ústav informačních systémů (UIFS)
Student:	**Čech Matěj**
Program:	Informační technologie
Název:	**Detekce maligních domén s využitím reputačních systémů v nástroji DomainRadar**
Kategorie:	Bezpečnost
Akademický rok:	2024/25

Zadání:

1. Seznamte se s reputačními systémy pro hodnocení doménových jmen, URL a IP adres (SenderScore, Cisco Talos Intelligence, Barracuda Central, CESNET NERD apod.).
2. Nastudujte dosavadní výzkum v detekci maligních domén na projektu MV ČR FETA a seznamte se s nástrojem DomainRadar.
3. Po konzultaci s vedoucím navrhněte rozšíření nástroje DomainRadar, které bude využívat informací z reputačních systémů pro zvýšení přesnosti detekce maligních domén.
4. Navržené rozšíření implementujte.
5. Experimentálně ověřte přínos vašeho rozšíření.
6. Diskutujte dosažené výsledky.

Literatura:

- Hendrik, Ferry, Kris Bubendorfer a Ryan Chard. "Reputation systems: A survey and taxonomy." *Journal of Parallel and Distributed Computing, č.* 75, 2015: s. 184-197.
- Antonakakis, Manos, Roberto Perdisci, David Dagon, Wenke Lee a Nick Feamster. "Building a dynamic reputation system for DNS". *19th USENIX Security Symposium (USENIX Security 10)*. 2010.
- Hranický, Radek, Adam Horák, Jan Polišenský, Kamil Jeřábek a Ondřej Ryšavý. Unmasking the Phishermen: "Phishing Domain Detection with Machine Learning and Multi-Source Intelligence". *Proceedings of IEEE/IFIP Network Operations and Management Symposium 2024*. Soul, 2024, s. 1-5.

Při obhajobě semestrální části projektu je požadováno:
Body 1 až 3.

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

Vedoucí práce:	**Hranický Radek, Ing., Ph.D.**
Vedoucí ústavu:	Kolář Dušan, doc. Dr. Ing.
Datum zadání:	1.11.2024
Termín pro odevzdání:	14.5.2025
Datum schválení:	22.10.2024

Abstrakt

Tato práce přináší programové řešení pro detekci maligních doménových jmen založeném na výstupech z široké škály reputačních systémů. Klíčovým přínosem je detailní návrh a popis rozšíření existujícího nástroje DomainRadar o modul pro klasifikaci domén, který využívá data získaná z reputačních systémů. Součástí práce je rovněž experimentální zhodnocení dostupných reputačních systémů z hlediska schopnosti detekovat škodlivá doménová jména. V rámci práce byly vytvořeny a natrénovány dva samostatné klasifikátory modelu LightGBM, přičemž jeden je zaměřen na detekci phishingových domén a druhý na domény určené k distribuci malware. Na testovací sadě dosahuje vytvořený klasifikátor phishingových domén přesnosti 99 % a klasifikátor malwarových domén přesnosti 99.16 %.

Abstract

This thesis presents a software solution for detecting malicious domain names based on the output of a wide range of reputation systems. The key part is detailed design and implementation of an extension to the existing DomainRadar tool, adding a module for a domain classifier that utilizes data from reputation systems. The thesis also includes experimental evaluation of available reputation systems in terms of their ability to detect malicious domain names. As a part of the thesis, two separate LightGBM classifiers were created and trained – one focused on detecting phishing domain names and the other on detecting domain names used to distribute malware. The classifiers were evaluated using a test data set on which the phishing classifier achieved accuracy of 99 % and the classifier for malware domain names reached accuracy of 99.16 %.

Klíčová slova

reputační systémy, phishing, malware, DomainRadar, doménové jméno, IP, klasifikace, detekce, AbuseIPDB, Cloudflare Radar, CriminalIP, Fortiguard, Google Safe Browsing, Greynoise, Hybrid Analysis, NERD, Opentip Kaspersky, Project Honeypot, Pulsedive, Threatfox, URLVoid, VirusTotal

Keywords

reputation systems, phishing, malware, DomainRadar, domain name, IP, classification, detection, AbuseIPDB, Cloudflare Radar, CriminalIP, Fortiguard, Google Safe Browsing, Greynoise, Hybrid Analysis, NERD, Opentip Kaspersky, Project Honeypot, Pulsedive, Threatfox, URLVoid, VirusTotal

Citace

ČECH, Matěj. *Detekce maligních domén s využitím reputačních systémů v nástroji DomainRadar*. Brno, 2025. Bakalářská práce. Vysoké učení technické v Brně, Fakulta informačních technologií. Vedoucí práce Ing. Radek Hranický, Ph.D.

Detekce maligních domén s využitím reputačních systémů v nástroji DomainRadar

Prohlášení

Prohlašuji, že jsem tuto bakalářskou práci vypracoval samostatně pod vedením pana Ing. Radka Hranického, Ph.D. a uvedl jsem všechny literární prameny, publikace a další zdroje, ze kterých jsem čerpal.

.....
Matěj Čech
12. května 2025

Poděkování

Velké díky patří vedoucímu mé práce, Ing. Radkovi Hranickému, Ph.D., za jeho skvělý přístup, ochotu a v neposlední řadě i za jeho cenné rady.

Obsah

Seznam obrázků

Kapitola 1

Úvod

Každým dnem vznikají nové webové stránky, které mají za cíl lidem uškodit. A tyto stránky jsou každým dnem více a více rafinovanější a běžný uživatel často nemusí být schopen je odlišit od legitimních stránek. Proto je potřeba se před nimi nějak bránit. Aby bylo možné tyto stránky odlišit od běžného provozu, je potřeba brát v potaz více detailů. Jelikož čím více relevantních dat z různých zdrojů je k dané webové stránce k dispozici, tím přesnější a spolehlivější může být její klasifikace. Jedním z přístupů, který může značně přispět k odhalení nebezpečných stránek, je využívání dat z reputačních systémů. Tyto systémy shromažďují a následně poskytují širokou škálu dat, která odráží i historii dané webové stránky.

Nástroj DomainRadar sbírá typické znaky škodlivých webových stránek z několika zdrojů. Podle těchto dat se následně sám rozhoduje, zda danou webovou stránku označí za benigní či maligní. Pro zajištění vyšší přesnosti klasifikace se nabízí využít právě reputační systémy, které mohou pomoci přinést lepší rozlišení mezi maligními a benigními webovými doménami. Proto je mým cílem zjistit, zda lze zvýšit přesnost detekce maligních domén při použití agregovaných reputačních dat a následně implementovat další klasifikační faktor do projektu DomainRadar. Tento faktor bude využívat právě data volně dostupných reputačních systémů, což může přinést zlepšení v detekci škodlivých doménových jmen a umožní nástroji DomainRadar lépe posoudit povahu dané webové stránky.

Současná řešení umožňují pouze získat informace o dané entitě z velké škály reputačních systémů a ty následně uživateli zobrazí. Výhodou takového řešení je, že to analytikovi ušetří práci, aby se nemusel pracně dotazovat o dané entitě vícero reputačních systémů. Ale velkou nevýhodou je, že i tak bude muset manuálně projít jednotlivé výsledky a sám je analyzovat. Zde přichází na řadu má práce, která výsledky o dané entitě jednotlivých reputačních systémů zpracuje a vyhodnotí do jednoznačné klasifikační podoby. A v zapojení se systémem DomainRadar proběhne ohodnocení entity automaticky na základě příchodu požadavku. Tudíž uživatel bude jednoznačně vědět, zdali reputační systémy danou entitu označili jako maligní či benigní.

Přínosy práce

Hlavním přínosem této práce je rozšíření nástroje DomainRadar o klasifikaci phishingových a malwarových doménových jménech na základě dat z reputačních systémů. Data z reputačních systémů se projevila jako další vhodné kritérium, pomocí kterého lze doménové jméno ohodnotit a tak zpřesnit detekci maligních domén v nástroji. Vybrané reputační systémy

byly nejdříve experimentálně zhodnoceny, aby bylo možné zjistit jejich přesnost v rozlišení maligních domén od benigních. Následně bylo v práci navrženo rozšíření jednotlivých částí aktuálního systému DomainRadar, které bylo poté do nástroje implementováno. V neposlední řadě byly vytvořeny dva klasifikační modely, jeden určený pro detekci phishingových domén a druhý pro detekci domén určených pro distribuci malware. Tyto modely byly zhodnoceny na testovacích sadách, u kterých klasifikátor phishingových domén dosahuje přesnosti 99 % a klasifikátor malwarových domén nabývá přesnosti 99.16 %. Oba klasifikátory na základě získaných informací mají slibné výsledky a přínosy v detekci maligních doménových jménech a to nejen pro nástroj DomainRadar, ale i pro detekci škodlivých domén obecně.

Struktura práce

Práce je rozdělena do následujících kapitol. Kapitola 2 představuje již existující nástroje, které využívají reputační systémy pro klasifikaci doménových jmen, udává popis jejich fungování a také zhodnocuje jejich výhody a nevýhody. V kapitole 3 je popsán princip, na kterém staví reputační systémy a dále popisuje, jakým způsobem získávají informace. Projekt DomainRadar je popsán v kapitole 4, krom základních informací jsou zde uvedeny a popsány i důležité části tohoto projektu, které využívá tato práce. V kapitole 5 jsou popsány některé volně dostupné reputační systémy, také zde jsou uvedeny jejich charakteristiky a poskytovaná data, která budou užitečná pro tuto práci. Dále zde jsou vypsány výhody a nevýhody těchto systémů. Experimentální zhodnocení těchto reputačních systémů je provedeno v kapitole 6. V kapitole 7 je uvedeno navržené řešení, které bude implementováno do nástroje DomainRadar. Implementace sběrové části je popsána v kapitole 8 a v kapitole 9 je popsáno vytvoření klasifikátorů domén na základě strojového učení. Experimentální ověření navrhnutého a implementovaného rozšíření je uvedeno v kapitole 10. Závěrečná kapitola 11 nakonec vše shrnuje.

Kapitola 2

Rozbor existujících řešení

V této kapitole bude představeno několik existujících řešení, která využívají reputační systémy pro analýzu a klasifikaci doménových jmen. Tato řešení využívají různé přístupy pro odhalení potenciálních hrozeb, přičemž některá řešení se zaměřují na detekci podezřelých webových stránek, zatímco jiné poskytují analytikům přístup k informacím z mnoha reputačních systémů. Každý popisovaný nástroj bude stručně představen z hlediska jeho principu fungování, výhod a nevýhod.

Příkladem existujícího řešení je studie od Chia-Mei Chen a kol. [11], kteří vytvořili systém, jenž zkoumá uživatelovy webové dotazy. Systém je rozdělen na dvě části: identifikace podezřelých webových stránek na základě reputace doménových jmen a poté následné podrobnější zkoumání pouze těch webových stránek, které byly označeny jako maligní [11]. Takto ušetří výpočetní čas a zdroje, aby nemuseli zkoumat všechny webové stránky, na které se uživatelé dotazují ve skutečném provozu na síti. Experimentálně zjistili, že přesnost takového filtru je 94 %. A díky rozdělení zkoumání webových stránek na dvě části ušetřili více než dvanáctinásobek výpočetního času v porovnání, kdyby zkoumali vše. Experiment simulovali v reálné síti s 560 tisíci URL dotazy za den. Během nasazení této práce na reálné síti taktéž odhalili maligní webové stránky, které nebyly identifikovány na veřejné černé listině.

Jedním z již existujících programů, které sbírají data poskytnuta reputačními systémy, je program Cortex [29], který je vyvíjen týmem TheHive Project. Tento program je veden jako open source a je poskytován pod licencí APGL-3.0. Cortex si klade za cíl ulehčit práci analytikovi tím, že mu poskytne jeden nástroj, ze kterého může získat informace z několika zdrojů. Cortex integruje mnoho kolektorů dat a to nejen z reputačních systémů. Dále poskytuje uživateli možnost vybrat si služby, ze kterých budou data sbírána [30]. Nevýhodou je však fakt, že tento program dokáže pouze získat informace o daném doménovém jméně či IP adrese. Tato data uživateli zobrazí a dále už s nimi nijak nepracuje.

Dalším podobným programem je API framework PyOTI [27, 28], který umožňuje získat data o doménových jménech, emailových adresách, IP adresách a URL adresách z různých analyzátorů a reputačních systémů. Tento program využívá rozhraní API daných služeb, kterými se dotazuje na ohodnocení dané entity. Stejnou nevýhodou tohoto programu je skutečnost, že program pouze shromažďuje informace o dané entitě, ale nijak je dále nezpracovává. Tedy ta nejdůležitější část – rozhodnout, zda se jedná o maligní entitu – je na uživateli.

Zmíněné programy taktéž nedokážou klasifikovat doménová jména v reálném provozu na síti. To znamená, že analytik nemůže automaticky v reálném čase odhalit v síti přístup na škodlivou doménu. Má pouze možnost vložit již posbíraná doménová jména do pro-

gramu. Program Cortex umožňuje analýzu jak IP adres tak doménových jmen, ale v obou případech probíhá analýza na základě již získaných dat a průběh analýzy tak není navržený pro průběžné monitorování síťového provozu v reálném čase. Místo toho se spíše jedná o nástroj pro jednotlivé vyhodnocení entit pomocí různých analyzátorů. Podobně PyOTI slouží k analýze doménových jmen, ale neprovádí aktivní sledování síťového provozu ani automatickou klasifikaci domén v reálném čase. Zde navíc musí uživatel předat doménové jméno každému kolektoru dat z podporovaných služeb zvlášť a následně i data sám vyextrahovat, jelikož jsou ve formátu JSON a odpovídají celé odpovědi získané z API daného systému. Oba nástroje tedy představují spíše podpůrné programy pro analýzu maligních doménových jmen, IP adres či jiných entit než nástroje určené pro automatickou detekci a klasifikaci doménových jmen.

Kapitola 3

Reputační systémy

V dnešní době je běžné, že se podvodníci snaží využít nepozornosti lidí a zneužít jejich důvěry. Ať už tím, že nabízí něco, co se tváří, že je dokonalé, ale opak je pravdou. Příkladem mohou být internetová tržiště, kde lidé nabízejí k prodeji své produkty. Jakým způsobem zjistit, zda danému prodejci můžou věřit? Například pomocí zpětné vazby od lidí, kteří již s tímto prodejcem uzavřeli obchod. Tato zpětná vazba typicky obsahuje ohodnocení prodejce v uděleném počtu hvězd z maximálního stanoveného počtu. Avšak nikomu by se nechtělo číst tisíce recenzí. A zde nastupuje na řadu reputační systém, který dokáže zagregovat všechna hodnocení jednoho prodejce do jedné hodnoty. Výsledná hodnota udává reputační skóre, podle kterého se mohou další zákaznící rozhodovat, zda s kupujícím provedou transakci či nikoliv. Tudíž je zřejmé, že je výhodné pro prodejce si udržet co nejlepší reputaci [26]. A to zvláště i v dnešní době, kdy se lidé bežně rozhodují na základě předchozích zkušeností a recenzí cizích lidí. Je však nutné brát v potaz, že zpětná vazba od uživatelů může být ovlivněna falešným hlášením.

Reputační systém lze obecně chápat jako prostředek, který sbírá informace o entitách. Ty následně zpracovává a transformuje na ohodnocení důvěry v danou entitu. Tato informace je pouze odhadem na základě toho, jak se daná entita chovala v minulosti [17]. Tudíž reputační systém nedokáže garantovat, zdali daná entita je věrohodná v aktuálním okamžiku. Ale vyjadřuje pouze odhad míry jistoty na základě předchozích transakcí entity s jinými entitami [17].

Kdo nebo co je entita, definuje kontext, ve kterém je reputační systém definován. Entitou může být například osoba, organizace, ale i doménové jméno a IP adresa.

Aktéry, vystupující v reputačních systémech, definuje pan Hendrikx a kol. jako *Trustor* (Důvěřovatel), *Trustee* (Důvěřovaný) a *Recommender* (Doporučovatel) [17]. Kde Důvěřovatel má zájem zjistit, zda Důvěřovanému může opravdu věřit. To zjistí tak, že se dotáže Doporučovatele, zda již má nějakou zkušenost s Důvěřovaným a pokud ano, tak jakou. Na základě této informace se může Důvěřovatel rozhodnout, zda opravdu Důvěřovanému bude věřit či nikoliv [17].

3.1 Reputační systém pro doménová jména a IP adresy

Cílem takového systému je přiřadit ohodnocení doménovým jménům nebo IP adresám. Toto ohodnocení by mělo odrážet historické transakce daného subjektu. Výsledné ohodnocení poté může být využito například pro odfiltrování maligního provozu na počítačové síti či zablokování spambotů (*služeb určených pro odesílání spamu*).

3.2 Získávání informací

Nejjednodušším způsobem je již dříve zmíněna zpětná vazba uživatelů. Zde je však nutné počítat s tím, že uživatel může zadat informace, které neodpovídají realitě. Tudíž tato vazba není spolehlivým faktorem.

V kontextu doménových jmen a IP adres se nejčastěji jedná o kombinaci dat získaných z černých listin (*blacklist*), bílých listing (*whitelist*), honeypotů, síťových sond (*network probe*), skenerů a detekovaných anomálií [5]. Daný reputační systém následně využívá heureistiky či metody strojového učení, aby ze záznamů, ke kterým má přístup, mohl získat reputační skóre [9].

3.3 Účel reputačních systémů

Díky reputačním systémům je možné na základě výsledného ohodnocení identifikovat škodlivý provoz či nevyžádaný spam, ten zablokovat a tak ochránit koncového uživatele, který by se jinak mohl stát obětí podvodu. Dále mohou být reputační systémy využity k detekci maligních aktivit, jako je například distribuce malware. Reputační systémy tak přidávají další pomocnou vrstvu k identifikaci těchto škodlivých entit. Opačným případem může být nárůst důvěřivosti u webových stránek, které mají velice dobrou reputaci. Uživatelé tak mohou být o něco klidnější, když ví, že komunikují se seriózní stranou.

Kapitola 4

DomainRadar

Projekt s názvem DomainRadar je vyvíjen týmem FETA pod záštitou Ministerstva vnitra České republiky, přičemž do něj jsou zapojeny instituce CESNET, VUT a ČVUT.

Cílem tohoto projektu je umožnit monitorování komunikace na síti a odhalit maligní doménová jména. Toho je dosaženo pomocí získávání informací o daném doménovém jméně z různorodých zdrojů, jako je například RDAP, TLS [19] a GeoLite2. Následná klasifikace pomocí metod strojového učení dokáže určit míru škodlivosti [24].

Služba rovněž obsahuje webové rozhraní, kde si uživatel může zobrazit informace o výsledné klasifikaci a s jakou pravděpodobností se jedná o maligní doménu. Tato klasifikace pravděpodobnosti je dále členěná na podčásti a to: s jakou pravděpodobností se jedná o phishing, malware či DGA (`Domain Generation Algorithm`).

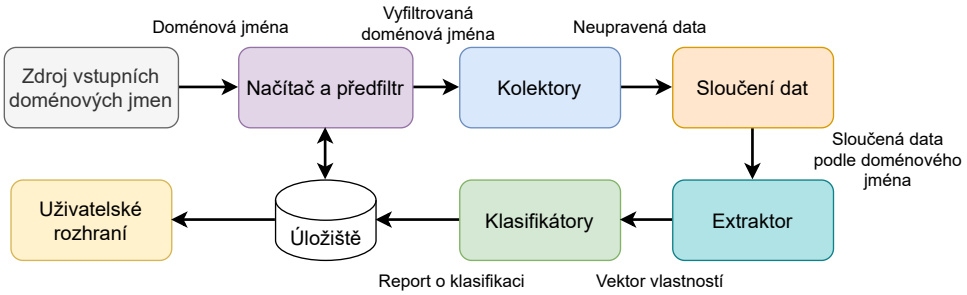
4.1 Architektura systému

Architektura celého systému je rozdělena do několika spolupracujících částí [24]. Aby bylo možné získat klasifikaci daného doménového jména, musí projít těmito částmi tak, jak je vyobrazeno na obrázku 4.1. Tedy ze zdroje vstupních doménových jmen (čímž může být například ruční zadání doménového jména v uživatelském rozhraní) jsou data načtena a vyfiltrovaná na základě nastavených pravidel. K vyfiltrovaným doménám jsou staženy informace z různých zdrojů pomocí kolektorů (fungování kolektorů je blíže popsáno v sekci 4.2.1). Jakmile jsou k dané doméně získány informace ze všech kolektorů, následuje sloučení těchto dat do jednoho objektu. Z tohoto objektu jsou vyextrahována data do podoby vektoru vlastností (*feature vector*), podle nichž následně klasifikátory vyvozují závěr o dané doméně. Výsledek je uložen do databáze a je možné si jej zobrazit v uživatelském rozhraní.

4.2 Zpracování požadavků

DomainRadar využívá k zpracování požadavků systém Apache Kafka[1] [24]. V tomto systému vystupují čtyři zásadní entity: *event*, *topic*, *producer* a *consumer*. Pod slovem *event* se rozumí zpráva, která sestává především z klíče, hodnoty a časové známky. Tyto zprávy jsou zařazeny do určitého tématu (*topic*). *Producer* vytváří zprávy (*event*), které zařadí do existujícího tématu a *consumer* může tyto zprávy číst. Aby však mohl zprávy číst, musí se nejdříve přihlásit k odběru daného tématu. *Consumer* je například kolektor dat z RDAP.

[1] Apache Kafka: https://kafka.apache.org/



Obrázek 4.1: Zjednodušené schéma toku dat v nástroji DomainRadar [24]

4.2.1 Zřetězená kolekce dat

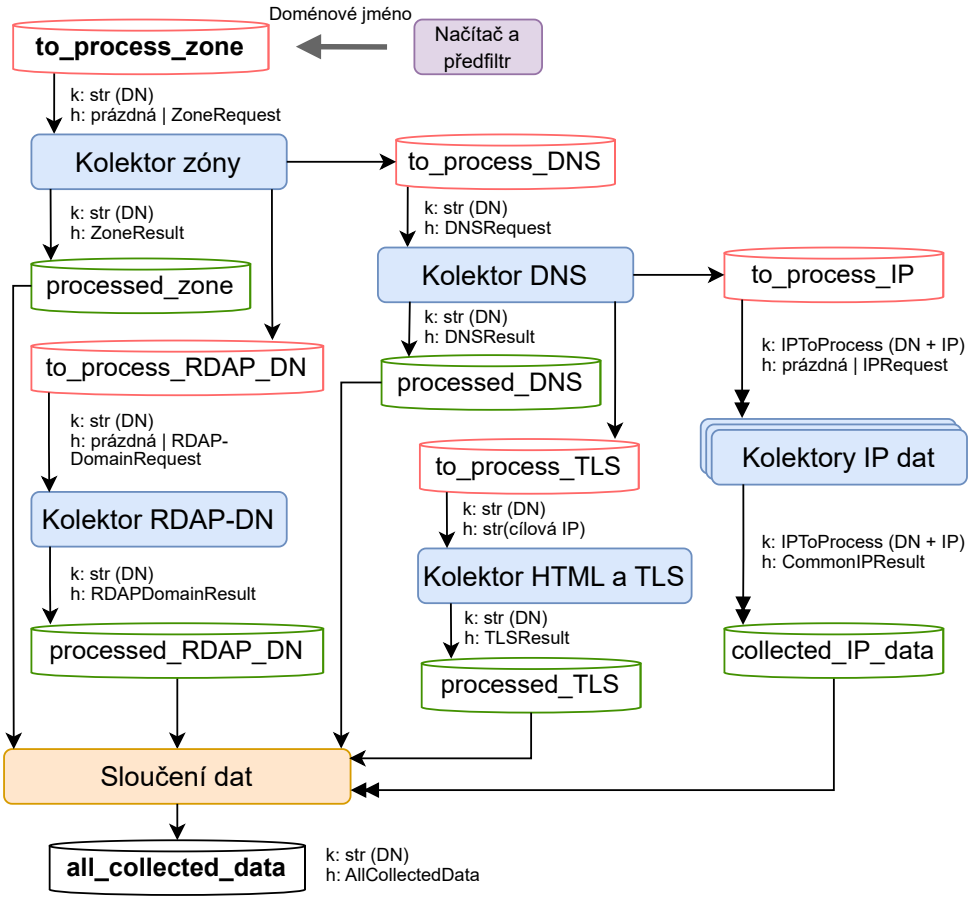
Aby byly k doménovému jménu posbírány informace jednotlivými kolektory, musí být vloženo do tématu `to_process_zone` [24]. Z tohoto tématu je přečteno kolektorem zón a výsledek je uložen do příslušných témat. Takto proces probíhá v několika krocích, dokud nejsou získána všechna data, která jsou následně i spojena. Celý proces zřetězené kolekce dat je zobrazen na schématu 4.2, kde válce s červeným obrysem představují zdrojová témata, odkud si jednotlivé kolektory získávají data, která mají zpracovat. Válce se zeleným obrysem značí cílová témata, kam příslušné kolektory ukládají získaná data o daném doménovém jméně respektive přidružené IP adrese. Kolektory dat jsou v diagramu znázorněny modrým obdélníkem a jejich funkce je následující [24]:

- `Kolektor zóny` zkontorluje, zda doménové jméno získané z tématu `to_process_zone` není `.arpa` a pokusí se danou doménu zpracovat. Údaje o zpracovaní kolektor ukládá do tématu `processed_zone`. V případě úspěchu kolektor navíc uloží doménové jméno do tématu `to_process_DNS` a `to_process_RDAP_DN`.
- `Kolektor RDAP-DN` získává registrační data o daném doménovém jméně pomocí protokolu RDAP (*Registration Data Access Protocol*). V případě neúspěchu je využita služba WHOIS. Výsledek je následně uložen do tématu `processed_RDAP_DN`.
- `Kolektor DNS` slouží k získání DNS záznamů o dané doméně. Všechny získané záznamy jsou uloženy do tématu `processed_DNS`, kde klíčem je doménové jméno. Záznamy typu `A` a `AAAA` jsou taktéž uloženy do tématu `to_process_IP`. Zde však nejsou uloženy jako jeden objekt, ale každá IP adresa je uložena zvlášť. V tomto případě je klíč doménové jméno společně s danou IP adresou.
- `TLS a HTML kolektor` otevírá TCP spojení se vstupní IP adresou na portu 443 a provádí takzvaný TLS handshake. V případě úspěchu ukládá kolektor získaná data, ke kterým například patří použitý protokol. Kolektor dále ukládá i HTML dané stránky.
- Jednotlivé kolektory IP dat jsou z diagramu pro přehlednost vypuštěny, místo toho jsou nahrazeny zástupným blokem. Jejich funkce spočívá v získání dat o dané IP adrese, kterou přečtou z tématu `to_process_IP`. Z jakého zdroje jsou data k IP adrese získána udává kolektor, příkladem využitého zdroje je například NERD[2]. Výsledná data jsou následně uložena do tématu `collected_IP_data`.

[2]NERD: https://nerd.cesnet.cz/

U tématu je taktéž znázorněn datový typ jak pro klíč záznamu (v obrázku zaznačeno jako `k`) tak pro hodnotu (v diagramu označeno jakožto `h`). Výsledná data jsou následně sloučena do jediného objektu, v diagramu tento proces slučování znázorňuje oranžový obdélník. Sloučená data jsou následně uložena do tématu `all_collected_data`, kde klíčem je právě doménové jméno a hodnotou jsou sloučená data. Toto schéma je pro tuto práci důležité, jelikož bude následně rozšířeno o témata i kolektory.

Nutno podotknout, že kolektor ze systému NERD je již v nástroji DomainRadar implementován [24], avšak získaná data nejsou následně nijak použita. V této práci tak budou využita ke klasifikaci společně s novými daty.

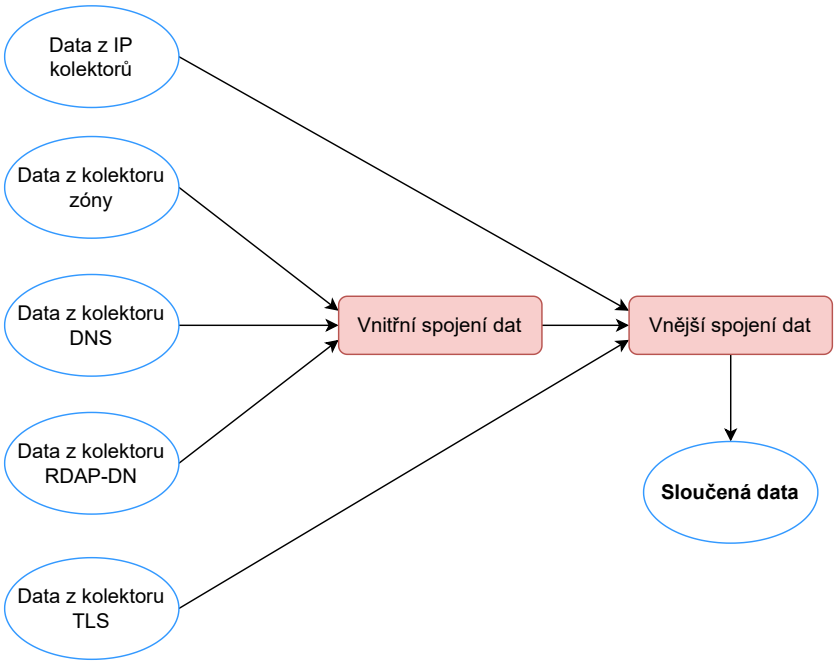


Obrázek 4.2: Schéma zachycující zřetězenou kolekci dat [24]. Válec s červeným obrysem značí zdrojové téma (*topic*). Válec se zeleným obrysem představuje cílové téma.

4.2.2 Sloučení získaných dat

Všechna posbíraná data k jednomu doménovému jménu jsou sloučena do jediného objektu [24] a ten je následně využit k extrakci příznaků. Zjednodušené schéma zachytávající proces slučování posbíraných dat je zobrazen na obrázku 4.3. Sloučená data jsou vyprodukována pouze až jsou získány výsledky ze všech kolektorů. V případě, že alespoň jeden z kolektorů DNS, zóny nebo RDAP-DN nevyprodukuje žádná data, tak vstupní doménové jméno nebude zahrnuto ve výstupu, to značí operace vnitřního spojení dat (*inner join*).

Pokud však kolektor TLS nebo některý z IP kolektorů nevyprodukuje data, tak i přes to budou sloučená data vytvořena – operace vnějšího spojení (*outer join*).



Obrázek 4.3: Zjednodušené schéma procesu slučování dat [24]

4.2.3 Důležitá témata

Následující témata (*topic*) budou v této práci zmíněna, proto zde následuje jejich popis [24].

- `to_process_zone` – vstupní bod, do kterého je jako klíč zadáno doménové jméno, o kterém mají být zjištěny informace.
- `to_process_IP` – do tohoto tématu jsou vloženy IP adresy, které byly zjištěny pomocí DNS z daného doménového jména. Klíčem této zprávy je dvojice doménové jméno a IP adresa.
- `collected_IP_data` – zde jsou uloženy zjištěné informace o dané IP adrese jednotlivými kolektory, které odebírají téma `to_process_IP`.

4.2.4 Uložení získaných dat kolektory do databáze

Získaná data kolektory jsou uložena do perzistentní databáze pomocí nástroje Kafka Connect. Seznam kolektorů je definován v tabulce `Collector`, kde jednotlivé záznamy mají identifikační číslo `id`, unikátní název kolektoru v sloupci `collector` a poté booleovskou hodnotu `is_ip_collector`, která určuje zda daný kolektor získává data o IP adresách. Výsledky jsou následně uloženy do tabulky s názvem `Collection_Result`, která mimo jiné obsahuje referenci na kolektor (sloupec `source_id`), který daná data získal a získaná surová data daným kolektorem pro dané doménové jméno či IP adresu.

4.2.5 Model Result a CommonIPResult

Jeden z klíčových modelů, který byl implementován v projektu DomainRadar, je model `CommonIPResult`. Tento model je důležitý pro tuto práci, jelikož bude využíván vytvořenými kolektory reputačních systému, které klasifikují entity na základě IP adres. Druhým klíčovým modelem je `Result`, který bude v této práci rozšířen dalším modelem. Z tohoto důvodu jsou modely níže popsány a uvedeny ve výpisu 4.1.

Model `CommonIPResult` rozšiřuje bázový model `Result` [24]. Tudíž výsledek operace kolektoru nebude obsahovat pouze samotná data, která vybere a uloží daný kolektor, ale taktéž i stavový kód (`statusCode`). Tento kód představuje výsledek operace kolektoru, kde hodnota 0 značí úspěch. Nenulové hodnoty značí neúspěch buď generického charakteru (např. vypršení časového limitu) či specifického charakteru pro daný kolektor. Položka `error` může obsahovat chybové hlášení, ale pouze v případě, když operace skončila neúspěchem. `lastAttempt` obsahuje časové razítko, kdy byla operace dokončena.

Položka `collector` v modelu `CommonIPResult` obsahuje identifikační název kolektoru, který výsledná data vyprodukoval [24]. Objekt `data` je generický objekt, který si každý kolektor předefinuje specifickým modelem, jenž obsahuje položky, které daný kolektor poskytuje.

```
class Result:
    statusCode: int
    error: str | None
    lastAttempt: timestamp

class CommonIPResult(Result):
    collector: str
    data: object | None
```

Výpis 4.1: Model Result a CommonIPResult [24]

4.3 Klasifikace doménových jmen

Úkolem klasifikační části systému je na základě získaných dat určit míru rizika, kterou daná doména představuje. Výsledkem je tedy vygenerovaný klasifikační report obsahující ohodnocení na základě několika klasifikátorů. Klasifikátory jsou rozděleny do třech skupin:

- klasifikátory domén generovaných DGA,
- klasifikátory domén sloužících k phishingu,
- klasifikátory domén k šíření malware.

Výsledky klasifikátorů jsou poté agregovány do třech hodnot právě podle těchto skupin. Celková míra rizika je následně určena pomocí rozhodovací neuronové sítě a pomocí sady heuristik. V poslední fázi je vygenerován report ve formátu JSON, který obsahuje celkovou míru hrozby a výsledky jednotlivých klasifikátorů.

Kapitola 5

Popis volně dostupných reputačních systémů

Následující sekce popisují vybrané, volně dostupné, reputační systémy, jejich vlastnosti a cenu.

5.1 abuseipdb.com

Služba AbuseIPDB umožňuje pomocí webového rozhraní klasifikovat doménová jména, IPv4 i IPv6 adresy [2]. Služba dále nabízí rozhraní API, pomocí kterého lze však získávat informace pouze o validních IPv4 nebo IPv6 adresách.

Vhodné informace poskytnuté rozhraním API

Pomocí rozhraní API této služby lze získat mnoho informací. Mezi vhodné informace pro mou práci však patří tyto [2]:

- `isWhitelisted` – booleovská hodnota značící, zda daná IP adresa je obsažena na alespoň jedné bílé listině této služby,
- `abuseConfidenceScore` – nebinární hodnota v rozsahu 0 až 100, která udává v procentech, jak moc si je služba jistá, že se jedná o maligní IP adresu,
- `isTor` – binární hodnota udávající, zda se jedná o uzel v síti Tor,
- `totalReports` – počet nahlášení, která služba dostala o dané IP adrese od uživatelů,
- `numDistinctUsers` – kolik unikátních uživatelů danou IP adresu nahlásilo,
- `lastReportedAt` – časové razítko ve formátu stanoveným ve standardu RFC 3339 [22]
- `reports` – seznam jednotlivých nahlášení, které mimo jiné obsahují časové razítko nahlášení, kategorizovaný důvod nahlášení a krátký komentář.

Ceník služby

Služba je zdarma při využití webového rozhraní a dotazování se pomocí něj. Pro dotazování se pomocí rozhraní API poskytuje abuseipdb.com následující balíčky, které jsou popsány v tabulce 5.1 [2] společně s cenou za měsíc.

Název balíčku	Počet dotazů za den	Cena za měsíc
Individual (*bez verifikace*)	1 000	Zdarma
Individual (*po verifikaci*)	3 000	Zdarma
Basic	10 000	$25
Premium	50 000	$99

Tabulka 5.1: Ceník služby abuseipdb.com platný ke dni 27. listopadu 2024

5.2 radar.cloudflare.com

Systém Cloudflare Radar poskytuje vyhledávání klasifikačních informací týkajících se doménových jmen a IPv4 adres [12]. Vyhledávání je umožněno jednak pomocí webového rozhraní, ale i pomocí služby API, kterou systém nabízí.

Vhodné informace poskytnuté rozhraním API

Jelikož v odpovědi této služby převažují statistické údaje o doménovém jméně, respektive IP adresách, je pro tuto práci vhodnou informací pouze jedna položka:

- `malicious` – booleovská hodnota udávající, zda daná entita je považována za maligní [12].

V případě, že systém nemá o dané entitě žádný záznam, výsledky získané pomocí API jsou prázdné [12]. Tudíž neobsahují ani položku `malicious` a lze ji tak rozlišit od těch, o kterých záznam má, ale nepovažuje je za maligní.

Ceník služby

Přístup k webovému rozhraní a API je poskytován zdarma. Pro využití API služby je však nutné si vytvořit uživatelský účet. Rozhraní API je následně limitováno na 1 200 dotazů za 5 minut na uživatelský účet, nikoliv na API klíč [12].

5.3 criminalip.io

Reputační systém CriminalIP nabízí ohodnocení IP adresy přes API i webové rozhraní. Umí však pouze ohodnocovat IPv4 adresy nikoliv IPv6 adresy [4]. Pro využití rozhraní API je však potřeba mít vytvořen uživatelský účet.

Vhodné informace poskytnuté rozhraním API

Tato služba poskytuje prostřednictvím API poměrně velkou část reputačních a klasifikačních informací o dané IPv4 adrese. Pro tuto práci jsou však přínosné převážně tyto informace [4]:

- `issues` – problémy zjištěné službou k dané IP adrese:
  - `is_scanner` – booleovská hodnota značící, zda daná IPv4 adresa je považována za síťový skener,

  - `is_snort` – booleovská hodnota udávající, zda IP adresa byla detekována systémem Snort.

- `score` – skóre přidělené službou dané IPv4 adrese:
  - `inbound` – skóre udávající potencionální hrozbu směrem do sítě,
  - `outbound` – skóre udávající potencionální hrozbu směrem ze sítě.

Hodnoty, kterých může skóre `inbound` a `outbound` nabývat, jsou následující [4]:

- `Safe` – IP adresa je zcela legitimní,
- `Low` – IP adresa je skoro legitimní,
- `Moderate` – IP adresa jeví maligní příznaky,
- `Dangerous` – pravděpodobně je IP adresa maligní,
- `Critical` – velká pravděpodobnost, že IP adresa je maligní.

Ceník služby

Typy balíčků, které služba poskytuje, jsou popsány v tabulce 5.2.

Název balíčku	Dotaz/měsíc	Cena za měsíc
Free	50	Zdarma
Lite	70 000	$65
Medium	700 000	$349
Pro	1 000 000	$1 089

Tabulka 5.2: Ceník služby criminalip.io platný ke dni 22. března 2025 [4]

5.4 fortiguard.com

Fortiguard, přesněji služba AntiSpam, nabízí jednak webové rozhraní, pomocí kterého lze identifikovat, zda tento reputační systém označil IP adresu, URL, adresu elektronické pošty či hash za spam, a jednak nabízí i API, pomocí kterého lze zjistit stejné informace [14]. Služba poskytnuta skrze webovou stránku a rozhraní API je poskytována bezplatně a není třeba žádného uživatelského účtu.

Vhodné informace poskytnuté rozhraním API

Tato služba poskytuje jen a pouze jednu informaci a tou je položka `spam`, která nabývá booleovské hodnoty [14]. Položka `spam` udává, zda daná entita byla službou Fortiguard označena za spam.

5.5 safebrowsing.google.com

Reputační systém Google Safe Browsing lze využít pouze skrze dostupné rozhraní API [15]. Toto rozhraní dovoluje klasifikovat doménová jména a IPv4 adresy.

Vhodné informace poskytnuté rozhraním API

Jediná informace, která je užitečná pro tuto práci, je položka `threatType`. Ta udává, za jaký typ hrozby je daná entita považována službou [15]. Položka `threatType` může nabývat jedné z těchto hodnot:

- `THREAT_TYPE_UNSPECIFIED` – jedná se o hrozbu, avšak služba není schopna klasifikovat typ hrozby,
- `MALWARE` – stránka hostuje maligní software, který má uživateli uškodit,
- `SOCIAL_ENGINEERING` – entita využívající taktik, jako je například phishing,
- `UNWANTED_SOFTWARE` – stránka distribuuje software, který nemusí být přímo maligní, ale negativně ovlivňuje uživatelův zážitek,
- `POTENTIALLY_HARMFUL_APPLICATION` – aplikace či software, který může uživateli uškodit.

V případě, že ani jeden tento typ není obsažen v odpovědi z API, tak služba předpokládá, že se jedná o benigní entitu, nebo nemá o ní žádné záznamy [15].

Ceník služby

Služba poskytuje všechny své verze rozhraní API zcela zdarma [15]. Vyjímkou je pouze komerční užití této služby, v takovém případě je nutno využít jinou odnož služby.

5.6 greynoise.io

Systém GreyNoise poskytuje webové prostředí i rozhraní API [16]. Tato služba umožňuje zadávat ke klasifikaci pouze IPv4 adresy.

Vhodné informace poskytnuté rozhraním API

Vhodné informace, které poskytuje API, jsou následující [16]:

- `noise` – booleovská hodnota udávající, zda daná IP adresa skenovala senzor sítě GreyNoise,
- `riot` – booleovská hodnota, která značí, zda daná IP adresa je součástí RIOT datasetu,
- `classification` – klasifikace dané IP adresy (může nabývat jedné z těchto hodnot: `benign`, `unknown`, `malicious`).

Ceník služby

Komunitní verze API je poskytovaná zdarma [16] a bez potřeby vytvoření uživatelského účtu. Služba je limitovaná na 50 dotazů za týden a limity jsou sdíleny napříč webovým prostředím a rozhraním API. V případě dotazování bez vytvořeného uživatelského účtu, je služba limitovaná více.

5.7 hybrid-analysis.com

Reputační systém Hybrid Analysis nabízí jednak webové rozhraní a jednak rozhraní API, které umožňuje zjistit informace o doménovém jméně, IP adrese a souboru [20].

Vhodné informace poskytnuté rozhraním API

API poskytuje tyto následující vhodné informace [20]:

- `result` – seznam výsledků k dané entitě, o kterých má služba informace. Každý výsledek obsahuje tyto hodnoty [20]:
  - `verdict` – řetězec udávající verdikt (možné verdikty: `no specific threat`, `suspicious`, `malicious`, `whitelisted`),
  - `threat_score` – `null` nebo číselná hodnota od 0 po 100, kde 100 je nejhorší možné skóre.

Ceník služby

Služba je poskytována zcela zdarma. Pro využití je však potřeba mít vytvořený uživatelský účet. Rozhraní API je limitováno na 200 dotazů za minutu a 2 000 dotazů za hodinu.

5.8 nerd.cesnet.cz

Služba NERD poskytuje jednak vyhledávání pomocí webového rozhraní a jednak i pomocí rozhraní API. Vyhledávání je možné primárně pomocí IP adres, avšak API nabízí i možnost vyhledávat pomocí doménového jména. Výhodou rozhraní API je možnost dávkového dotazování, kdy jeden dotaz obsahuje více IP adres, o kterých je žádoucí získat informace [8, 7].

Vhodné informace poskytnuté rozhraním API

Mezi vhodné informace získané pomocí API patří [8]:

- `bl` – seznam černých listin, na kterých je daná IP adresa uvedena,
- `rep` – hodnota od 0 do 1, udávající reputační skóre, kde 0 je nejlepší a 1 nejhorší hodnocení daného subjektu.

Ceník služby

Využití poskytovaných služeb je zdarma a to jak webového tak API rozhraní. Pro získání API klíče je však nutné se zaregistrovat pomocí účtu vedeného u oprávněné instituce [8]. Rozhraní API je rovněž limitováno na jeden dotaz za sekundu.

5.9 opentip.kaspersky.com

Reputační systém Opentip Kaspersky nabízí webovou aplikaci i rozhraní API, pomocí kterých je možné provést kontrolu klasifikace doménového jména, IPv4 adresy a souboru [6].

Vhodné informace poskytnuté rozhraním API

Informace poskytnuté rozhraním API, které jsou pro tuto práci užitečné jsou následující [6]:

- `Zone` – zóna, do které byla entita službou zařazena,
- `Categories` – kategorie, které byly dané entitě službou přiřazeny,
- `CategoriesWithZone` – kategorie, které byly dané entitě službou přiřazeny, a zóna dané kategorie.

Zóna může nabývat následujících hodnot [6]:

- `Red` – entita je považována za nebezpečnou,
- `Orange` – entitě by se nemělo důvěřovat,
- `Yellow` – entita je považována za adware,
- `Grey` – entita nebyla kategorizována,
- `Green` – u entity nebyly detekovány žádné hrozby.

Ceník služby

Služba je poskytovaná bezplatně a to jak webový analyzátor tak i rozhraní API. U tohoto systému však nikde není uveden limit API, dle mých zkušeností to však je přibližně 2 000 dotazů za den.

5.10 projecthoneypot.org

Reputační systém Project Honeypot nabízí k dispozici webové rozhraní, pomocí kterého lze zjistit informace jen o IPv4 adresách [31]. API rozhraní služba neposkytuje, ale nabízí se možnost stránku s reportem IP adresy stáhnout a následně ji zpracovat.

Vhodné informace poskytnuté službou

Služba má čtyři skupiny, do kterých může být IP adresa zahrnuta [31]: `Harvesters`, `Spam Servers`, `Dictionary Attackers` a `Comment Spammers`. Přičemž jedna IP adresa může být zahrnuta do vícero těchto skupin naráz.

5.11 pulsedive.com

Služba Pulsedive poskytuje webové i API rozhraní pro získání klasifikačních hodnot IP adresy a doménového jména [25].

Vhodné informace poskytnuté rozhraním API

Reputační systém poskytuje následující ohodnocení entity [25], které je přinosné pro tuto práci: `risk`, `risk_recommended` a `manualrisk`. Pulsedive v dokumentaci neuvádí, co přesně se za těmito položkami skrývá. `risk` a `risk_recommended` může nabývat hodnot: `none`, `low`, `medium`, `high`, `critical`, `retired` a `unknown`. Položka `manualrisk` nabývá buď hodnoty `1` nebo `0`.

Ceník služby

Služba je zdarma při využití webového rozhraní a dotazování se pomocí něj. Pro dotazování se pomocí rozhraní API poskytuje systém balíčky popsané v tabulce 5.3.

Název balíčku	Dotaz/sekundu	Dotaz/den	Dotaz/měsíc	Cena za měsíc
Free Account	1	50	500	Zdarma
Team	Bez omezení	Bez omezení	25 000	$300
Business	Bez omezení	Bez omezení	100 000	$900

Tabulka 5.3: Ceník služby pulsedive.com platný ke dni 22. března 2025 [25]

5.12 threatfox.abuse.ch

Systém Threatfox nabízí webovou stránku a rozhraní API, pomocí kterých lze zjistit informace o doménovém jméně, IPv4 adrese a hashi [1] (služba je hromadně označuje jako `IOC`). Tento systém se zaměřuje hlavně na detekci malwaru a umožňuje uživatelům nahlašovat `IOC` službě do databáze. Služba taktéž expiruje `IOC` starší než šest měsíců, aby se systém vyhnul nežádoucímu falešnému označení `IOC` za maligní.

Vhodné informace poskytnuté rozhraním API

Informace, které mohou být užitečné pro tuto práci, a které jsou uvedeny v odpovědi API [1], jsou následující:

- `threat_type` – řetězcové označení typu hrozby (například *botnet_cc*),
- `confidence_level` – číselná hodnota udávající míru jistoty systému od 0 do 100.

Ceník služby

Webové rozhraní i rozhraní API je poskytováno zcela zdarma [1]. Pro využití API je však potřeba si vytvořit bezplatný účet na stránkách systému.

5.13 urlvoid.com

Služba UrlVoid poskytuje webové rozhraní, pomocí kterého lze vyhledávat informace pouze o doménových jménech. Neposkytuje však rozhraní API k získání informací, avšak je možné stáhnout stránku s reputací daného subjektu a tu následně zpracovat samostatně. Služba poskytuje ohodnocení na základě 39 skenovacích nástrojů třetí strany [23].

Vhodné informace poskytnuté službou

Jedinou vhodnou informací je položka `Detections Counts`, která udává, kolik skenovacích nástrojů odhalilo nějaký problém u dané webové domény. Tato hodnota se pohybuje v rozmezí od 0 do 39 (popřípadě `null`, pokud o dané doméně služba nevede žádné záznamy) [23].

5.14 virustotal.com

Služba VirusTotal rovněž poskytuje API rozhraní i webové rozhraní, kterým je možné zjistit ohodnocení IP adresy, webové domény, ale i souboru. VirusTotal poskytuje mnoho užitečných informací k vyhledávané entitě [32]. K určení ohodnocení je využito 94 analyzačních nástrojů. Výhodou je, že nástroje neoznačují danou entitu pouze jako bezpečná/škodlivá, ale hodnocení je rozděleno do vícero kategorií. Reputaci mohou taktéž přidělovat registrovaní uživatelé na základě hodnocení. Toto hodnocení je postaveno na systému hlasování, kde uživatel danou entitu označuje jako bezproblémovou/problémovou. Výsledná hodnota je dostupná jakožto komunitní skóre.

Vhodné informace poskytnuté rozhraním API

Mezi vhodné informace získané pomocí API patří [32]:

- `last_analysis_stats` – výsledky poslední analýzy, tyto výsledky jsou rozděleny do těchto podkategorií:
  - `malicious` – počet analyzátorů, které daný subjekt označilo jako maligní,
  - `suspicious` – počet analyzátorů, které daný subjekt označilo jako podezřelý,
  - `undetected` – počet analyzátorů, které u daného subjektu neobjevily žádný problém,
  - `harmless` – počet analyzátorů, které daný subjekt označilo jako neškodný,
  - `timeout` – počet analyzátorů, které nedokázaly odpovědět včas.
- `reputation` – číselná hodnota odpovídající kladnému počtu hlasů od registrovaných uživatelů,
- `total_votes` – počet hlasů registrovaných uživatelů v dané kategorii
  - `harmless` – počet hlasů od registrovaných uživatelů, kteří danou entitu označili jako neškodnou,
  - `malicious` – počet hlasů od registrovaných uživatelů, kteří danou entitu označili jako maligní.
- `last_analysis_result` – seznam výsledků od jednotlivých analyzačních nástrojů.

Ceník služby

Služba je zdarma při využití webového rozhraní a dotazování se pomocí něj. Pro dotazování se pomocí rozhraní API poskytuje virustotal.com balíčky popsané v tabulce 5.4.

Název balíčku	Dotaz/den	Dotaz/minutu	Cena za měsíc
Public	500	4	Zdarma
Premium	Bez omezení	Bez omezení	Nutno kontaktovat VirusTotal

Tabulka 5.4: Ceník služby virustotal.com platný ke dni 27. listopadu 2024 [32]

5.15 Komparativní srovnání reputačních systémů

V tabulce 5.5 jsou komparativně srovnány jednotlivé vybrané reputační systémy. Je zde uvedeno, zda daný reputační systém dokáže ohodnotit IP adresy, přesněji zda dokáže ohodnotit IPv4 adresy a IPv6 adresy. Dále je zde uvedeno, zda dokáže ohodnotit doménová jména (v tabulce zkráceně označeno jako *DN*). Jednou z hlavních vlastností vybraných reputačních systémů je, zda daný systém dokáže rozlišit benigní doménová jména od těch, o kterých nemá systém žádné informace. Tuto vlastnost v tabulce zachycuje sloupec *Rozlišitelnost žádných informací*. Dále je zde ke každému reputačnímu systému uvedeno omezení počtu provedených dotazů, který daný systém má pro verzi dostupného plánu, který je poskytován službou zdarma.

Na první pohled je zřejmé, že některé reputační systémy mají velké omezení počtu provedených dotazů pro získání reputačních dat o entitě z dané služby. I přes to, že podporují ohodnocení vícero typů entit a dokážou rozlišit benigní entity od těch, o kterých nemají žádné informace, jsou pro celkovou klasifikaci méně přínosné, jelikož tyto limity by mohly nežádoucím způsobem ovlinit klasifikaci a to z důvodu absence těchto dat při ohodnocování daného doménového jména.

Reputační systém	**Podpora**			**Rozlišitelnost žádných informací**	**Omezení dotazů**
	IPv4	**IPv6**	**DN**		
abuseipdb.com	Ano	Ne	Ne	Ne	1 000/den
radar.cloudflare.com	Ano	Ne	Ano	Ano	1 200/5 minut
criminalip.io	Ano	Ne	Ne	Ano	50/měsíc
fortiguard.com	Ano	Ano	Ano	Ne	Žádné
safebrowsing.google.com	Ano	Ne	Ano	Ne	10 000/den
greynoise.io	Ano	Ne	Ano	Ano	50/týden
hybrid-analysis.com	Ano	Ano	Ano	Ano	2 000/hodinu
nerd.cesnet.cz	Ano	Ne	Ano	Ano	1/sekundu
opentip.kaspersky.com	Ano	Ne	Ano	Ano	2 000/den
projecthoneypot.org	Ano	Ne	Ne	Ano	Žádné
pulsedive.com	Ano	Ano	Ano	Ano	50/den
threatfox.abuse.ch	Ano	Ne	Ano	Ano	Žádné
urlvoid.com	Ne	Ne	Ano	Ano	Žádné
virustotal.com	Ano	Ano	Ano	Ano	500/den

Tabulka 5.5: Komparativní srovnání vybraných reputačních systémů

Kapitola 6

Experimentální zhodnocení dostupných reputačních systémů

K předběžnému zjištění úspěšnosti detekce maligních doménových jmen jsem použil sadu dat, která byla vytvořena v rámci projektu FETA [18]. Dále jsem využil službu PhishTank[1] pro phishingové domény a službu ThreatFox[2] pro domény označené jako malware.

Z těchto externích zdrojů jsem sestavil tři oddělené datové sady, každou zaměřenou na jednu z kategorií: benigní, phishing, malware. Každá datová sada obsahuje 100 doménových jmen, přičemž výběr byl proveden na základě konkrétního kritéria. Doménové jméno jsem zařadil do testovací sady pouze v případě, že měla uveden alespoň jeden DNS záznam typu A. Tento přístup byl zvolen z toho důvodu, že některé dostupné reputační systémy klasifikují entity na základě IPv4 adresy, nikoli přímo na základě doménového jména. Díky tomuto postupu je zajišťeno, že všechny testované reputační systémy mají stejné množství vstupních dat. Tento přístup taktéž umožňuje objektivní porovnání výsledků napříč všemi reputačními systémy.

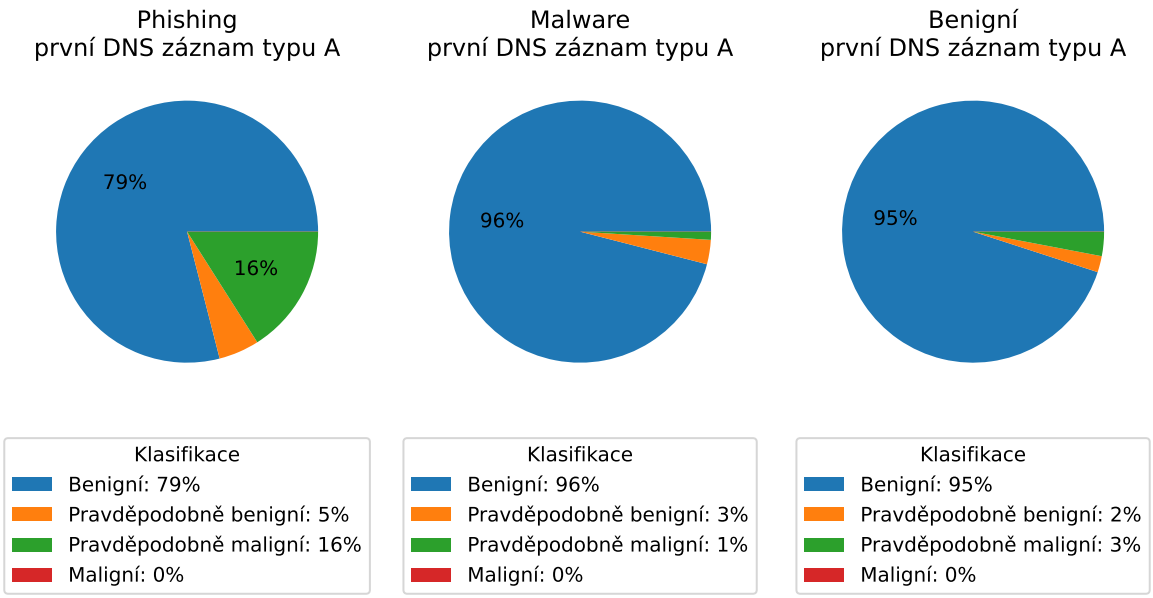
6.1 abuseipdb.com

Na obrázku 6.1 je zobrazena klasifikace službou abuseipdb.com jednotlivých IPv4 adres. Jelikož služba neudává hladiny, podle kterých se entity zařadí mezi danou klasifikaci, bylo potřeba najít vhodné hranice, kdy co nejvíce maligních IP adres je označeno za maligní, ale na druhou stranu benigní IP adresy nesmí být falešně označeny za maligní. Experimentováním jsem došel k těmto následujícím hladinám:

- **Benigní** – `abuseConfidenceScore` $= 0$,
- **Pravděpodobně benigní** – $1 \leq$ `absueConfidenceScore` ≤ 19,
- **Pravděpodobně maligní** – $20 \leq$`absueConfidenceScore` ≤ 49,
- **Maligní** – $50 \leq$ `absueConfidenceScore` ≤ 100.

Zde je na první pohled zřejmé, že i přes to, že se jedná o maligní doménové jméno, byla IP adresa ohodnocena jako benigní. To je způsobeno tím, že pokud služba nemá o dané entitě záznam, automaticky ji přiděluje reputační skóre 0. Což poté může zkreslovat výsledky.

[1] PhishTank: https://phishtank.org/
[2] ThreatFox: https://threatfox.abuse.ch/



Obrázek 6.1: Klasifikace 100 phishingových, malwarových a benigních IPv4 adres službou abuseipdb.com

6.2 radar.cloudflare.com

Obrázek 6.2 zachycuje ohodnocení doménových jmen službou Cloudflare Radar po jednotlivých kategoriích. Vyobrazená klasifikace je rozhodnuta na základě položky `malicious` získané z odpovědi API a to dle následujících hodnot:

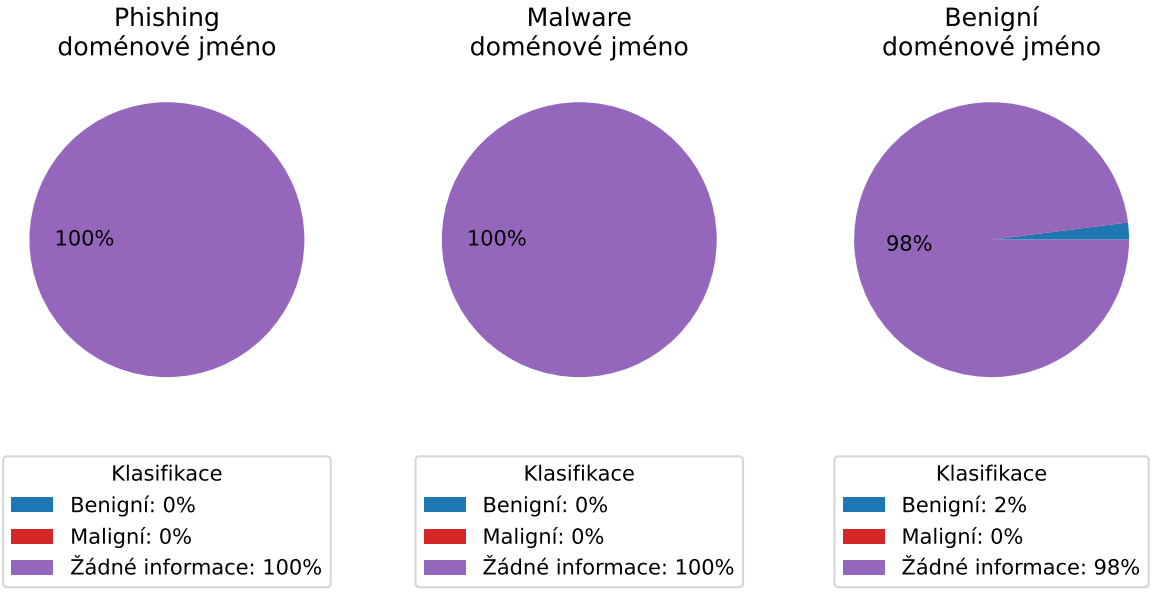
- **Benigní** – `malicious` = `False`,
- **Maligní** – `malicious` = `True`,
- **Žádné informace** – `malicious` není obsažena v odpovědi.

6.3 criminalip.io

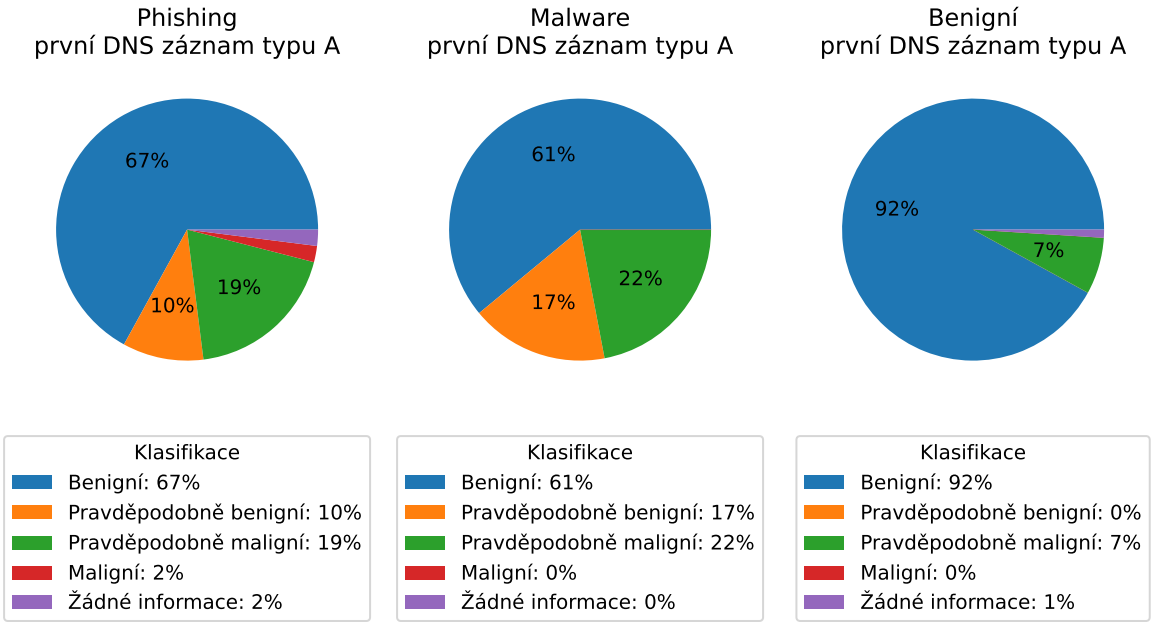
Graf 6.3 vyobrazuje klasifikaci IPv4 adres ze všech tří skupin. Klasifikace na tomto grafu je rozhodnuta pomocí položky `outbound` v sekci `score`, která je obsažena v odpovědi API. Hladiny byly vybrány na základě textového popisu jednotlivých typů skóre, které služba uvádí, a jsou následující:

- **Benigní** – `outbound` = `Safe`,
- **Pravděpodobně benigní** – `outbound` = `Low`,
- **Pravděpodobně maligní** – `outbound` = `Moderate` nebo `Dangerous`,
- **Maligní** – `outbound` = `Critical`,
- **Žádné informace** – získaná odpověď je prázdná.

Popis ohodnocení `Moderate` a `Dangerous` je velice podobný, proto je sloučen do jedné skupiny.



Obrázek 6.2: Klasifikace 100 phishingových, malwarových a benigních doménových jmen službou radar.cloudflare.com

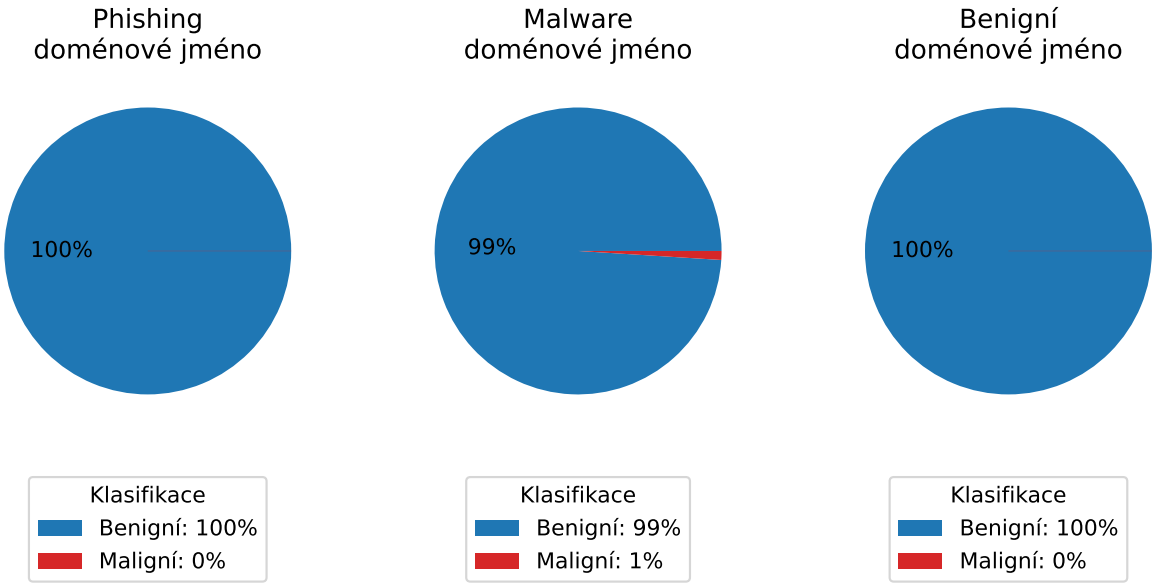


Obrázek 6.3: Klasifikace 100 phishingových, malwarových a benigních IPv4 adres službou criminalip.io

6.4 fortiguard.com

Obrázek 6.4 zobrazuje klasifikaci doménových jmen ze všech tří skupin doménových jmen službou Fortiguard. Klasifikace je rozhodnuta pomocí položky `spam` v získané odpovědi pomocí API tohoto systému. Skupiny pro klasifikaci zobrazené v grafu jsou stanoveny následovně:

- **Benigní** – `spam` = `False`,
- **Malignní** – `spam` = `True`,



Obrázek 6.4: Klasifikace 100 phishingových, malwarových a benigních doménových jmen službou fortiguard.com

6.5 safebrowsing.google.com

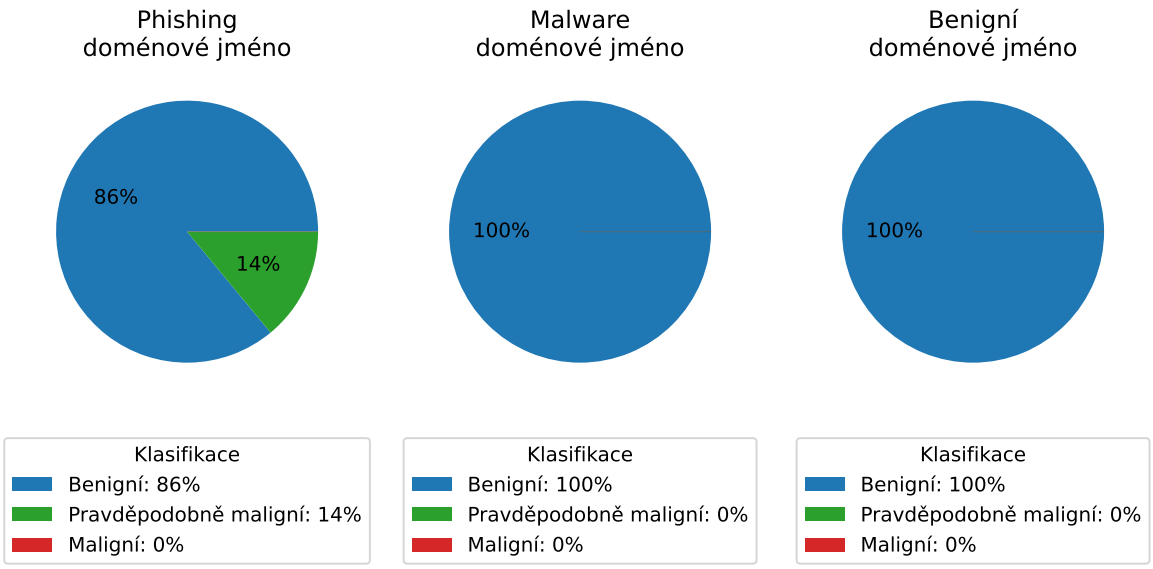
Na obrázku 6.5 je zobrazena klasifikace doménových jmen službou Google Safe Browsing. Klasifikace je rozhodnuta na základě sečtení všech typů hrozeb získaných ze systému viz sekce 5.5 (níže v této sekci souhrnně označeno jako `součet hrozeb`). Hladiny pro klasifikaci jsou zvoleny následující:

- **Benigní** – `součet hrozeb` = 0.0,
- **Pravděpodobně malignní** – `součet hrozeb` = 1,
- **Malignní** – `součet hrozeb` = 2.

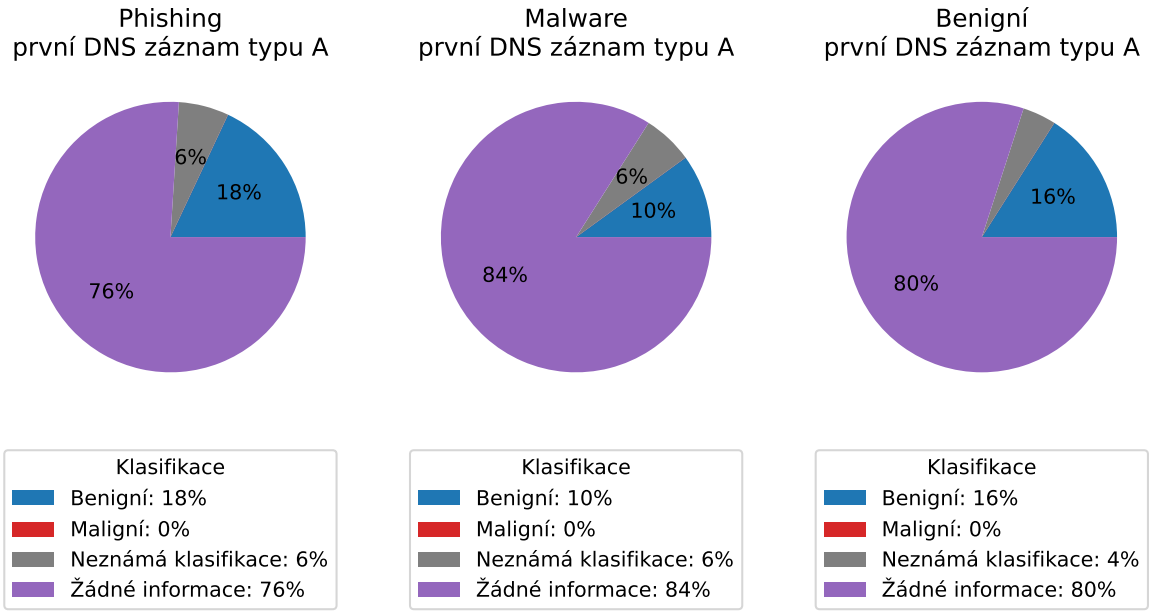
6.6 greynoise.io

Obrázek 6.6 zachycuje klasifikaci IPv4 adres systémem Greynoise. Hladiny byly vybrány na základě položky `classification` získané pomocí rozhraní API služby a pomocí jejich hodnoty. Hladiny byly tak zvoleny následovně:

- **Benigní** – `classification` = `benign`,
- **Malignní** – `classification` = `malicious`,
- **Neznámá klasifikace** – `classification` = `unknown`,
- **Žádné informace** – získaná odpověď je prázdná.



Obrázek 6.5: Klasifikace 100 phishingových, malwarových a benigních doménových jmen službou safebrowsing.google.com

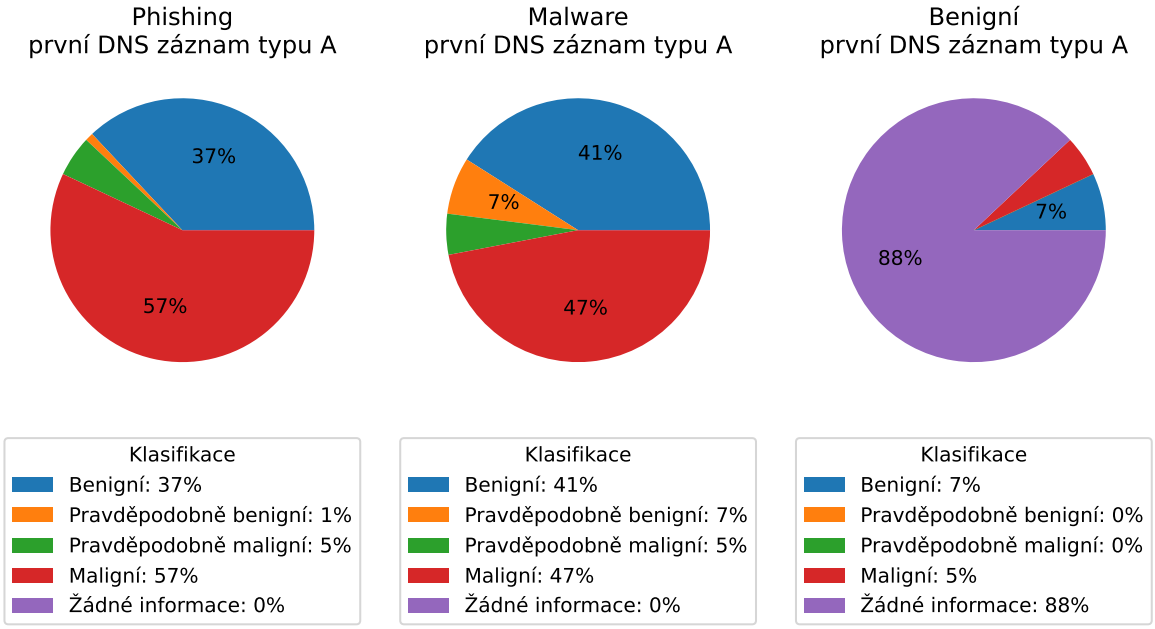


Obrázek 6.6: Klasifikace 100 phishingových, malwarových a benigních IPv4 adres službou greynoise.io

6.7 hybrid-analysis.com

Klasifikace IPv4 adres službou Hybrid Analysis je vyobrazena na obrázku 6.7, který obsahuje ohodnocení všech zmíněných kategorií. Jelikož služba při ohodnocení IP adresy vrací seznam výsledků, je použita průměrná hodnota všech položek `threat_score` v celém seznamu (v této sekci bude souhrnně označeno jako `průměrné skóre`). Rozdělení do tříd je v obrázku 6.7 provedeno následovně:

- **Benigní** – $0 \leq$ `průměrné skóre` ≤ 20,
- **Pravděpodobně benigní** – $20 <$ `průměrné skóre` ≤ 40,
- **Pravděpodobně maligní** – $40 <$ `průměrné skóre` ≤ 60,
- **Maligní** – $60 <$ `průměrné skóre` ≤ 100,
- **Žádné informace** – skóre není uvedeno, nebo seznam výsledků je prázdný.

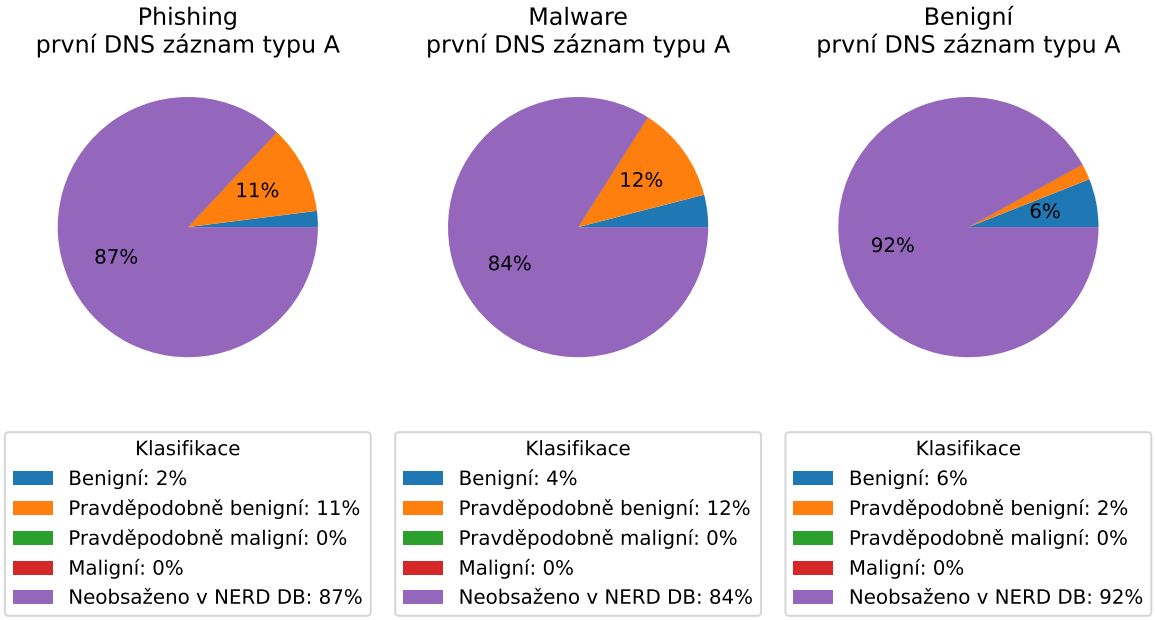


Obrázek 6.7: Klasifikace 100 phishingových, malwarových a benigních IPv4 adres službou hybrid-analysis.com

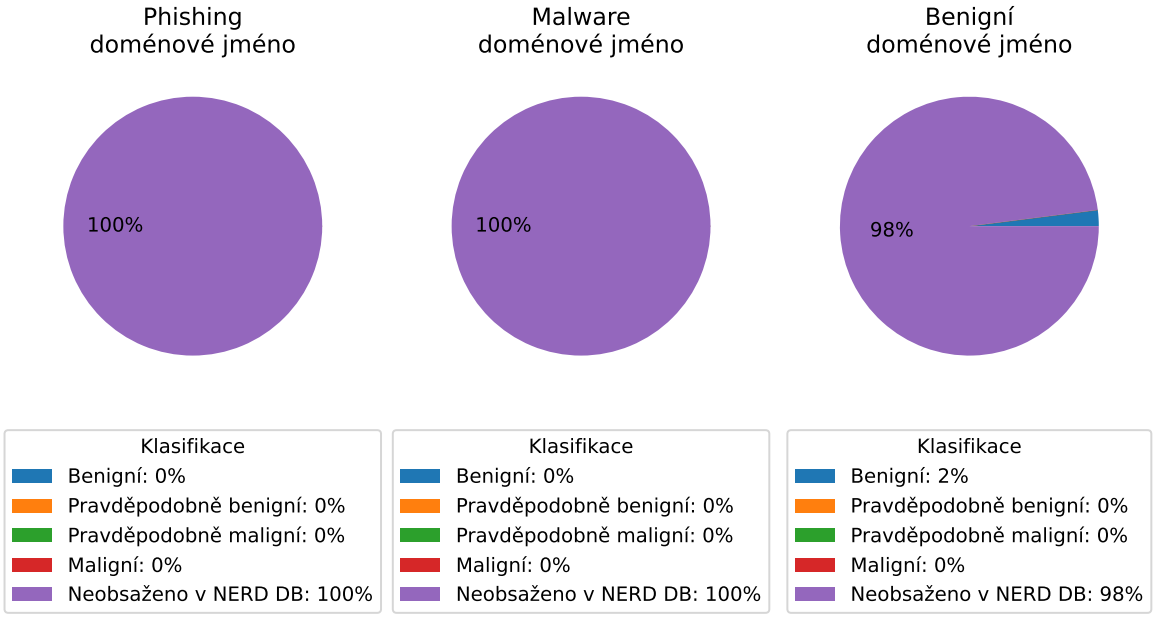
6.8 nerd.cesnet.cz

Obrázek 6.8 vyobrazuje klasifikaci IPv4 adres z daných datových sad. Na obrázku 6.9 je zobrazená klasifikace doménových jmen. Poměrně velká většina těchto doménových jmen nemá v reputačním systému žádný záznam, jelikož vyhledávání pomocí doménových jmen není primárním účelem tohoto systému. Výhodou této služby je odlišení benigních adres od těch, o kterých služba nemá žádné informace. Zvolené hladiny pro klasifikaci jsou takovéto:

- **Benigní** – `rep` $= 0.0$,
- **Pravděpodobně benigní** – $0.0 <$ `rep` ≤ 0.2,
- **Pravděpodobně maligní** – $0.2 <$ `rep` ≤ 0.8,
- **Maligní** – $0.8 <$ `rep` ≤ 1.



Obrázek 6.8: Klasifikace 100 phishingových, malwarových a benigních IPv4 adres službou nerd.cesnet.cz

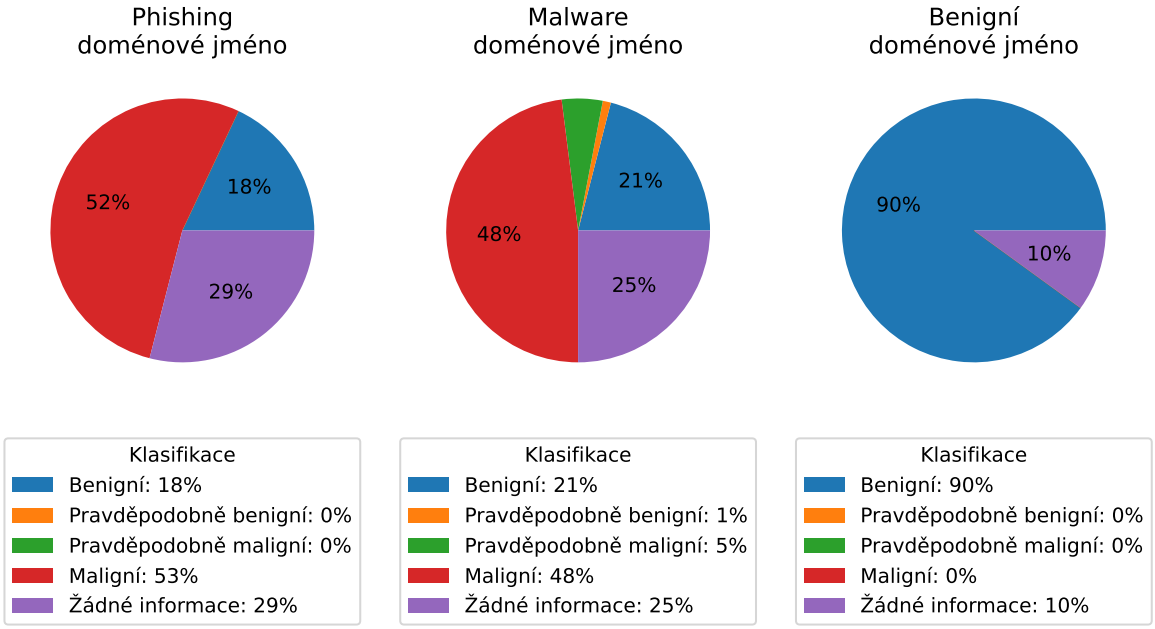


Obrázek 6.9: Klasifikace 100 phishingových, malwarových a benigních doménových jmen službou nerd.cesnet.cz

6.9 opentip.kaspersky.com

Provedená klasifikace doménových jmen, ze všech tří kategorií, při užití služby Opentip Kaspersky je zobrazena na obrázku 6.10. Jednotlivé kategorie klasifikace byly rozděleny podle položky `Zone`, do které byla daná doména zařazena službou (nikoliv `Zone` kategorie) a to dle následujícího kritéria:

- **Benigní** – `Zone` = `Green`,
- **Pravděpodobně benigní** – `Zone` = `Yellow`,
- **Pravděpodobně maligní** – `Zone` = `Orange`,
- **Maligní** – `Zone` = `Red`,
- **Žádné informace** – `Zone` = `Grey`.



Obrázek 6.10: Klasifikace 100 phishingových, malwarových a benigních IPv4 adres službou opentip.kaspersky.com

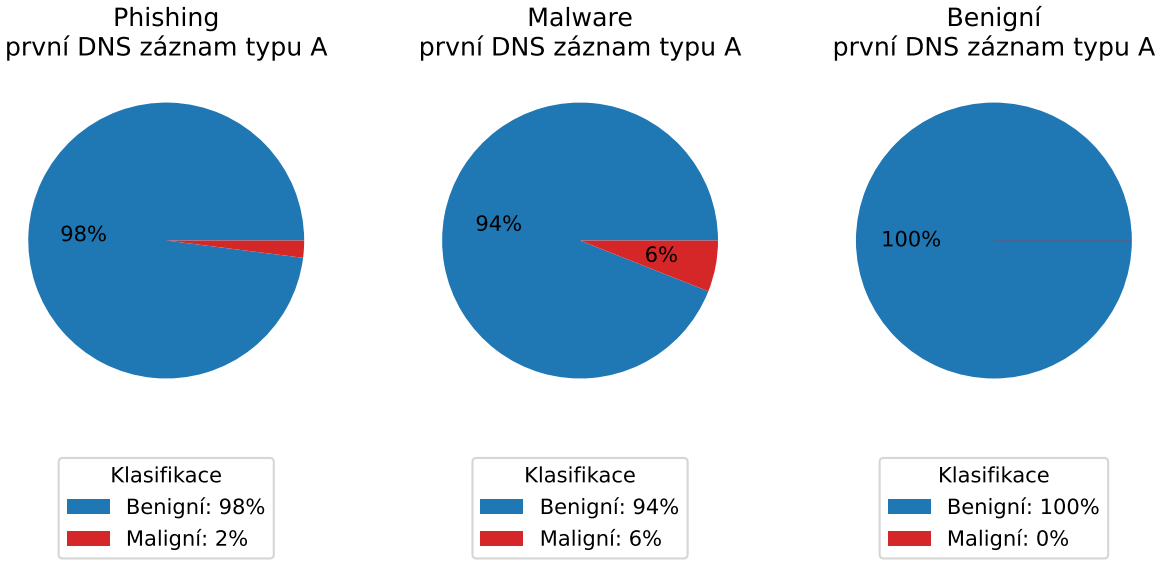
6.10 projecthoneypot.org

Klasifikace IPv4 adres službou Project Honeypot je dle jednotlivých kategorií zobrazena na obrázku 6.11. Jelikož služba zařazuje IP adresu do několika kategorií viz sekce 5.10, tak je IP adresa na obrázku klasifikována dle následujících podmínek:

- **Benigní** – IP adresa není službou zařazena do žádné kategorie,
- **Maligní** – IP adresa je službou zařazena alespoň do jedné kategorie.

6.11 pulsedive.com

Na obrázku 6.12 je vyobrazena klasifikace doménových jmen službou Pulsedive z jednotlivých kategorií. Klasifikace byla rozhodnuta pomocí položky `risk` získané pomocí služby API. Hladiny klasifikace byly zvoleny pomocí popisu jednotlivých hodnot položky `risk` a některé byly sloučeny, jelikož jejich popis je velmi podobný. Tyto hladiny jsou následující:



Obrázek 6.11: Klasifikace 100 phishingových, malwarových a benigních IPv4 adres službou projecthoneypot.org

- **Benigní** – `risk` = `none`,
- **Pravděpodobně benigní** – `risk` = `low`,
- **Pravděpodobně malignní** – `risk` = `medium` nebo `high`,
- **Malignní** – `risk` = `critical`,
- **Neznámá klasifikace** – `risk` = `unknown` nebo `retired`,
- **Žádné informace** – získaná odpověď neobsahuje položku `risk`.

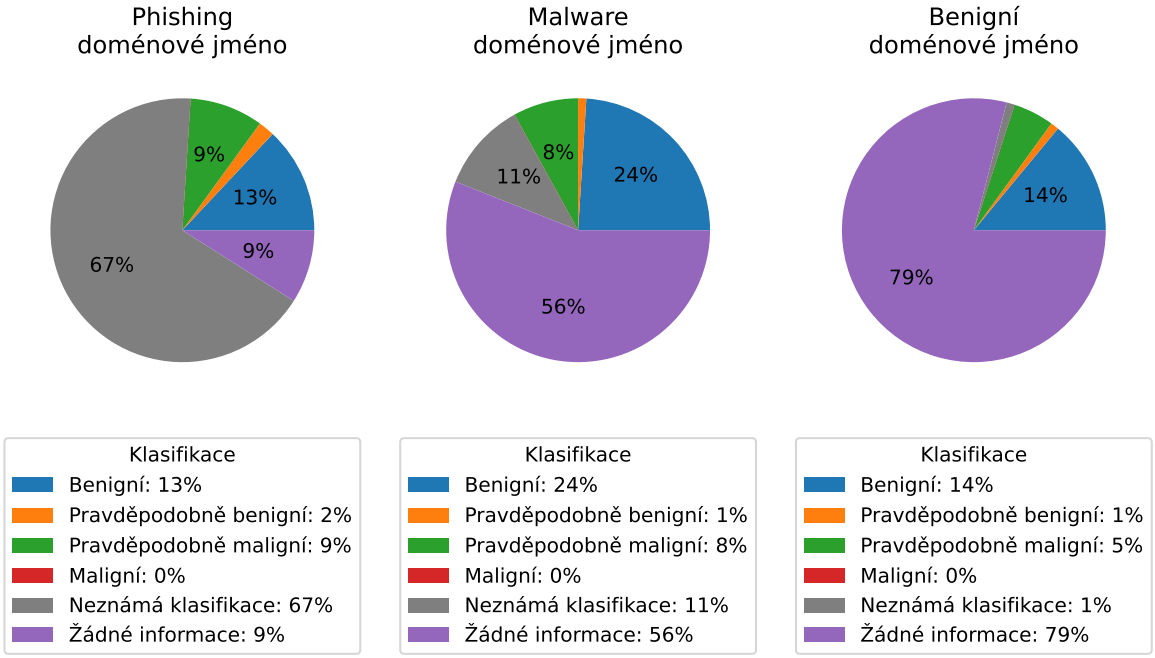
6.12 threatfox.abuse.ch

Klasifikace doménových jmen z jednotlivých datových sad službou threatfox.abuse.ch je zobrazena na obrázku 6.13. O doménových jménech nejsou žádné informace nejspíše z toho důvodu, protože služba po určité době maže záznamy o těchto doménách. Ke klasifikaci byla využita položka `malware` v získané odpovědi ze systému a to následovně:

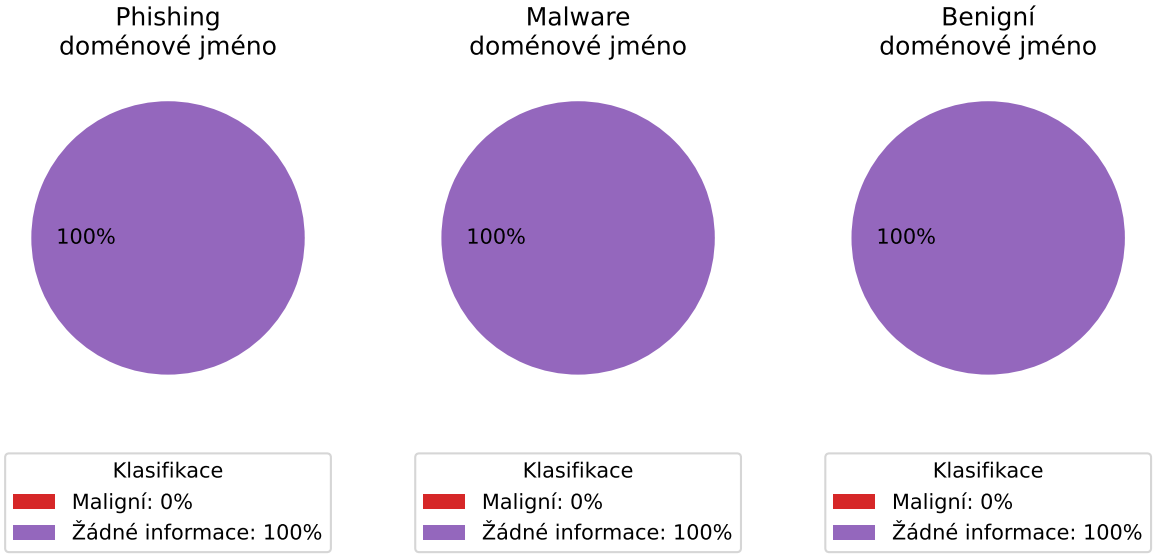
- **Benigní** – položka `malware` není v odpovědi uvedena,
- **Malignní** – položka `malware` je v odpovědi uvedena.

6.13 urlvoid.com

Na obrázku 6.14 je zobrazeno rozklasifikování doménových jmen z jednotlivých datových sad pomocí služby urlvoid.com. Ke klasifikaci byla využita hodnota `Detection Counts` uvedená na webové stránce reportu dané IP adresy. Zvolené hladiny pro klasifikace jsou následující:

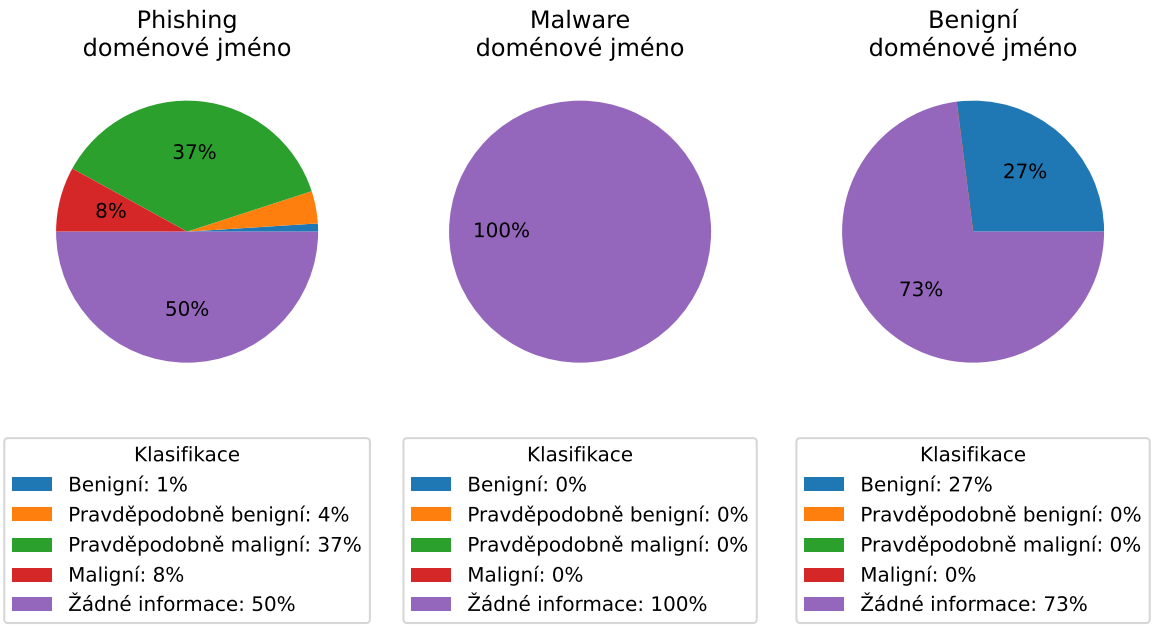


Obrázek 6.12: Klasifikace 100 phishingových, malwarových a benigních doménových jmen službou pulsedive.com



Obrázek 6.13: Klasifikace 100 phishingových, malwarových a benigních doménových jmen službou threatfox.abuse.ch

- **Benigní** – `Detection Counts` $= 0$,
- **Pravděpodobně benigní** – $1 \leq$ `Detection Counts` ≤ 2,
- **Pravděpodobně maligní** – $3 \leq$ `Detection Counts` ≤ 6,
- **Maligní** – $7 \leq$ `Detection Counts` ≤ 39.



Obrázek 6.14: Klasifikace 100 phishingových, malwarových a benigních IPv4 adres službou urlvoid.com

6.14 virustotal.com

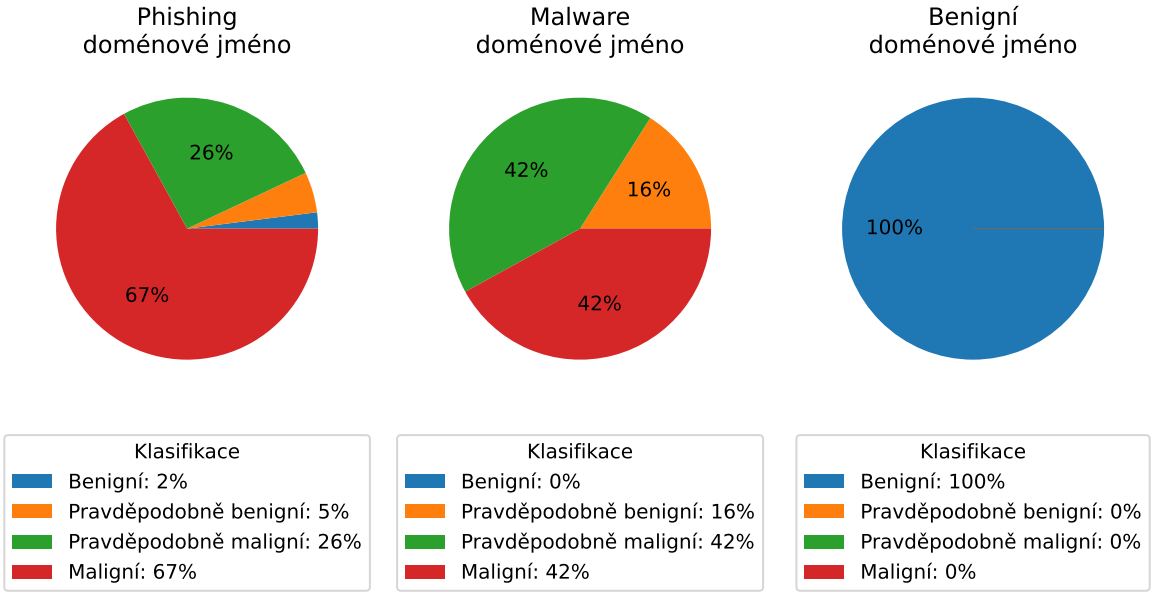
Obrázek 6.15 zobrazuje klasifikaci doménových jmen ze všech tří předpřipravených datových sad. Obrázek 6.16 naopak vyobrazuje klasifikaci prvních IPv4 adres těchto doménových jmen. Již na první pohled je zřejmé, že tato služba poskytuje u doménových jmen přesné a poměrně vhodné výsledky. U IPv4 adres jsou výsledky méně přesné. Zvolené hladiny pro klasifikace jsou následující:

- **Benigní** – `last_analysis_stats.malicious` $= 0$,
- **Pravděpodobně benigní** – $1 \leq$ `last_analysis_stats.malicious` ≤ 4,
- **Pravděpodobně maligní** – $5 \leq$ `last_analysis_stats.malicious` ≤ 9,
- **Maligní** – $10 \leq$ `last_analysis_stats.malicious` ≤ 100.

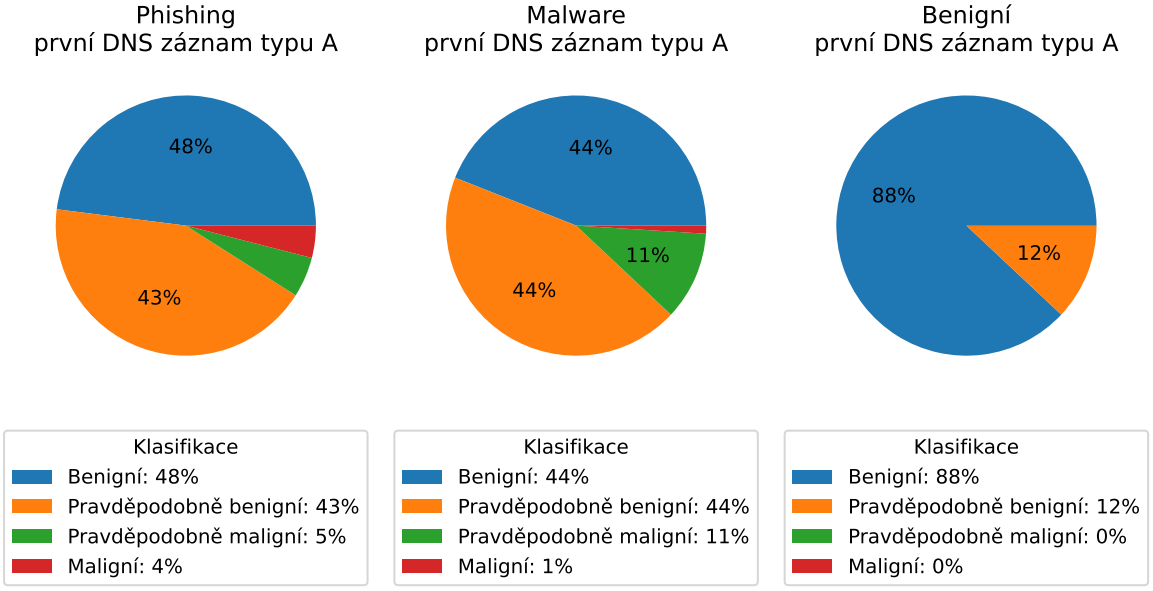
6.15 Zhodnocení

Z grafů klasifikace doménových jmen či jejich IP adres reputačními systémy, je zřejmé, že některé reputační systémy si v rozpoznání škodlivých domén vedly zásadně lépe než jiné. Tento výsledek naznačuje, že reputační systémy s vyšší mírou přesnosti a spolehlivosti by měly mít při klasifikaci větší váhu.

Výhodou některých reputačních systémů je, že dokáží oddělit entity, o kterých nemají žádné informace od těch, které jsou benigní. Díky tomuto přístupu nebude výsledek při klasifikaci zkreslený a umožní přesnější hodnocení entity. V případě klasifikace dat z reputačních systémů, které tyto dva typy nerozlišují, tedy pokud o dané entitě nemají žádný záznam, je automaticky interpretována jako benigní, bude nutné, abych já tyto typy rozlišil,



Obrázek 6.15: Klasifikace 100 phishingových, malwarových a benigních doménových jmen službou virustotal.com



Obrázek 6.16: Klasifikace 100 phishingových, malwarových a benigních IPv4 adres službou virustotal.com

aby nevzniklo nežádoucí zkreslení. Pokud systém takovou entitu nesprávně klasifikuje jako benigní, mohlo by to vést k falešně pozitivním výsledkům, což by mohlo ovlivnit celkovou přesnost klasifikace.

Kapitola 7

Návrh rošíření nástroje DomainRadar

Hlavní část bude spočívat ve vytvoření jednotlivých kolektorů pro daný reputační systém. Každý kolektor bude v rámci systému Apache Kafka vystupovat jakožto *producer* a zároveň i jako *consumer*. *Producer* část kolektoru bude vytvářet zprávy obsahující získaná data z reputačního systému o jednotlivých doménových jménech nebo IP adresách. Na druhé straně část *consumer* bude přijímat zprávy z témat (*topic*), ve kterých bude uvedena IP adresa či doména, o které má daný reputační systém zjistit informace.

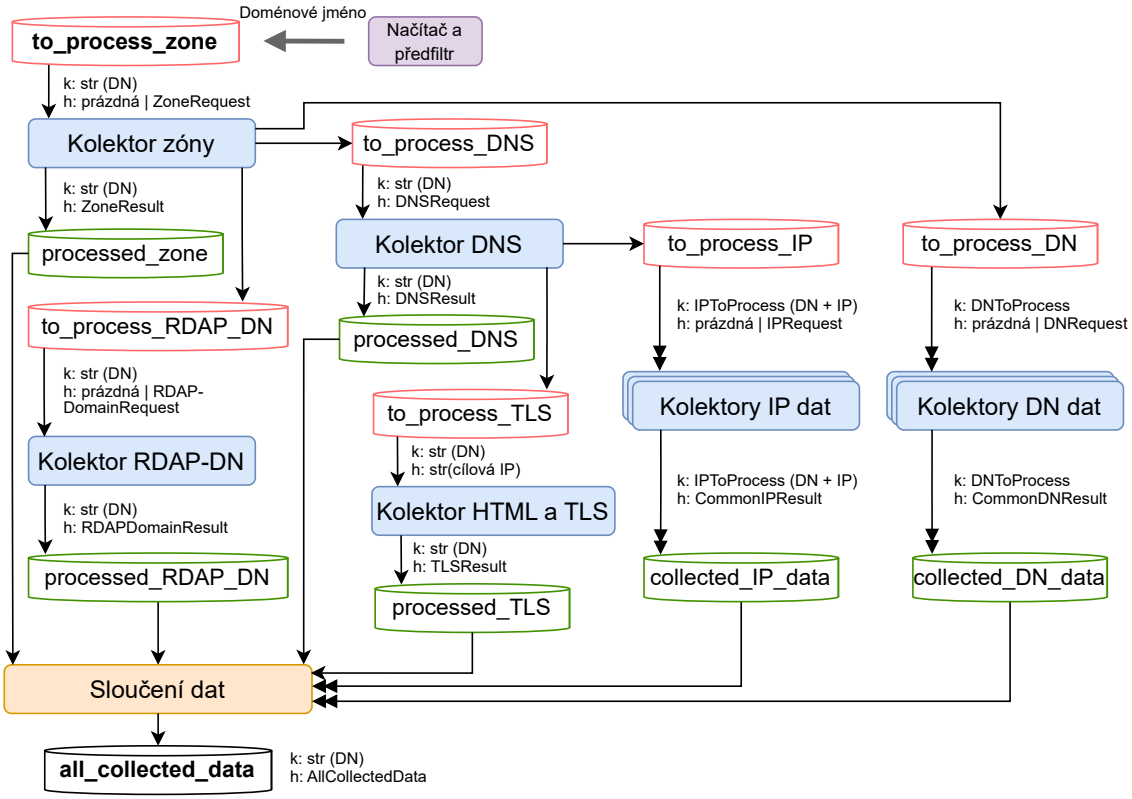
7.1 Rozšíření modelu zřetězené kolekce dat v nástroji DomainRadar

Jelikož je potřeba implementovat i kolektory reputačních systémů, které pracují s doménovými jmény, bude nutné v DomainRadaru rozšířit model zřetězené kolekce. Nabízí se dvě možná řešení, kterým tento problém může být vyřešen. Prvním je přidat zdrojová i cílová témata pro každý kolektor reputačních systémů, který pracuje s doménovými jmény. Samotný kolektor by poté odebíral zdrojové téma, se kterým by vystupoval jako *consumer* a data ukládal to cílového tématu jakožto *producer*. Jelikož není potřeba získávat data od některých kolektorů reputačních systémů dříve než od jiných, protože jednotlivé reputační systémy nezávisí na odpovědi jiného reputačního systému (tj. nezpracovávají odpověď od jiného systému), bylo by toto řešení zbytečně komplikované. Problém by mohl nastat, kdyby bylo žádoucí nástroj DomainRadar rozšířit o další reputační systém. V ten moment by bylo potřeba přidat další zdrojové a cílové téma a zajistit správné zřetězení dat. Z těchto důvodů je toto řešení velice nevhodné.

Druhým možným a dle mého úsudku lepším řešením je vytvořit jedno společné zdrojové téma, od kterého by kolektory reputačních systémů odebíraly data, tedy vystupovaly by jako *consumer*, a jedno společné cílové téma, do kterého by kolektory ukládaly data jakožto *producer*. Díky tomu nebude při vytváření nového kolektoru reputačních systémů (či jiného kolektoru) pracujícího s doménovými jmény třeba přidávat nezbytně dvě nová témata, ale zaregistroval by se pouze jako *consumer* a jako *producer* do příslušných společných témat. Tento způsob řešení by tedy umožnil jednodušší rozšířitelnost nástroje DomainRadar.

Nejen pro účely lepší budoucí integrace do nástroje je zvolena druhá možnost řešení. Návrh rozšířeného modelu je vyobrazen na obrázku 7.1. Nově vytvořeným společným zdrojovým tématem je téma `to_process_DN`. Klíčem je nově vytvořený model `DNToProcess`,

který je popsán v sekci 7.2.1, a hodnota je buď prázdná nebo model `DNRequest`, který je popsán v sekci 7.2.2. Společným zdrojovým tématem, které je nově vytvořené, je téma s názvem `collected_DN_data`. Do tohoto tématu budou kolektory reputačních systémů pracující s doménovými jmény ukládat data, kde klíč je model `DNToProcess`, jehož hodnota je stejná jako při přečtení z tématu `to_process_DN`. Stejná hodnota klíče je z toho důvodu, aby následně při slučování dat bylo možné rozlišit data od různých doménových jmen. Hodnotou, která je do tématu uložená, je model `CommonDNResult`. Tento model je blíže popsán v sekci 7.2.3. Z důvodu přehlednosti jsou z tohoto návrhu vypuštěny kolektory reputačních systémů a místo toho jsou nahrazeny zástupným blokem. Samotné zapojení kolektorů je blíže znázorněno v sekci 7.1.1.

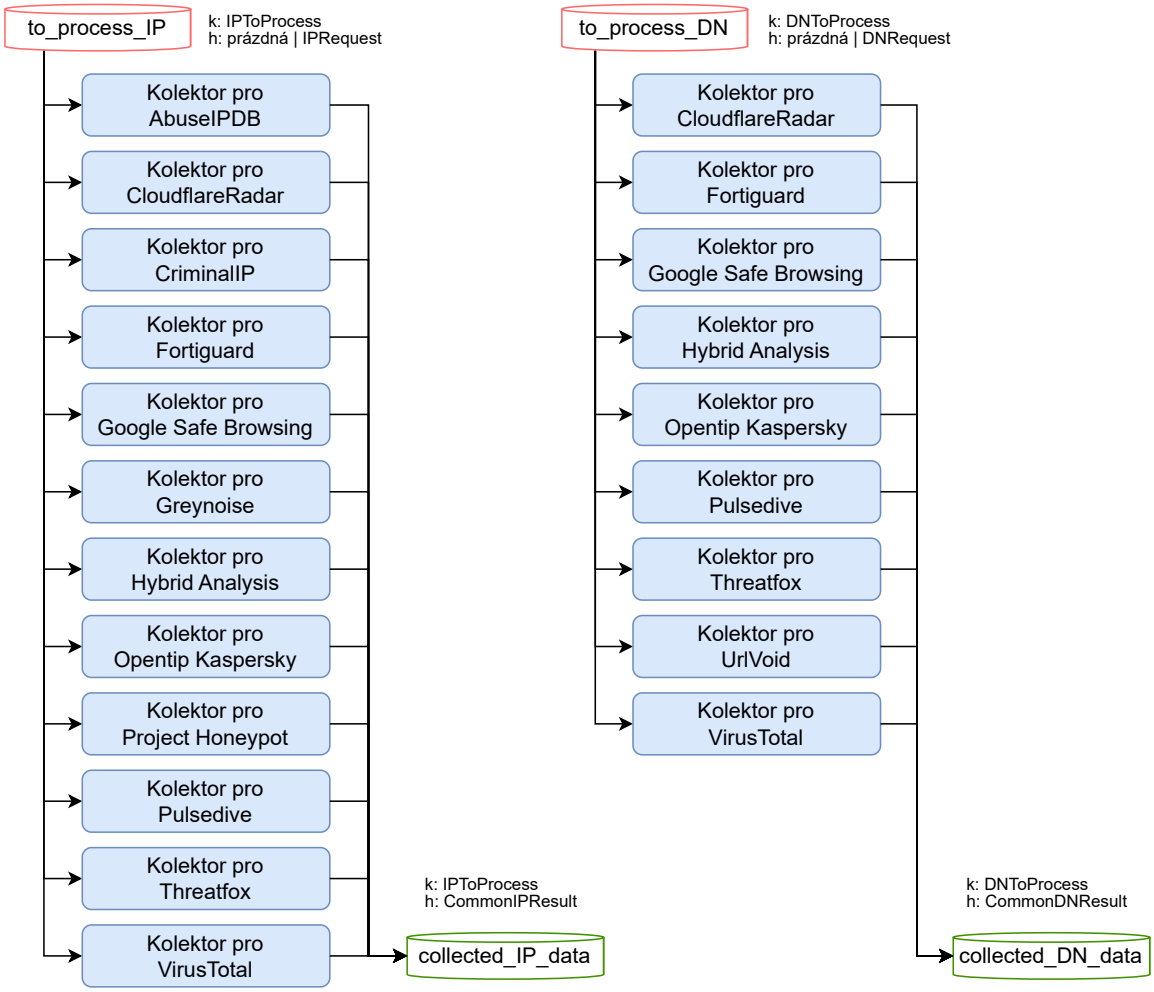


Obrázek 7.1: Návrh rozšířeného modelu zřetězené kolekce dat v nástroji DomainRadar. Válec s červeným obrysem značí zdrojové téma. Válec se zeleným obrysem představuje cílové téma.

7.1.1 Zapojení kolektorů reputačních systémů do zřetězené kolekce dat

Návrh zapojení kolektorů pro reputační systémy jak pro IP adresy, tak pro doménová jména je zobrazen na obrázku 7.2. Jedná se pouze o zobrazení úseku zřetězené kolekce dat, která má za úkol především zachytit, jaký kolektor bude odebírat jaké téma v systému Apache Kafka. Tudíž kolektory reputačních systémů, které umí zpracovat IP adresy, budou odebírat téma `to_process_IP`, k němuž budou přistupovat jako *consumer*. Získaná a zpracovaná data z reputačních systémů budou tyto kolektory ukládat do tématu `collected_IP_data` (tedy budou pracovat jako *producer*). V případě, že reputační systém, dokáže podávat informace ohledně doménových jmen, bude kolektor, který tento daný reputační systém zpracovává,

odebírat téma `to_process_DN`, odkud zjistí doménové jméno, ke kterému mají být získány data reputačním systémem. Data, která jsou kolektorem získaná z reputačního systému, jsou následně uložena do tématu `collected_DN_data`.



Obrázek 7.2: Návrh vytvoření a provázání jednotlivých kolektorů. Válec s červeným obrysem značí zdrojové téma, válec se zeleným obrysem představuje cílové téma a modrý obdélník značí kolektor.

7.2 Nově definované modely pro zřetězenou kolekci dat

Tato sekce se zabývá definicí nově vzniklých modelů, které slouží k jednotnému ukládání a čtení dat mezi tématy při zřetězené kolekci dat.

7.2.1 Definice modelu DNToProcess

Model `DNToProcess` slouží jako klíč a identifikuje doménové jméno. Při konzumaci dat z tématu `to_process_DN` je jeho účel takový, aby kolektor věděl, o kterém doménovém jméně má získat informace z reputačního systému. V případě ukládání dat do tématu `collected_DN_data` je účelem tohoto modelu, aby výsledná data byla snadněji rozlišitelná

mezi všemi daty a bylo tak možné získat data právě k jedné doméně, kterou model zastupuje. Vytvořený model je definován následovně:

```
class DNToProcess:
    dn: str
```

7.2.2 Definice modelu DNRequest

Model `DNRequest` obsahuje položku `collectors`, která udává jaké kolektory reputačních systémů mají danou zprávu, jejíž je instance modelu `DNRequest` hodnotou, zpracovat. V případě, kdy je namísto modelu `DNRequest` v hodnotě zprávy použita prázdná hodnota (`null`), je naznačeno, že zprávu mají zpracovat všechny kolektory. Tento vytvořený model je definován následovně:

```
class DNRequest:
    collectors: set[str] | None
```

7.2.3 Definice modelu CommonDNResult

Tato sekce se zaměřuje na definici nového datového typu `CommonDNResult`, který bude implementován do nástroje DomainRadar. Tento datový typ bude klíčovým prvkem pro některé kolektory využívající volně dostupné reputační systémy. Přesněji se jedná o kolektory dat pracující s reputačnímy systémy klasifikující na základě doménových jmen.

Model bude sloužit k uložení výsledku operace kolektorů pracujících s doménovými jmény. Tudíž bude možné odlišit, zda výsledek pochází z kolektoru, který měl na vstupu IP adresu (`CommonIPResult`) či doménové jméno (nově definovaný model `CommonDNResult`). Tento model rozšiřuje bázový model `Result`, který byl popsán v sekci 4.2.5. Tudíž model `CommonDNResult` bude obsahovat stejné položky jako jeho bázový model `Result`.

Položka `collector` značí textový identifikační název daného kolektoru, který výsledná data vyprodukoval. Objekt `data` je zde jako generický objekt, který si každý kolektor pracující s doménovým jménem na vstupu předefinuje specifickým modelem. Tento specifický model bude přímo určený právě danému kolektoru a bude obsahovat položky, které daný kolektor poskytuje.

Nově definovaný model `CommonDNResult` odpovídá následující definici:

```
class CommonDNResult(Result):
    collector: str
    data: object | None
```

7.3 Návrh kolektorů pro reputační systémy

Tato sekce se zaměřuje na návrh jednotlivých kolektorů reputačních systémů, které budou implementovány v rámci projektu DomainRadar. Kolektory budou klíčovými komponentami, které budou zajišťovat sběr informací z reputačních systémů (respektive jeden kolektor bude zajišťovat sběr z jednoho reputačního systému). Návrh kolektoru bude obsahovat několik kroků. Prvním krokem bude určení, jakým způsobem bude daný kolektor komunikovat s reputačním systémem. V závislosti na typu kolektoru a na typu reputačního systému mohou být získávána různá data, tudíž dalším krokem bude specifikování, která

data bude daný reputační systém shormažďovat. Následovat bude specifikace objektu, který bude obsahovat tato získaná data.

Obecně budou jednotlivé kolektory získávat data pomocí rozhraní API, pokud jej daná služba nabízí k dispozici. API rozhraní umožňuje strukturovanou výměnu dat, což zajišťuje vyšší efektivitu a rychlost. V případě, kdy daná služba rozhraní API nenabízí, bude nutné přistoupit k alternativnímu řešení získávání dat. Bude nutné příslušnou stránku, kde je vypsáno reputační hodnocení (případně i jiná vhodná informace), stáhnout a následně tyto potřebné informace o daném subjektu z dokumentu vyextrahovat. Tento alternativní přístup je časově zdlouhavější, ale i přesto je tento způsob získávání dat vhodný, jelikož umožňuje získat další informace pro klasifikaci dané entity.

Pro omezení času, který by kolektor strávil nad extrakcí dat při využití sběru dat z dokumentu stránky, bude vhodné nastavit přiměřenou dobu, kterou bude mít daný kolektor na získání potřebných informací. Pokud se kolektoru nepodaří do stanovené doby získat požadovaná data, bude daná extrakce automaticky ukončena. V takovém případě bude součástí výstupu chybová hláška *Timeout*, která značí, že čas vypršel. Tak se zajistí, že kolektor nestráví extrakcí spoustu času, který by mohl silně ovlivnit celou klasifikaci dané entity.

7.3.1 Kolektor dat ze systému abuseipdb.com

Kolektor, který bude získávat data ze systému abuseipdb.com, bude využívat volně dostupné rozhraní API ke komunikaci s tímto reputačním systémem. Tento kolektor bude odebírat téma `to_process_IP` ze systému Apache Kafka, tudíž výsledná hodnota uložená do tématu `collected_IP_data` bude typu `CommonIPResult<AbuseipdbData>`. Třída `AbuseipdbData` definuje data, která bude tento kolektor shromažďovat a ukládat. Tato třída je popsána následovně:

```
class AbuseipdbData:
    abuseConfidenceScore: int
    isWhitelisted: bool
    isTor: bool
    totalReports: int
```

Všechny položky uvedené v modelu `AbuseipdbData` jsou stejnojmenné informace získané pomocí API systému AbuseIPDB.

7.3.2 Kolektor dat ze systému radar.cloudflare.com

Reputační systém Cloudflare Radar umožňuje klasifikovat IPv4 adresy a doménová jména. Tudíž kolektor pro tento reputační systém bude odebírat témata dvě a to `to_process_IP` a `to_process_DN`. Výsledky budou typu `CommonIPResult<CloudflareRadarData>` respektive `CommonDNResult<CloudflareRadarData>`. V případě, že z tématu `to_process_IP` bude odebrána IPv6 adresa, budou data `CommonIPResult` nastavená na `None`. Nově definovaný model `CloudflareRadarData` je definován následovně:

```
class CloudflareRadarData:
    malicious: bool
```

Položka `malicious` v modelu `CloudflareRadarData` odpovídá hodnotě získané pomocí API reputačního systému.

7.3.3 Kolektor dat ze systému criminalip.io

Reputační systém criminalip.io dovoluje využít rozhraní API pouze pro získávání informací ohledně IPv4 adres. Z tohoto důvodu bude kolektor pro tento systém odebírat téma `to_process_IP` a produkovat bude typ `CommonIPResult<CriminalIPData>`. Pokud by IP adresou získanou z `to_process_IP` byla IPv6 adresa, budou data `CommonIPResult` nastavená na `None`. Model `CriminalIPData` popisující ukládaná data ze systému criminalip.io je definován takto:

```
class CriminalIPData:
    scoreInbound: str
    scoreOutbound: str
```

7.3.4 Kolektor dat ze systému fortiguard.com

Kolektor dat pro systém Fortiguard bude odebírat dvě témata. Prvním odebíraným tématem je `to_process_IP`, při kterém budou mít uložená data do výstupního tématu hodnotu typu `CommonIPResult<FortiguardData>`. Druhým odebíraným tématem bude `to_process_DN`. Zde bude uložená hodnota zprávy do výstupního tématu se zpracovanými daty mít typ `CommonDNResult<FortiguardData>`. Nový model `FortiguardData` pro data z tohoto systému je definován následovně:

```
class FortiguardData:
    spam: bool
```

Hodnota položky `spam` v modelu `FortiguardData` odpovídá stejnojmenné položce uvedené v odpovědi systému poskytnuté prostřednictvím rozhraní API.

7.3.5 Kolektor dat ze systému safebrowsing.google.com

Systém Google Safe Browsing umožňuje používat jejich služby skrze API, pomocí kterého mohou být získány data k IPv4 adresám a k doménovým jménům. Proto bude kolektor sbírající data ze systému Google Safe Browsing odebírat témata `to_process_IP` a `to_process_DN`. Data uložená do příslušného výstupního tématu tak budou mít hodnotu typu `GoogleSafeBrowsingData`. Tento model je popsán ve výpisu takto:

```
class GoogleSafeBrowsingData:
    threatTypeUnspecifiedCnt: int
    malwareCnt: int
    socialEngineeringCnt: int
    unwantedSoftwareCnt: int
    potentiallyHarmfulApplicationCnt: int
```

7.3.6 Kolektor dat ze systému greynoise.io

Reputační systém greynoise.io umožňuje získávat informace ohledně IPv4 adres, proto bude kolektor implementující sběr z tohoto systému odebírat téma `to_process_IP`. Zpráva uložená do tématu `collected_IP_data` bude mít hodnotu `CommonIPResult<GreynoiseData>`. Jelikož systém nepodoruje IPv6 adresy, tak v případě, že odebraná IP adresa ze zdrojového tématu bude verze 6, data `CommonIPResult` budou nastavená na `None`. Nově vytvořený model `GreynoiseData` s daty, které budou od reputačního systému získány, je definován tímto výpisem:

```
class GreynoiseData:
    noise: bool
    riot: bool
    classification: str
```

7.3.7 Kolektor dat ze systému hybrid-analysis.com

Reputační systém hybrid-analysis.com nabízí rozhraní API, pomocí kterého mohou být vyhledávány výsledky k IP adrese i k doménovému jménu. Kolektor dat z tohoto systému bude využívat toto rozhraní, a tudíž bude přihlášen k odběru dvou témat: `to_process_IP` a `to_process_DN`. Výstupem tohoto kolektoru bude v závislosti na typu vstupu (tedy zda se jedná o IP adresu či doménu) typ `CommonIPResult<HybridAnalysisData>` nebo `CommonDNResult<HybridAnalysisData>`. Model `HybridAnalysisData` je definován v následujícím výpisu takto:

```
class HybridAnalysisData:
    maliciousCnt: int
    suspiciousCnt: int
    noThreatCnt: int
    whitelistedCnt: int
    worstScore: int
    bestScore: int
    avgScore: int
    nullScoreCnt: int
```

V modelu `HybridAnalysisData` položky `maliciousCnt`, `suspiciousCnt`, `noThreatCnt` a `whitelistedCnt` jsou počítadla, kolikrát se daná klasifikace (tedy název položky bez `Cnt`) objevila v odpovědi reputačního systému v položce `verdict`. Údaj `worstScore` a údaj `bestScore` znamená nejhorší skóre respektive nejlepší skóre zaznamenané mezi seznamem odpovědí a to pod položkou `threat_score` z API. Hodnota `avgScore` následně představuje průměrnou hodnotu všech uvedených ohodnocení. `nullScoreCnt` představuje počet, kolikrát byla daná entita nahlášena, avšak není přidělené skóre, tedy `threat_score` je rovno `None`.

7.3.8 Kolektor dat ze systému opentip.kaspersky.com

Jelikož systém Opentip Kaspersky dokáže ohodnotit IPv4 adresy a doménová jména, tak kolektor získavající data z tohoto systému bude odebírat jak téma `to_process_IP` tak i téma `to_process_DN`. Model s daty získanými z tohoto systému o dané entitě je definován tímto výpisem:

```
class OpentipKasperskyData:
    zone: str
    categories: list[str]
    categoryZones: list[str]
```

Položka `zone` v modelu `OpentipKasperskyData` je stejnojmenná položka, která se nachází v odpovědi API ze systému Opentip Kaspersky. Co se týče položky `categories`, tak ta obsahuje všechny kategorie, do kterých byla daná entita systémem zařazena. `categoryZone` pak následně obsahuje zónu pro kategorie a to v takovém pořadí, jaké kategorii odpovídají ze seznamu `categories`.

7.3.9 Kolektor dat ze systému projecthoneypot.org

Jelikož systém projecthoneypot.org nenabízí API, bude muset kolektor získávající informace z tohoto systému stránku s reportem o dané IPv4 adrese stáhnout a následně bude zapotřebí z ní vyextrahovat, zda systém zařadil IP adresu do některé skupiny maligních IP adres. Jelikož systém dokáže ohodnotit pouze IP adresy, tak bude kolektor pracující se systémem projecthoneypot.org odebírat pouze téma `to_process_IP`. S ohledem na to, že systém dokáže zpracovat pouze IPv4 adresy, tak při odběru IPv6 adresy ze vstupního tématu, bude do výstupního tématu uložen `CommonIPResult`, který bude mít data nastavená na `None`. Model, který nese informaci, zda daná IPv4 adresa byla označena systémem za maligní, je popsán v tomto výpise:

```
class ProjectHoneypotData:
    malicious: bool
```

7.3.10 Kolektor dat ze systému pulsedive.com

Data z reputačního systému pulsedive.com budou kolektorem sbírána pomocí rozhraní API. Tento systém poskytuje reputační ohodnocení jak pro doménová jména, tak i pro IP adresy, tudíž kolektor bude odebírat téma `to_process_IP` i `to_process_DN`. Výstupem tohoto kolektoru bude tedy buď typ `CommonIPResult<PulsediveData>` nebo typ `CommonDNResult<PulsediveData>` v závislosti na vstupu. Z API budou vybrány pouze ty položky, které jsou užitečné pro klasifikaci, tedy především reputační skóre. Jednotlivé položky v modelu `PulsediveData` odpovídají položkám získatelným pomocí rozhraní API. Samotný model `PulsediveData` je definován následovně:

```
class PulsediveData:
    risk: str
    riskRecommended: str
    manualRisk: int
```

7.3.11 Kolektor dat ze systému threatfox.abuse.ch

Systém Threatfox nabízí službu API, proto jej bude kolektor pracující s tímto reputačním systémem využívat. Jelikož systém dokáže ohodnotit IPv4 adresu a doménová jména, bude kolektor odebírat téma `to_process_IP` a `to_process_DN`. V případě, že odpověď nebude obsahovat požadované informace, bude v modelu u dané položky uloženo `None`. V případě odběru IPv6 adresy ze vstupního tématu `to_process_IP` bude do výstupního tématu uložen `CommonIPResult` s položkou `data` nastavenou na `None`. Data, která bude kolektor ze systému sbírat, jsou popsána modelem `ThreatfoxData`, který je definován následovně:

```
class ThreatfoxData:
    threatType: str | None
    malware: str | None
    confidenceLevel: int | None
    tags: list[str]
```

Data v modelu `ThreatfoxData` jsou stejnojmenná data, která jsou získatelná z rozhraní API systému threatfox.abuse.ch.

7.3.12 Kolektor dat ze systému urlvoid.com

Kolektor pracující s reputačním systémem urlvoid.com nebude moci využívat rozhraní API, jelikož jej systém neposkytuje. Tento kolektor bude tudíž muset využívat alternativní možnosti, kdy bude muset stránku stáhnout a následně vyextrahovat počet udávající, v kolika případech bylo dané doménové jméno detekováno systémem. Jelikož kolektor bude odebírat téma `to_process_DN`, tak výsledný typ tohoto kolektoru bude uložen jako `CommonDNResult<UrlvoidData>`. Získaná data tímto kolektorem budou následně uložena dle třídy, která je uvedena ve výpisu následovně:

```
class UrlvoidData:
    detectionCounts: int | None
```

7.3.13 Kolektor dat ze systému virustotal.com

Kolektor získávající data z reputačního systému virustotal.com bude využívat poskytované rozhraní API k získávání informací o doménách a IP adresách. Vzhledem k tomu, že tento kolektor bude odebírat dvě témata, konkrétně `to_process_IP` pro IP adresy a `to_process_DN` pro doménová jména, tak výsledný typ bude záviset na typu entity, kterou kolektor zpracovává. Pokud se jedná o analýzu IP adresy, výsledná data budou nabývat typu `CommonIPResult<VirustotalData>`. V případě, že se jedná o analýzu doménového jména, tak výsledný typ bude `CommonDNResult<VirustotalData>`. Model `VirusTotal` je definován ve výpisu následovně:

```
class VirustotalData:
    reputation: int
    malicious: int
    suspicious: int
    undetected: int
    harmless: int
```

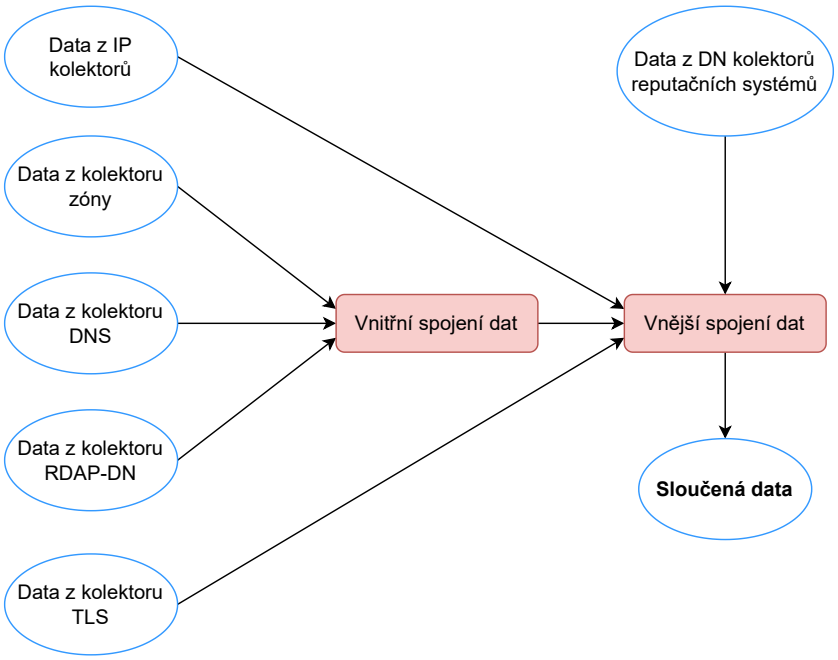
Položka `reputation` odpovídá položce se stejným jménem, která je poskytnuta v odpovědi API reputačním systémem. Všechny ostatní položky odpovídají stejnojmenným položkám získaných pomocí API v objektu `last_analysis_stats`

7.4 Úprava procesu sloučení dat

Z důvodu přidání tématu `collected_DN_data` bude nutné upravit proces sloučení dat. Bez upravení tohoto procesu by totiž výsledná sloučená data neobsahovala informace o doménovém jméně získaných z reputačních systémů. Navrhnutá změna je zachycena na obrázku 7.3.

7.5 Klasifikace na základě získaných dat

Pro to, aby získaná data byla k něčemu užitečná, je nutné z těchto dat vyvodit nějaký závěr. Nabízí se dvě možná řešení, jak výsledného ohodnocení dosáhnout. Prvním možným řešením je stanovit heuristiku, která by umožňovala vypočítat a stanovit výslednou klasifikaci daného doménového jména. Bylo by však nutné jednotlivým reputačním systémům přiřadit váhy, které by odrážely vlastnosti daného reputačního systému. Zejména by se tedy váha musela stanovit na základě přesnosti, spolehlivosti a konzistenci v podávaném ohodnocení.



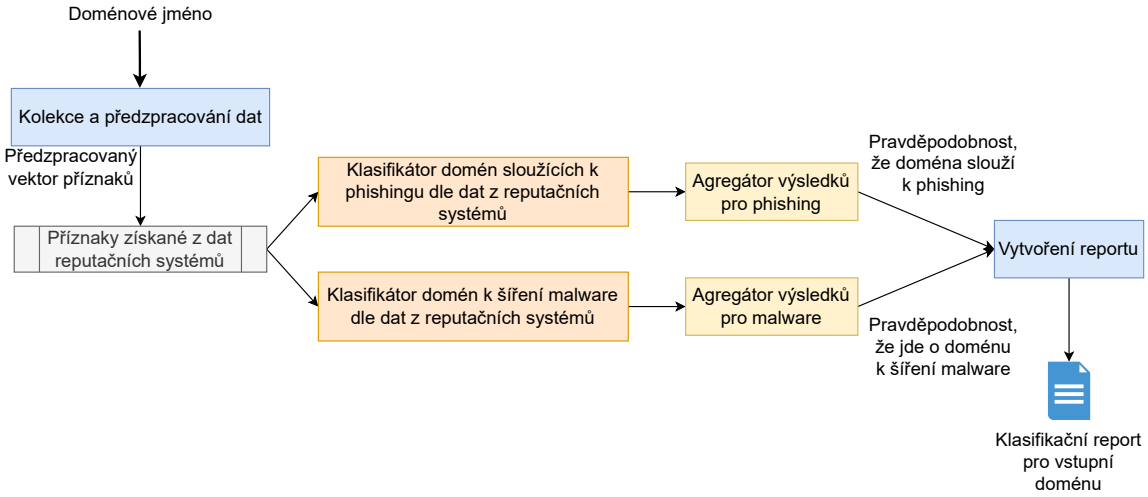
Obrázek 7.3: Schéma zachycující rozšířený proces sloučení dat

To by v praxi znamenalo, že reputační systémy, které dokáží přesněji a spolehlivěji rozlišit, zda dané doménové jméno či IP adresa je maligní či benigní, budou mít při klasifikaci větší váhu. Tímto krokem by se zajistilo, že výsledná klasifikace bude lépe odrážet skutečnost a nebude negativně ovlivněna systémy s nižší přesností. Další problém nastupuje při transformaci získaných dat. Jelikož ne každý reputační systém poskytuje přímo reputační skóre, ale namísto toho uvádí například zařazení do nějaké definované skupiny. Heuristika by tak nemohla pouze provést například vážený průměr reputačního skóre, ale musela by vhodným způsobem interpretovat možné zařazení do jednotlivých tříd. Tyto problémy by se mohly negativně projevit právě při použití většího množství dat z reputačních systémů. Najít vhodnou heuristiku, které by podávala požadovaný přesný výsledek, by bylo komplikované.

Druhou možností je využít řešení, které by pro klasifikaci využívalo metod strojového učení. V takovém případě by bylo potřeba najít vhodný model, který by podával konzistentní výsledky na základě dat z reputačních systémů. Tato metoda by v případě rozšíření nástroje DomainRadar o případný další reputační systém umožnila snažší rozšíření klasifikace v nástroji.

Z důvodu většího množství různorodých dat mi proto dává smysl využít právě druhé metody a vytvořit model pomocí metody strojového učení, který by zajišťoval klasifikaci na základě dat z reputačních systémů. Respketive bude potřeba vytvořit modely dva, jelikož nástroj DomainRadar rozlišuje klasifikaci pro doménová jména sloužící k phishingu od klasifikace domén, které mají za účel šířit malware. Jak bylo zmíněno dříve, nástroj DomainRadar taktéž klasifikuje rodiny algoritmů DGA, ty však reputační systémy nedokáží odlišit, proto je nebudou klasifikovat dalším samostatným klasifikátorem. Jeden nově vytvořený model by tak klasifikoval phishingová doménová jména na základě dat z reputačních systémů a druhý by obstarával klasifikaci domén určených pro šíření malware pomocí dat z reputačních systémů. Návrh tohoto řešení je vyobrazen na obrázku 7.4. Modely se budou podílet na celkové klasifikaci doménového jména tak, že jejich výsledná hodnota

(kterou představuje pravděpodobnost, že se jedná o pozitivní třídu, tedy maligní doménu) bude přidána již do existujícího agregátoru výsledků pro phishing respektive pro malware. Výslednou klasifikaci těchto dvou nově vytvořených modelů bude možné si zobrazit ve výsledném klasifikačním reportu.



Obrázek 7.4: Návrh klasifikace doménového jména z dat získaných z reputačních systémů

Data pro klasifikaci reputačního výsledku doménového jména

V nástroji DomainRadar je doménové jméno vstupem k celkové klasifikaci. U reputačních systémů, které pracují přímo s doménovým jménem, bude výběr dat pro klasifikaci relativně přímočarý, jelikož každý reputační systém poskytuje maximálně jednu sadu informací o daném doménovém jméně. Vybraná data jsou vypsána v tabulce B.1.

Data pro klasifikaci reputačního výsledku IP adres

Jelikož každé doménové jméno může mít různý počet DNS záznamů typu A a AAAA, bude nutné tento fakt zohlednit při klasifikaci a získaná data nejdříve zagregovat tak, aby výsledkem byla pouze jedna sada souhrnných hodnot a ne tolik sad, kolik má dané doménové jméno asociovaných IP adres. U dat pocházejících od IP adres se nabízejí dvě možnosti, jak provést tuto agregaci reputačního skóre:

- Všechny získané reputační skóre zprůměrovat do jedné hodnoty.
- Vybrat nejhorší reputační skóre, které daný kolektor k dané IP adrese získal.

Průměrná hodnota dává smysl však pouze u reputačních systémů, které dokáží odlišit IP adresy, o kterých nemají žádný záznam, od těch, které jsou bezproblémové. Pokud by se počítal průměr včetně hodnot, které neodpovídají skutečnosti, mohlo by dojít ke značnému zkreslení výsledků, což je nežádoucí. Proto u reputačních systémů, které tuto skutečnost nedokáží odlišit, bude nutné zvolit druhou možnost a vybrat tak to nejhorší přidělené reputační skóre. Tímto způsobem se efektivně zajistí, aby výsledná hodnota lépe odpovídala skutečnosti a produkovala spolehlivější výsledky. Všechny vybrané příznaky jsou uvedeny v tabulce B.2.

Kapitola 8

Implementace sběrové části

Pro uložení zprávy do tématu `to_process_DN` je upraven kolektor zóny, který je napsán v jazyce Python, a který je obsažen ve složce `domainradar/colext/python` mezi ostatními kolektory. Uložení zprávy do tohoto tématu proběhne pouze v tom případě, když je zóna úspěšně nalezena. Veškerá ostatní implementace sběrové části dat z reputačních systémů je napsaná v programovacím jazyce Java a je obsažena ve složce `domainradar/colext/java`.

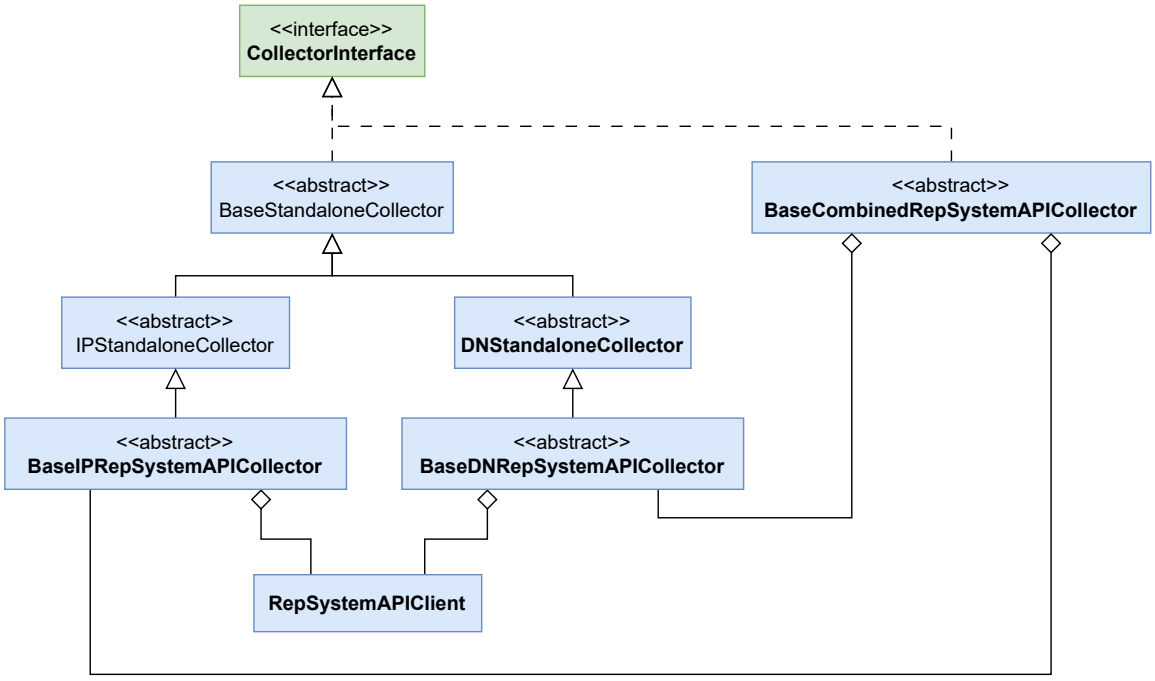
8.1 Kolektory dat z reputačních systémů

Kolektory, které sbírají data z reputačních systémů, jsou implementovány jako třída v již existujícím balíku `cz.vut.fit.domainradar.standalone.collectors`. Pro abstrahování společné logiky je vytvořeno několik abstraktních tříd, které lze vidět na obrázku 8.1. Tučně zvýrazněné třídy jsou nově vytvořené touto prací, zatímco nezvýrazněné třídy již existují. Funkce jednotlivých nově vzniklých tříd je následující:

- `DNStandaloneCollector` rozšiřuje základní třídu metodami specifickými pro kolektory doménových jmen. Tato třída rovněž inicializuje jeden Kafka *producer*.
- `RepSystemAPIClient` zajišťuje komunikaci s API reputačního systému.
- `BaseIPRepSystemAPICollector` se přihlašuje k odběru tématu `to_process_IP` a definuje metody, které musí kolektor dat o IP adrese z reputačních systému pracující s API implementovat, aby mohly být následně využity ke zpracování dotazu třídou `RepSystemAPIClient`.
- `BaseDNRepSystemAPICollector` se přihlašuje k odběru tématu `to_process_DN` a definuje metody, které musí kolektor dat o doménových jménech z reputačních systému pracující s API implementovat, aby mohly být využity ke zpracování dotazu třídou `RepSystemAPIClient`.
- `BaseCombinedRepSystemAPICollector` zajišťuje integraci kolektorů pro reputační systémy, které umí pracovat jak s IP adresy tak i s doménovými jmény.

Jelikož třída `BaseCombinedRepSystemAPICollector` nerozšiřuje pomocnou abstraktní třídu `BaseStandaloneCollector`, tak by kolektor pro reputační systém implementující třídu `BaseCombinedRepSystemAPICollector` nebyl považován za plnohodnotný kolektor.

Aby to bylo možné, je vytvořeno rozhraní `CollectorInterface`, které definuje nutné metody k implementaci. Díky tomu je tak možné provádět stejné operace nad kolektory, jejichž nepřímá nadtřída je `BaseStandaloneCollector` a nad kolektory, které dědí třídu `BaseCombinedRepSystemAPICollector`.

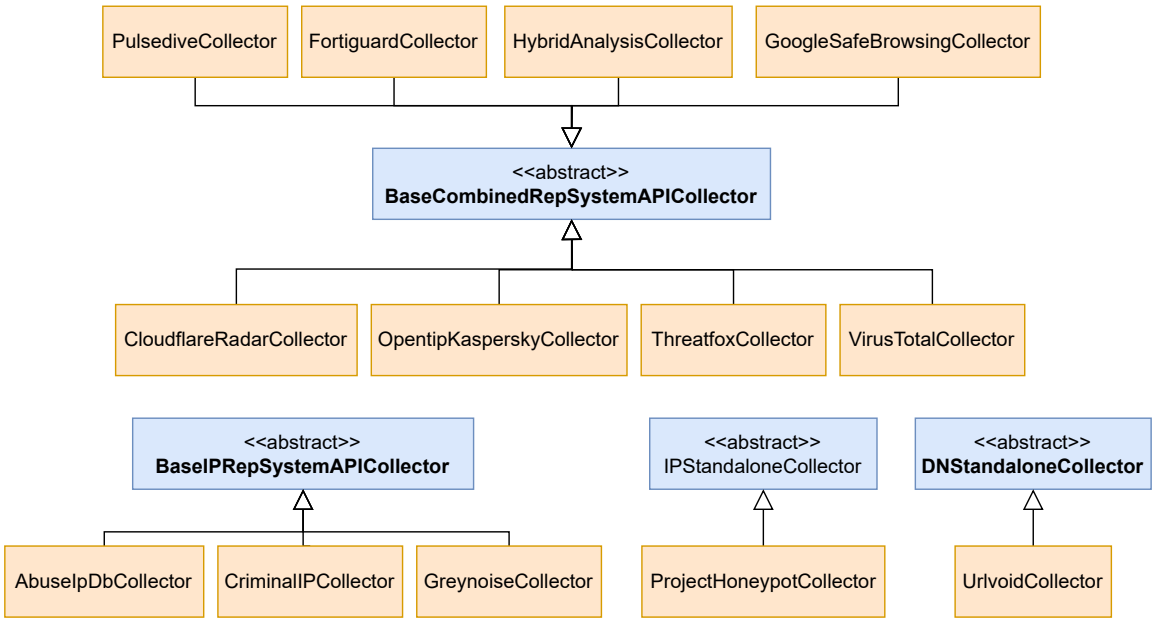


Obrázek 8.1: Diagram zachycující pomocné abstraktní třídy pro kolektory dat z reputačních systémů

Implementace samotných kolektorů

Samotné kolektory a jejich přímé abstraktní nadtřídy jsou zachyceny na diagramu 8.2. Modré obdélníky znázorňují abstraktní třídy abstrahující určitou část implementace kolekce dat (blíže popsáno v sekci 8.1). Oranžové obdélníky představují samotné kolektory dat z reputačních systémů, kde název kolektoru odpovídá názvu služby, ze které daný kolektor získává data, následovaný slovem `Collector`. Kolektory dat rozšiřující bázovou třídu `BaseCombinedRepSystemAPICollector` sbírají data o doménových jménech i IP adresách a ke sběru dat využívají rozhraní API poskytnuté danou službou. Třídy kolektorů, které rozšiřují bázovou třídu `BaseIPRepSystemAPICollector` sbírají data pouze o IP adresách a to pomocí API příslušné služby. Kolektory `ProjectHoneypotCollector` a `UrlvoidCollector` dědí bázovou abstraktní třídu `IPStandaloneCollector`, respektive bázovou abstraktní třídu `DNStandaloneCollector`, jelikož oba systémy, ze kterých tyto kolektory získávají data, neposkytují rozhraní API. Tyto kolektory pracují tak, že stáhnou webovou stránku, ze které následně vyextrahují potřebné informace. Toho je dosaženo s využítm knihovny Jsoup[1]. Na druhé straně kolektory pracující s API využívají pro zpracování JSON objektů balík org.json[2].

[1] Jsoup: https://jsoup.org/
[2] org.json: https://github.com/stleary/JSON-java



Obrázek 8.2: Diagram zachycující implementované kolektory s jejich přímou nadtřídou

Datové modely pro ukládaní dat

Implementace datových modelů pro data z reputačních systémů se nachází v nově vytvořeném balíčku `cz.vut.fit.domainradar.common.models.repsystems`. Modely jsou implementovány jako záznamy (klíčové slovo `record`) v jazyce Java.

Konfigurace kolektorů

Kolektory dat z reputačních systémů využívají konfigurační soubory, jejichž názvy jsou definovány v souboru `CollectorConfig.java`, kde jsou taktéž uvedeny výchozí hodnoty parametrů, v případě že konfigurační soubor daný parametr neobsahuje.

Uložení získaných dat kolektory do databáze

Získaná data jednotlivými kolektory dat z reputačních systémů jsou taktéž uložena do databáze. Název kolektoru v sloupci `collector` v tabulce `Collector` odpovídá názvu služby, ze které jsou data sbírána. V případě, že kolektor pracuje se službou, která dokáže získávat informace o doménových jménech i IP adresách, jsou vytvořeny v tabulce `Collector` záznamy dva. Pro doménová jména má záznam nastevený název kolektoru na název služby následovaný textem `-dn` a booleovská hodnota sloupce `is_ip_collector` je nastavena na `false`. Pro IP adresy má záznam nastavený hodnotu sloupce `collector` na název služby následovaný textem `-ip` a položka `is_ip_collector` má hodnotu `true`. Toto rozdělení je provedeno z toho důvodu, aby bylo možné rozlišit data v databázi právě u kolektorů dat z reputačních systémů, které umí ohodnotit jak doménová jména tak IP adresy. Vložení těchto záznamů kolektorů do databáze je přidáno v souboru `15_seed.sql`.

Aby se ukládala data o doménových jménech získaných pomocí kolektorů dat z reputačních systémů, bylo potřeba taktéž doplnit v konfiguračním souboru pro nástroj Kafka Connect `30_postgres-sink-dn-collectors.properties` nově vytvořené cílové téma `collected_DN_data`.

Spuštění kolektorů

Každý kolektor dat z reputačních systémů se v nástroji DomainRadar spouští jakožto samostatný Docker kontejner. Tedy při spuštění nástroje DomainRadar pomocí příkazu `docker compose up -d`, který je nutné provést ve složce `domainradar/infra`, se společně s ostatními součástmi nástroje spustí i kolektory dat z reputačních systémů.

8.2 Přidání dat z reputačních systémů do sloučených dat

Pro přidání IP dat získaných z reputačních systémů do celkových sloučených dat je upravena položka `TAGS` a `COLLECTOR_NAMES` v třídě `TagRegistry`, která se nachází v balíku `cz.vut.fit.domainradar.serialization`. Pro sloučení dat o doménových jménech získaných kolektory reputačních systémů k ostatním získaným datům o daném doménovém jméně je vytvořena třída `RepSystemDNEntriesProcessFunction` (ta je přidaná do existujícího balíku `cz.vut.fit.domainradar`), která se o toto sloučení stará. Tato třída postupně sbírá získaná data, dokud nezíská data od kolektorů všech reputačních systémů pracujících s doménovými jmény. V moment kdy přijme všechna data, tak je uloží společně s již dříve získanými daty do nově vzniklé agregační třídy `KafkaDomainWithRepSystemAggregate`. Pro určení kolektorů, které pracují s reputačními systémy zpracovávající doménová jména, jsou přidány do třídy `TagRegistry` dvě položky. První z nich je položka `REP_SYSTEM_DN_TAGS` a druhá položka `REP_SYSTEM_DN_COLLECTORS_NAMES`, v nich jsou obsaženy názvy právě těchto kolektorů mapované na číslo a zpětné mapování.

8.3 Extrakce příznaků

Transformační funkce, které se starají o extrakci příznaků z dat získaných z reputačních systémů, jsou implementovány v souborech `rep_dn.py` (extrakce příznaků o doménovém jméně) a `rep_ip.py` (extrakce příznaků o IP adresách). Tyto soubory jsou ve složce `domainradar/colext/python/extractor/extractor/transformations`. Všechny extrahované příznaky z reputačních systémů o doménovém jméně jsou vypsány v tabulce B.1, dále jsou v tabulce B.2 vypsány všechny příznaky z reputačních systémů o IP adresách.

Mapování zóny ze systému opentip.kaspersky.com

Jelikož zóna poskytnutá k doménovému jménu či IP adrese je řetězec, je potřeba jej namapovat na číselnou hodnotu. Mapování je tedy následující:

- Zóna `Green` je mapována na číslo 0,
- zóna `Grey` je mapována na číslo 1,
- zóna `Yellow` je mapována na číslo 2,
- zóna `Orange` je mapována na číslo 3,
- zóna `Red` je mapována na číslo 4,
- jakákoliv jiná hodnota je mapována na číslo -1.

Mapování risku ze systému pulsedive.com

Položky `risk` a `risk_recommended` získané ze systému pulsedive.com jsou řetězce, tedy jejich možná hodnota je mapována na číselnou hodnotu následovně:

- Risk `none` je mapován na číslo 0,
- risk `low` je mapován na číslo 1,
- risk `medium` je mapován na číslo 2,
- risk `high` je mapován na číslo 3,
- risk `critical` je mapován na číslo 4,
- risk `retired` je mapován na číslo 5,
- risk `unknown` je mapován na číslo 6,
- jakákoliv jiná hodnota je mapována na číslo -1.

Mapování položky malicious ze systému radar.cloudflare.com

Aby bylo možné rozlišit, kdy položka `malicious` získaná ze systému Cloudflare Radar byla součástí odpovědi a kdy nikoliv, jsou data z kolektoru mapována takto:

- Položka `malicious` s hodnotou `True` je mapována na číslo 1,
- položka `malicious` s hodnotou `False` je mapována na číslo 0,
- položka `malicious` s hodnotou `None` je mapována na číslo -1,
- jakákoliv jiná hodnota je mapována na číslo -1.

8.4 Mock API

Jelikož některé API použitých reputačních systémů velice omezují počet dotazů, které mohou být systému zaslány, tak pro testovací účely bylo vytvořeno mock API. Toto mock API umožňuje vyzkoušet funkčnost jednotlivých kolektorů bez nutnosti obstarat si API klíč u daného reputačního systému a nachází se ve složce `domainradar/mock_api`. Toto API poskytuje odpovědi se stejnou strukturou jako API reputačního systému, avšak jsou zde uvedeny pouze položky, jež jsou kolektory sbírány. Přiřazené hodnoty v odpovědi mock API však neodpovídají skutečným datům, které by vrátilo API reputačního systému. Tato data nabývají jedné náhodně vybrané hodnotě ze všech možných hodnot, která daná položka daného reputačního systémů může nabývat. Toto vytvořené rozhraní taktéž zachovává strukturu cesty koncového bodu (*endpoint*) API reputačního systému pro získání dat.

Jelikož toto mock API je zabaleno do vlastního kontejneru a tento kontejner se spouští společně s celým nástrojem DomainRadar, tak pro napojení kolektoru na vytvořené mock API stačí změnit `BASE` adresu v kódu daného kolektoru z adresy API reputačního systému na `http://mock-api:5000/<název systému>/<endpoint>`. Příkladem u kolektoru dat ze systému VirusTotal je změna z adres:

`https://www.virustotal.com/api/v3/ip_addresses/`
`https://www.virustotal.com/api/v3/domains/`

na:

`http://mock-api:5000/virustotal/api/v3/ip_addresses/`
`http://mock-api:5000/virustotal/api/v3/domains/`

Pro správnou funkčnost je taktéž potřeba, aby v konfiguračním souboru daného kolektoru (složka `domainradar/infra/client_properties/`) byl nastaven token API klíče, ten může být nastaven na jakoukoliv hodnotu. Toto nastavení je potřeba z toho důvodu, že kdyby token nebyl nastaven, kolektor je považován za vypnutý.

Kapitola 9

Trénování a zhodnocení modelů

Z důvodu rozdělení klasifikace v nástroji DomainRadar na phishingové domény a domény existující za účelem šířit malware, je zapotřebí vytvořit modely dva. V této kapitole je popsán výběr modelů, proces trénování a jejich zhodnocení.

Trénování modelu pro detekci phishingových domén bylo provedeno na datové sadě obsahující 3 000 benigních domén a 3 000 phishingových domén. Model pro detekci malwarových domén byl trénován na sadě se stejnými 3 000 benignímí domény a s 3 000 malwarovými domény. Trénovací sada obsahuje 4 200 domén a testovací sada má 1 800 domén. Z důvodu omezeného počtu možných provedených dotazů některých reputačních systémů je získání dat o vícero doménových jménech a jejich IP adresách časově náročné. Vektor použitých příznaků je uveden v příloze B.

9.1 Výběr modelu

Pro účely zjištění, který model pro klasifikaci použít, byla využita knihovna PyCaret[1]. Díky ní bylo zjištěno na vícero datových sadách, že nejlepší a nejkonzistnější výsledky podává model `Decision Tree` a model `LightGBM`. Proto byly oba modely otestovány podrobněji. Ohodnocení na základě několika metrik a porovnání modelů z jednoho běhu trénování knihovnou PyCaret při využití datové sady obsahující příznaky o malwarových a benigních doménových jménech lze vidět v tabulce 9.1.

Model	Přesnost	AUC	Úplnost	Přesnost pozitivní třídy
LightGBM	0.9893	0.9979	0.9893	0.9894
Ada Boost	0.9881	0.9973	0.9881	0.9882
Logistic Regression	0.9879	0.9976	0.9879	0.9881
Decision Tree	0.9855	0.9884	0.9855	0.9856
K Neighbours	0.9693	0.9909	0.9693	0.9697
Naive Bayes	0.9669	0.9828	0.9669	0.9676
Ridge	0.9652	0.9963	0.9652	0.9675
SVM – Linear Kernel	0.9560	0.9903	0.9560	0.9580

Tabulka 9.1: Porovnání některých modelů na datové sadě obsahující malwarové a benigní doménová jména pomocí knihovny PyCaret

[1]PyCaret: https://pycaret.org/

Jelikož byl model `Decision Tree` v obou případech velice závislý na jediné vlastnosti (viz příloha C), tak by nebylo vhodné tento model použít a proto nejsou v této kapitole jeho výsledky uvedeny. V případě potenciálního neúspěchu získat data ze služby VirusTotal by klasifikace byla pravděpodobně špatná a výsledná klasifikace v nástroji DomainRadar by byla tímto zkreslena. Model `LightGBM` projevil daleko lepší závislost na vlastnostech, která je rozpoložena mezi daty z vícero reputačních systémů. Navíc model `LightGBM` obecně nabízí rychlý proces trénování bez velké ztráty přesnosti klasifikace výsledným modelem [21] a má nízkou spotřebu paměti. Model `LightGBM` je taktéž schopen pracovat s velkým počtem příznaků. Z těchto důvodů byly jak pro phishingové domény tak pro malwarové domény zvoleny a natrénovány modely `LightGBM`, které jsou zakomponovány do nástroje DomainRadar a jejich vlastnosti jsou zobrazeny níže.

9.2 Použité metriky k ohodnocení výsledného modelu

K zjištění efektivity a správného rozhodování klasifikačního modelu je potřeba určit nějakou metriku. K tomu se běžně používají následující vzorce a termíny [13, 33, 3]:

- `Threshold` – Hodnota mezi 0 a 1 umožňující modelu provést klasifikační rozhodnutí.
- `True Positive` (zkráceně `TP`) – Případ, kdy skóre je větší než hodnota `threshold` a opravdové ohodocení je 1.
- `False Positive` (zkráceně `FP`) – Případ, kdy skóre je větší než hodnota `threshold` a opravdové ohodnocení je 0.
- `True Negative` (zkráceně `TN`) – Případ, kdy skóre je nižší nebo rovno stanovené hodnotě `threshold` a opravdové ohodnocení je 0.
- `False Negative` (zkráceně `FN`) – Případ, kdy skóre je nižší nebo rovno stanovené hodnotě `threshold` a opravdové ohodnocení je 1.
- `Confusion matrix` (matice záměn) – Maticové zobrazení hodnot `TP`, `FP`, `TN` a `FN`.
- `Recall` (úplnost) – Hodnota udávající podíl pozitivních případů, které byly predikovány jako pozitivní. Tato metrika je vypočítána následujícím matematickým vzorcem:

$$Recall = \frac{TP}{TP + FN}$$

- `Precision` (přesnost pozitivní třídy) – Hodnota značící podíl správně identifikovaných pozitivních případů ze všech případů, které byly označeny za pozitivní. Metrika se počítá následovně:

$$Precision = \frac{TP}{TP + FP}$$

- `Accuracy` (přesnost) – Hodnota, která udává podíl případů, které byly správně identifikovány. Hodnota je počítána dle tohoto vzorce:

$$Accuracy = \frac{TP + TN}{TP + FN + TN + FP}$$

- `Receiver operating characteristic curve` (zkráceně `ROC`) – Grafické znázornění míry skutečně pozitivních ohodnocení oproti míře falešně pozitivních ohodnocení.

- `Area under the ROC curve` (zkráceně `AUC`) – Tato hodnota udává plochu pod křivkou `ROC` a sumarizuje jak moc dobrý je test bez ohledu na hodnotu `threshold`.
- `Log-loss` – Tato hodnota je založena na predikčních pravděpodobnostech, kde nižší hodnota značí lepší predikce.

9.3 Ladění hyperparametrů

Jak pro model detekce phishingových domén tak pro model určený k detekci malwarových domén bylo provedeno ladění hyperparametrů za účelem nalezení nejlepší konfigurace. Toho bylo dosaženo metodou *Grid Search* (mřížkové vyhledávání) z knihovny scikit-learn[2]. Hodnoty hyperparametrů, které byly při vyhledávání využity, jsou uvedeny v tabulce 9.2. Hodnotící metrikou pro určení nejlepší kombinace parametrů byla přesnost (*Accuracy*). Při vyhledávání nejlepší kombinace byla využita takzvaná strategie 5-fold cross-validation, při které je trénovací sada rozdělena do pěti disjunktních podmnožin [10]. Model je následně trénován na čtyřech podmnožinách, které společně představují trénovací sadu a poslední podmnožina je určená jako validační. Tento proces je opakován tolikrát, dokud každá podmnožina nebyla použita jako validační.

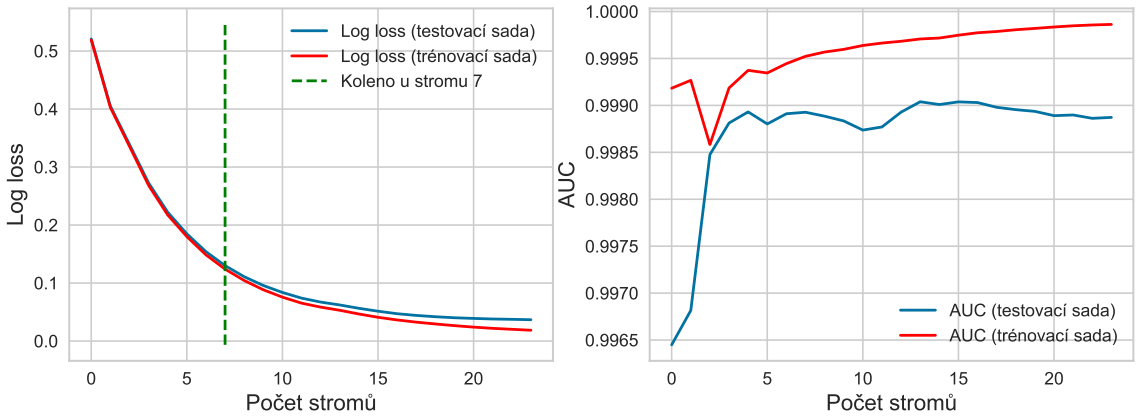
`max_depth`	8, 9, 10, 11, 12, 13, 14, 15, 20
`min_split_gain`	0.001, 0.01, 0.05
`n_estimators`	4, 6, 7, 8, 9, 10, 12, 14
`num_leaves`	5, 10, 15, 20, 25

Tabulka 9.2: Hodnoty hyperparametrů využitých k nalezení nejlepší kombinace s využitím mřížkového vyhledávání

9.4 LightGBM model pro detekci phishingových domén

V grafu 9.1 je zobrazena metrika `Log-loss` a `AUC` pro model detekce phishingových doménových jmen. Díky těmto grafům bylo jednodušší následně uzpůsobit parametry výsledného modelu. Aby výsledný model nebyl přetrénovaný, byly sledovány rozdíly výsledků na trénovací a testovací sadě na grafu zachycující Log-loss. Výsledné parametry jsou uvedeny v tabulce 9.3 a byly získány pomocí metody mřížkového vyhledávání, které je popsáno v sekci 9.3. Zvolené parametry pro ladění hyperparametrů jsou uvedeny v tabulce 9.2. Na obrázku 9.2 je vyhodnocení natrénovaného modelu se získanými parametry zobrazené pomocí matice záměn (*Confusion matrix*), kde vstupem byla testovací doménová jména. Třída 0 v tomto obrázku značí benigní domény a třída 1 označuje domény phishingové. Graf 9.3 dále zachycuje důležitost příznaků z reputačních systémů natrénovaného modelu. Výsledný natrénovaný model na testovací sadě dosahuje těchto hodnot: $Recall \approx 0.9830$, $Precision \approx 0.9965$, $Accuracy = 0.99$ a $AUC \approx 0.9993$.

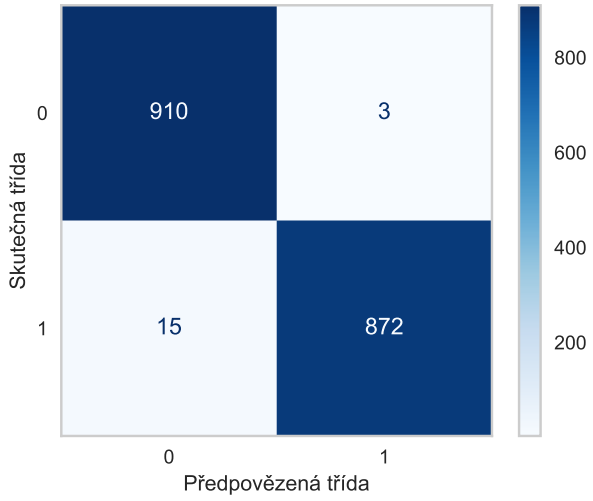
[2] scikit-learn: https://scikit-learn.org/



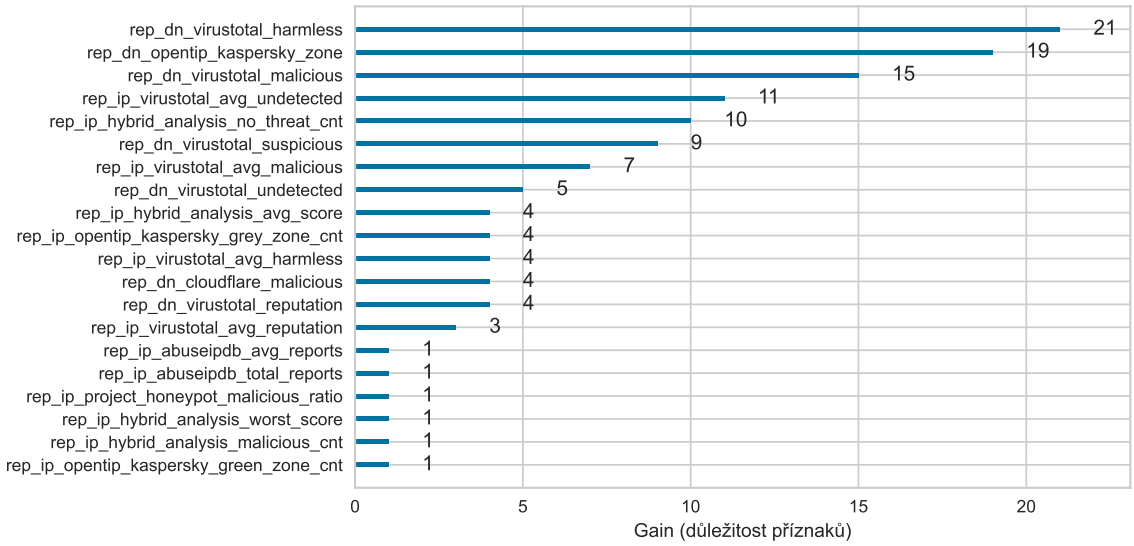
Obrázek 9.1: Metrika Log-loss a AUC modelu LightGBM pro klasifikaci phishingových domén závislá na počtu stromů

Parametr	**Nastavená hodnota**
objective	binary
boosting_type	gdbt
colsample_bytrate	1
learning_rate	0.2
max_depth	10
min_child_samples	20
min_split_gain	0.01
n_estimators	9
num_leaves	15

Tabulka 9.3: Parametry výsledného modelu LightGBM pro phishingové domény získané pomocí mřížkového vyhledávání



Obrázek 9.2: Confusion matrix výsledného modelu LightGBM pro phishingové domény (třída 0 = benigní domény, třída 1 = phishingové domény)



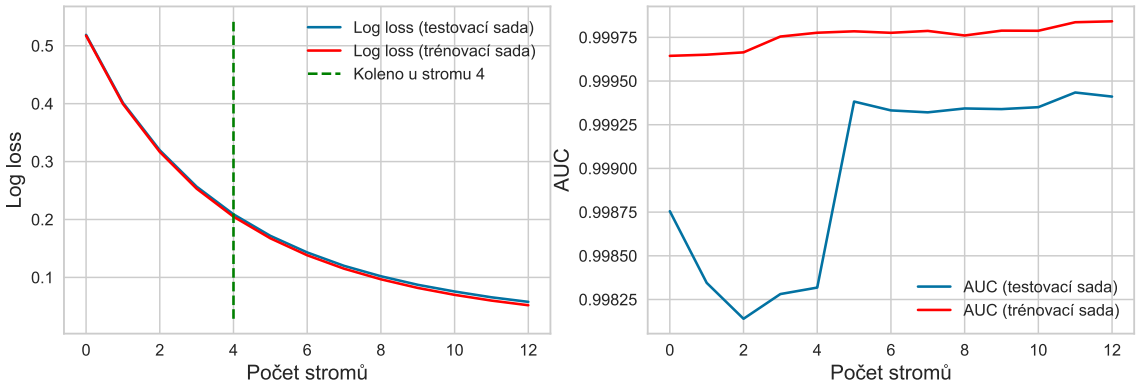
Obrázek 9.3: Důležitost příznaků výsledného modelu LightGBM pro phishingové domény

9.5 LightGBM model pro detekci malwarových domén

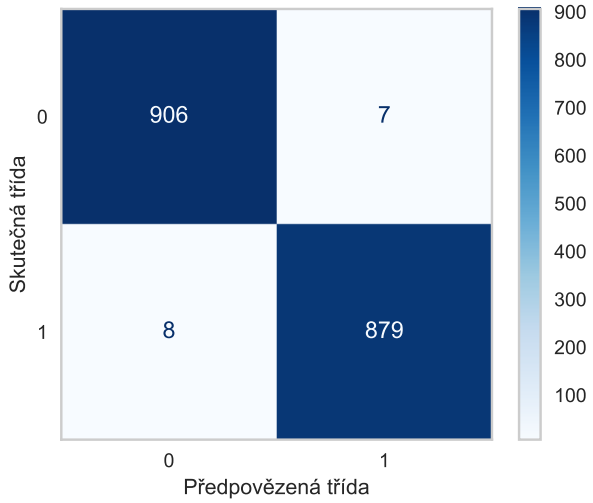
Graf 9.4 zobrazuje data metrik `Log-loss` a `AUC` pro model klasifikace malwarových doménových jmen. Znázornění těchto dat do grafů přispělo ke snazšímu uzpůsobení a ladění parametrů modelu pro trénování. Jednotlivé hodnoty parametrů byly získány pomocí metody mřížkového vyhledávání, které je společně s jeho využitím blíže popsáno v sekci 9.3. Využité hodnoty parametrů, které byly následně použity pro nalezení nejlepší konfigurace jsou uvedeny v tabulce 9.2. Konečné nastavené hodnoty modelu, nalezené pomocí metody mřížkového vyhledávání, jsou vypsané v tabulce 9.4. Aby se zajistilo, že výsledný model nebude přetrénovaný, byly sledovány rozdíly výsledků na trénovací a testovací datové sadě, které jsou zachyceny na Log-loss grafu. Obrázek 9.5 zachycuje *Confusion matrix* získanou z testovacích dat. Na tomto obrázku třída 0 zastupuje benigní domény a třída 1 naopak domény malwarové. Důležitost příznaků dat z reputačních systémů je zachycena grafem 9.6. Výsledný natrénovaný model na testovací sadě má tyto vlastnosti: $Recall \approx 0.9909$, $Precision \approx 0.9920$, $Accuracy = 0.9916$ a $AUC \approx 0.9983$.

Parametr	**Nastavená hodnota**
objective	binary
boosting_type	gdbt
colsample_bytrate	1
learning_rate	0.1
max_depth	10
min_child_samples	20
min_split_gain	0.001
n_estimators	7
num_leaves	20

Tabulka 9.4: Parametry výsledného modelu LightGBM pro malwarové domény po ručním ladění



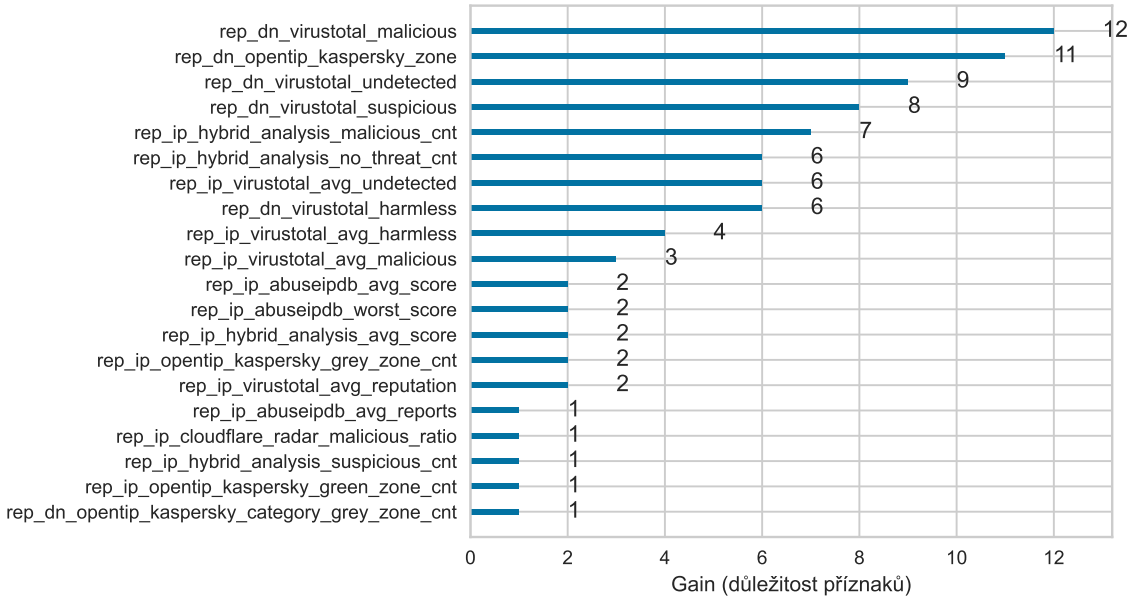
Obrázek 9.4: Metrika Log-loss a AUC modelu LightGBM pro klasifikaci malwarových domén závislá na počtu stromů



Obrázek 9.5: Confusion matrix výsledného modelu LightGBM pro malwarové domény (třída 0 = benigní domény, třída 1 = malwarové domény)

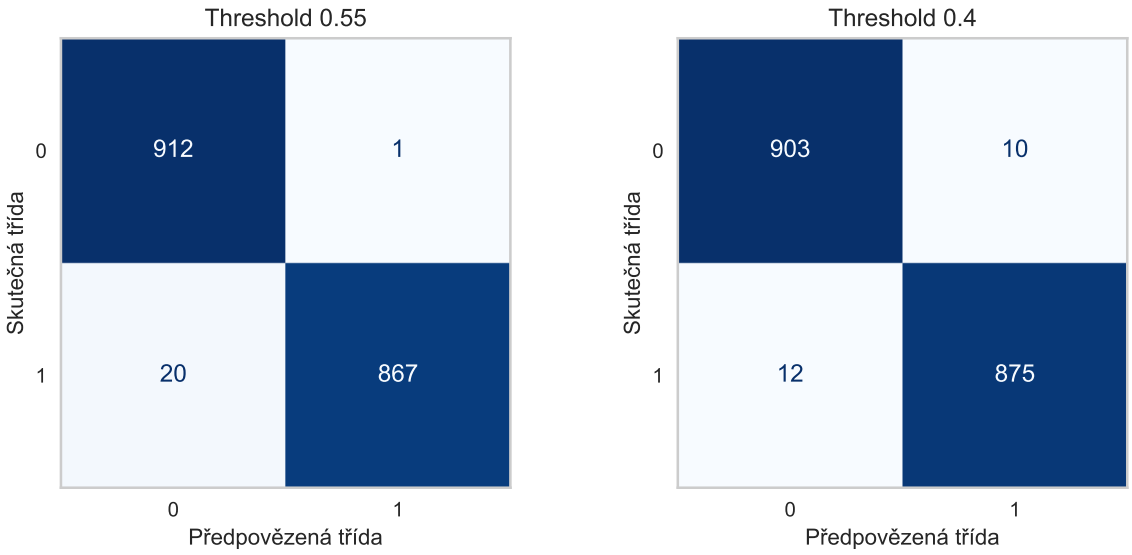
9.6 Citlivost vůči falešným výsledkům

Pro zjištění, zda natrénované modely nejsou příliš citlivé vůči falešně pozitivním či negativním výsledkům, byla u obou modelů změněna hodnota parametru `Threshold` z výchozí hodnoty $0,5$ na několik jiných různých hodnot. Na obrázku 9.7 lze vidět dvě matice záměn modelu LightGBM pro klasifikaci phishingových doménových jmen. I v tomto případě značí třída 0 benigní domény a třída 1 domény phishingové. Levá matice záměn odpovídá modelu s nastavenou hodnotou parametru `Threshold` na $0,55$. V tomto případě bylo více phishingových domén označeno za benigní, ale při tom počet benigních domén označených za phishing se moc nezměnil. Druhá matice záměn zachycuje výsledky modelu s nastavenou hodnotou parametru `Threshold` na $0,4$, jelikož při hodnotě $0,45$ nenastaly žádné změny. V tomto případě přibylo značně více počtu benigních domén označených za phishing, což je nežádoucí. Zatímco počet phishingových domén označených za benigní se moc nezměnil.



Obrázek 9.6: Důležitost příznaků výsledného modelu LightGBM pro malwarové domény

Jelikož výsledky klasifikátoru pro malwarové domény byly obdobné jako u klasifikátoru pro phishingové domény, nejsou zde zobrazeny matice záměn zvlášť i pro ně. Výchozí hodnota parametru se tak projevila jako nejlepší možnost, jelikož podávala nejpřesnější výsledky.



Obrázek 9.7: Matice záměn při jiném nastavení hodnoty parametru threshold u klasifikátoru phishingových doménových jmen

9.7 Zhodnocení

Trénování a testování klasifikátorů využívající model `LightGBM` probíhalo na stroji s procesorem Intel Core i7-10700K. Hledání nejlepší konfigurace hyperparametrů pomocí metody mřížkového vyhledávání trvalo u obou modelů necelých 12 minut (pro každý). Výsledné modely byly uloženy ve formátu joblib, kde model pro detekci phishingových doménových jmen zabírá téměř 27 kB a model pro detekci domén určených pro šíření malware nabývá velikosti necelých 21 kB. Testování rychlosti klasifikace bylo provedeno pro každý počet vzorků desetkrát s tím, že model i testovací sady byly již před testem načteny v paměti. Výsledné rychlosti klasifikace při tomto testování lze vidět v tabulce 9.5

Počet vzorků	**Průměrný čas běhu (ms)**	
	Phishingový klasifikátor	**Malwarový klasifikátor**
1	1.2033	1.1316
10	1.3046	1.4127
100	1.7049	1.7700
1000	1.8135	1.8081
6000	3.4332	3.7376

Tabulka 9.5: Průměrný čas potřebný ke klasifikaci různě velkých vzorků phishingových a malwarových doménových jmen

Oba výsledné klasifikátory používající model `LightGBM` dosahují velice slušných výsledků, což je zvláště pozoruhodné vzhledem k tomu, že většina reputačních systémů vykazovala výrazně nižší úspěšnost klasifikace (viz kapitola 6). Díky těmto výsledkům jsou tyto modely zakomponovány do nástroje DomainRadar, kde se podílí na klasifikaci doménových jmen společně s dalšími klasifikátory.

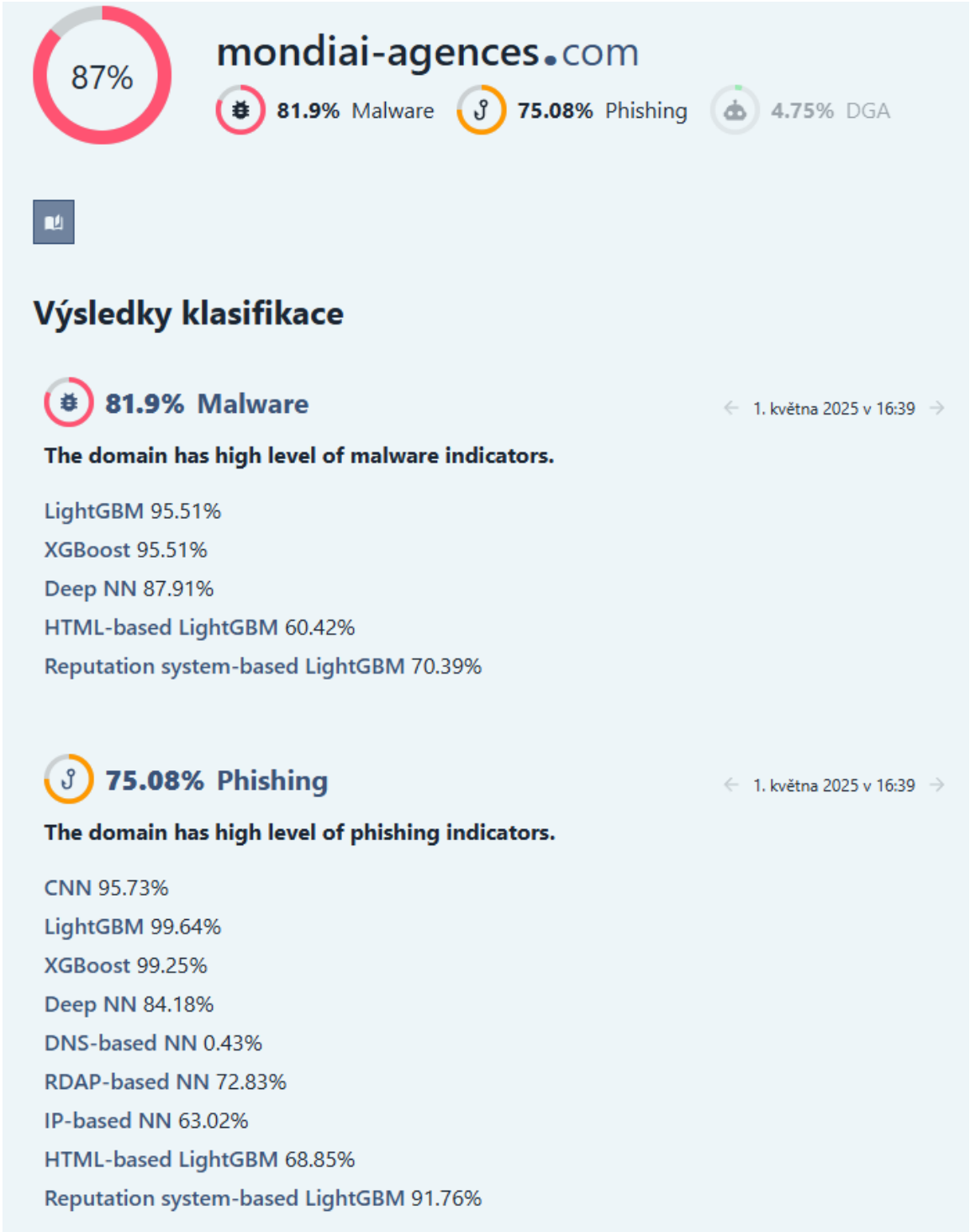
Kapitola 10

Experimentální ověření rozšíření

Experimentální ověření, že rozšíření nástroje DomainRadar o klasifikaci na základě dat z reputačních systémů funguje správně, spočívá v nasazení nástroje a vložení několika doménových jmen pomocí webového prostředí k celkové klasifikaci. Phishingové domény byly vybrány na základě potvrzeného nahlášení v systému PhishTank a malwarová doménová jména byla vybrána ze systému ThreatFox. Aby se bylo možné ujistit, že klasifikátory neovlivňují ohodnocení benigních domén, byly vybrány některé známé zpravodajské domény. Celkovou pravděpodobnost udělenou nástrojem DomainRadar vybraným doménám lze vidět v tabulce 10.1. Zde je taktéž uvedena celková pravděpodobnost udělená danému doménovému jménu nástrojem DomainRadar před rozšířením o klasifikaci na základě dat z reputačních systémů. Ukázka výsledného reportu jedné z vybraných domén, který lze spatřit ve webovém prostředí nástroje DomainRadar, je zobrazen na obrázku 10.1. Výsledky z klasifikátorů na základě dat z reputačních systémů jsou označeny jakožto `Reputation system-based LightGBM`. Na obrázku je možné vidět, že klasifikátor malwarových doménových jmen si je na 70.39 % jistý, že se jedná o malwarovou doménu a klasifikátor phishingových doménových jmen označil doménu jako phishing s 91.76% jistotou.

Typ doménového jména	Doménové jméno	Celková pravděpodobnost škodlivosti	
		Před rozšířením	Po rozšíření
phishing	mondiai-agences.com	76 %	87 %
phishing	groupuser.sbs	37 %	51 %
phishing	wedkl.cn	40 %	59 %
malware	mebwg.press	40 %	53 %
malware	tsoi-zhiv.com	76 %	77 %
malware	disciplipna.top	64 %	74 %
benigní	irozhlas.cz	0 %	0 %
benigní	denik.cz	0 %	0 %
benigní	bbc.com	0 %	0 %

Tabulka 10.1: Celková klasifikace doménových jmen před rozšířením a po rozšíření



Obrázek 10.1: Report z webového prostředí nástroje DomainRadar obsahující klasifikaci na základě dat z reputačních systémů

Zhodnocení a diskuze

Vytvořené klasifikátory dosahují velice slušných výsledků a ve spolupráci s ostatními klasifikátory v nástroji DomainRadar pomohly zvýšit přesnost konečného rozhodnutí výsledné pravděpodobnosti škodlivosti daného doménového jména. V kontextu práce to znamená, že se podařilo splnit požadovanou funkcionalitu, tedy detekci maligních domén s využitím reputačních systémů v nástroji DomainRadar. Taktéž bylo zjištěno, že agregovaná data z reputačních systémů jsou vhodná ke klasifikaci doménových jmen.

Pro uživatele nástroje DomainRadar rozšíření znamená další relevantní zdroj informací, díky kterému se mohou lépe spolehnout na výslednou klasifikaci poskytnutou nástrojem. A jelikož byly vybrány reputační systémy, které poskytují i bezplatný balíček, nemusí si uživatel zaplatit několik služeb, aby měl v nástroji DomainRadar tuto funkcionalitu.

Díky dosaženým výsledkům, které jsou velice slibné, mohou být vytvořené modely pro klasifikaci phishingových doménových jmen a domén určených k šíření malware užity i samostatně nebo v jiných nástrojích. Příkladem dalšího využití by mohlo být rozšíření do webového prohlížeče, které by využilo těchto modelů ke klasifikaci navštívené stránky uživatelem a mohlo by jej varovat, že přistupuje na doménu, která byla detekována jako maligní.

Námětem na rozšíření práce by mohlo být rozdělení sběru dat v nástroji DomainRadar na vícero stupňů. Doménové jméno by bylo klasifikováno v každém stupni postupně s přibývajícími daty s tím, že do dalšího stupně by se dostalo pouze v případě, když by klasifikaci v daném stupni nebylo možné spolehlivě určit. Klasifikace na základě dat z reputačních systémů by se tak mohla odsunout na pozdější stupeň, čímž by se do jisté míry mohla překonat restrikce počtu provedených dotazů, kterými jsou některé reputační systémy omezeny, protože k jasně benigním či maligním doménovým jménům by nebylo potřeba získávat data z reputačních systémů. Případně by mohly být jednotlivé kolektory dat z reputačních systémů rozděleny do vícero stupňů podle jejich restrikcí. To by znamenalo, že reputační systémy s menšími či žádnými restrikcemi by mohly být využity dříve a častěji, zatímco omezené reputační systémy by byly využity pouze v případě, kdy přesnou klasifikaci není možné spolehlivě určit.

Kapitola 11

Závěr

Cílem této práce bylo rozšířit nástroj DomainRadar o nový klasifikační faktor pro ohodnocování doménových jmen na základě získaných dat z reputačních systémů. Účelem tohoto faktoru bylo přispět k zpřesnění detekce škodlivých doménových jmen tím, že umožní využívat další informace o doménách a jejich přidružených IP adresách k následné klasifikaci daného doménového jména. Významnou výhodou je, že reputační systémy dokážou poskytnou nejen informace o aktuálním stavu doménového jména, ale taktéž i o jeho historii. Toto nástroji DomainRadar umožnilo lépe vyhodnotit riziko spojené s konkrétními doménami na základě historických aktivit, které by mohly jiným metodám uniknout.

Prvním zásadním bodem bylo vybrat reputační systémy, ze kterých budou čerpána data. V této práci bylo popsáno a experimentálně zhodnoceno celkem 14 volně dostupných reputačních systémů. Součástí těchto 14 reputačních systémů byl i systém nerd.cesnet.cz, z kterého jsou již data v nástroji DomainRadar získávána, avšak nejsou dále nijak využita. Dále bylo potřeba navrhnout, jakým způsobem budou kolektory dat z reputačních systémů v nástroji DomainRadar integrovány. Pro tyto účely vznikly dvě nová témata v systému Apache Kafka, které slouží ke kolekci dat o doménových jménech z reputačních systémů.

Dalším důležitým bodem bylo zvolit způsob, jakým bude klasifikace doménových jmen na základě dat z reputačních systémů prováděna. K těmto účelům byla vybrána klasifikace na základě strojového učení. Pro vybrání modelu byla použita knihovna PyCaret, kde vstupem bylo vícero datových sad, aby se bylo možné ujistit, že vybraný model bude podávat spolehlivé výsledky. Nejlépe hodnocenými modely se jak pro phishingové tak i pro malwarové domény staly modely `Decision Tree` a `LightGBM`. Natrénované modely `Decision Tree` však v obou případech vykazovaly vysokou závislost na jednom příznaku, a proto byl zvolen a natrénován model `LightGBM`, který následně prokazoval daleko lepší závislost na příznacích. Natrénovaný model pro detekci phishingových doménových jmen na testovací sadě dosáhl přesnosti 99 % a model pro klasifikaci malwarových domén měl na testovací sadě přesnost 99.16 %. Tyto modely byly následně zahrnuty do klasifikačního procesu nástroje DomainRadar.

V kontextu práce se podařilo zjistit, že agregovaná data z reputačních systémů jsou vhodná ke klasifikaci doménových jmen a také se podařilo splnit požadovanou funkcionalitu detekce maligních domén s využitím reputačních systémů v nástroji DomainRadar. Uživatel nástroje se tak bude moct lépe spolehnout na výslednou klasifikaci daného doménového jména. Jelikož dosažené výsledky obou vytvořených klasifikátorů jsou velice slibné, mohou být využity i mimo nástroj DomainRadar, například v rozšíření do webového prohlížeče, které by uživatele varovalo při přístupu na škodlivou doménu.

Cílem budoucí práce by mohlo být rozdělení sběru dat v nástroji DomainRadar na vícero stupňů, kde další stupeň by byl využit v případě, že klasifikaci doménového jména není možné jistě určit v daný moment. K jasně benigním či maligním doménám by tak nebylo třeba sbírat data z jiných zdrojů. To by umožnilo posunout sběr dat z reputačních systémů na pozdější stupeň a do jisté míry by se překonala restrikce počtu zaslaných dotazů, kterými jsou některé reputační systémy omezeny.

Literatura

[1] ABUSE.CH a SPAMHAUS. *ThreatFox: Share Indicators Of Comporomise (IOCs)* [online]. 2024. Dostupné z: https://threatfox.abuse.ch/. [cit. 2024-12-11].

[2] ABUSEIPDB LLC. *AbuseIPDB – IP address abuse reports – Making the Internet safer, one IP at a time* [online]. 2024. Dostupné z: https://www.abuseipdb.com/. [cit. 2024-12-11].

[3] AGGARWAL, A.; KASIVISWANATHAN, S.; XU, Z.; FEYISETAN, O. a TEISSIER, N. *Label Inference Attacks from Log-Loss Scores*. Květen 2021. Dostupné z: http://dx.doi.org/10.48550/arXiv.2105.08266.

[4] AI SPERA INC.. *CriminalIP: Cybersecurity Search Engine* [online]. 2024. Dostupné z: https://www.criminalip.io/. [cit. 2025-03-22].

[5] ANTONAKAKIS, M.; PERDISCI, R.; DAGON, D.; LEE, W. a FEAMSTER, N. Building a Dynamic Reputation System for DNS. In: *19th USENIX Security Symposium* [online]. Washington, D.C.: USENIX Association, Září 2010, s. 273–290. Dostupné z: https://www.usenix.org/legacy/event/sec10/tech/full_papers/Antonakakis.pdf. [cit. 2024-11-13].

[6] AO KASPERSKY LAB. *Kaspersky Threat Intelligence Portal* [online]. 2025. Dostupné z: https://opentip.kaspersky.com/. [cit. 2025-03-22].

[7] BARTOŠ, V. NERD: Network Entity Reputation Database. In: *Proceedings of the 14th International Conference on Availability, Reliability and Security*. New York, NY, USA: Association for Computing Machinery, 2019. ARES '19. ISBN 978-1-4503-7164-3. Dostupné z: https://doi.org/10.1145/3339252.3340512.

[8] BARTOŠ, V. *NERD – Network Entity Reputation Database* [online]. 2024. Dostupné z: https://nerd.cesnet.cz/. [cit. 2024-12-11].

[9] BARTOŠ, V.; ŽÁDNÍK, M.; HABIB, S. M. a VASILOMANOLAKIS, E. Network entity characterization and attack prediction. *Future Generation Computer Systems*, 2019, sv. 97, s. 674–686. Dostupné z: https://doi.org/10.1016/j.future.2019.03.016.

[10] BERRAR, D. Cross-Validation. In: RANGANATHAN, S.; GRIBSKOV, M.; NAKAI, K. a SCHÖNBACH, C., ed. *Encyclopedia of Bioinformatics and Computational Biology*. Oxford: Academic Press, 2019, s. 542–545. ISBN 978-0-12-811432-2. Dostupné z: https://doi.org/10.1016/B978-0-12-809633-8.20349-X.

[11] CHEN, C.-M.; HUANG, J.-J. a OU, Y.-H. Efficient suspicious URL filtering based on reputation. *Journal of Information Security and Applications*, 2015, sv. 20, s. 26–36. ISSN 2214-2126. Dostupné z: https://doi.org/10.1016/j.jisa.2014.10.005.

[12] Cloudflare, Inc.. *Cloudflare Radar* [online]. 2025. Dostupné z: https://radar.cloudflare.com/. [cit. 2025-03-22].

[13] Erickson, B. J. a Kitamura, F. Magician's corner: 9. Performance metrics for machine learning models. *Radiology: Artificial Intelligence*. Radiological Society of North America, 2021, sv. 3, č. 3, s. e200126. Dostupné z: https://doi.org/10.1148/ryai.2021200126.

[14] Fortinet, Inc.. *FortiGuard Labs: AntiSpam Service* [online]. 2025. Dostupné z: https://www.fortiguard.com/services/antispam. [cit. 2025-03-22].

[15] Google. *Google Safe Browsing* [online]. 2025. Dostupné z: https://safebrowsing.google.com/. [cit. 2025-03-22].

[16] GreyNoise, Inc.. *Greynoise Intelligence: Real-Time Intelligence For Modern Threats* [online]. 2025. Dostupné z: https://www.greynoise.io/. [cit. 2025-03-22].

[17] Hendrikx, F.; Bubendorfer, K. a Chard, R. Reputation systems: A survey and taxonomy. *Journal of Parallel and Distributed Computing*, 2015, sv. 75, s. 184–197. ISSN 0743-7315. Dostupné z: https://doi.org/10.1016/j.jpdc.2014.08.004.

[18] Hranický, R.; Horák, A. a Ondryáš, O. *A Dataset of Information (DNS, IP, WHOIS/RDAP, TLS, GeoIP) for a Large Corpus of Benign, Phishing, and Malware Domain Names 2024* [online]. 2024. Dostupné z: https://doi.org/10.5281/zenodo.13330074. [cit. 2024-11-13].

[19] Hranický, R.; Horák, A.; Polišenský, J.; Jeřábek, K. a Ryšavý, O. Unmasking the Phishermen: Phishing Domain Detection with Machine Learning and Multi-Source Intelligence. In: *NOMS 2024-2024 IEEE Network Operations and Management Symposium*. 2024, s. 1–5. Dostupné z: https://doi.org/10.1109/NOMS59830.2024.10575573.

[20] Hybrid Analysis. *Free Automated Malware Analysis Service* [online]. 2025. Dostupné z: https://www.hybrid-analysis.com/. [cit. 2025-03-22].

[21] Ke, G.; Meng, Q.; Finley, T.; Wang, T.; Chen, W. et al. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. In: Guyon, I.; Luxburg, U. V.; Bengio, S.; Wallach, H.; Fergus, R. et al., ed. *Advances in Neural Information Processing Systems*. Curran Associates, Inc., 2017, sv. 30.

[22] Newman, C. a Klyne, G. *Date and Time on the Internet: Timestamps* RFC 3339. RFC Editor, červenec 2002. Dostupné z: https://doi.org/10.17487/RFC3339.

[23] NoVirusThanks. *URLVoid: Check If a Website is Malicious/Scam or Safe/Legit* [online]. 2024. Dostupné z: https://www.urlvoid.com/. [cit. 2024-12-11].

[24] Ondryáš, O. *Efektivní rozsáhlý sběr informací o doménových jménech*. Brno, 2024. Diplomová práce. Vysoké učení technické v Brně, Fakulta informačních technologií. Vedoucí práce Hranický, R.

[25] Pulsedive LLC. *Threat Intelligence – Pulsedive* [online]. 2025. Dostupné z: https://pulsedive.com/. [cit. 2025-03-22].

[26] Resnick, P.; Zeckhauser, R. J.; Swanson, J. a Lockwood, K. The Value Of Reputation On Ebay. *A Controlled Experiment*, 2003, sv. 9, s. 79–101. Dostupné z: http://dx.doi.org/10.2139/ssrn.385206.

[27] Retail & Hospitality ISAC (RH-ISAC). *The Ease and Benefit of Automating Threat Intel with PyOTI* [online]. 2024. Dostupné z: https://rhisac.org/threat-intelligence/automation-with-pyoti/. [cit. 2024-12-15].

[28] Retail & Hospitality ISAC (RH-ISAC). *GitHub – RH-ISAC/PyOTI* [online]. 2024. Dostupné z: https://github.com/RH-ISAC/PyOTI. [cit. 2024-12-15].

[29] StrangeBee SAS.. *Cortex* [online]. 2024. Dostupné z: https://strangebee.com/cortex/. [cit. 2024-11-13].

[30] TheHive Project. *GitHub – TheHive-Project/Cortex* [online]. 2024. Dostupné z: https://github.com/TheHive-Project/Cortex. [cit. 2024-11-13].

[31] Unspam Technologies, Inc.. *The Web's Largest Community Tracking Online Fraud & Abuse* [online]. 2025. Dostupné z: https://www.projecthoneypot.org/index.php. [cit. 2025-03-22].

[32] VirusTotal. *Analyze suspicious files, domains, IPs and URLs to detect malware and other breaches* [online]. 2024. Dostupné z: https://www.virustotal.com/. [cit. 2024-12-11].

[33] Yang, T. a Ying, Y. AUC maximization in the era of big data and AI: A survey. *ACM computing surveys*. ACM New York, NY, 2022, sv. 55, č. 8, s. 1–37. Dostupné z: https://doi.org/10.1145/3554729.

Příloha A

Odpovědi API použitých reputačních systémů

Tato příloha obsahuje ukázku odpovědí získaných pomocí rozhraní API od použitých reputačních systémů. Pro jednoduchost jsou uvedeny pouze odpovídající maligním doménám respektive IP adresám. Jelikož jsem neobjevil ani doménu ani IP adresu, kterou by všechny reputační systémy označily za maligní, bude u každého reputačního systému uvedena doména nebo IP adresa, ke které se odpověď váže. API klíč je v dotazech nahrazen zástupným textem `<API key>`.

abuseipdb.com

Zvolená maligní IPv4 adresa: `3.33.130.190`

Dotaz

```
GET /api/v2/check?ipAddress=3.33.130.190 HTTP/1.1
Host: api.abuseipdb.com
Key: <API key>
```

Odpověď


```
{
    "data": {
        "ipAddress": "3.33.130.190",
        "isPublic": true,
        "ipVersion": 4,
        "isWhitelisted": false,
        "abuseConfidenceScore": 53,
        "countryCode": "US",
        "usageType": "Content Delivery Network",
        "isp": "Amazon Technologies Inc.",
        "domain": "amazon.com",
        "hostnames": [
            "a2aa9ff50de748dbe.awsglobalaccelerator.com"
        ],
        "isTor": false,
        "totalReports": 11,
        "numDistinctUsers": 5,
        "lastReportedAt": "2025-04-24T20:18:46+00:00"
    }
}
```

radar.cloudflare.com

Zvolené maligní doménové jméno: `only-fans.uk`

Dotaz

```
GET /client/v4/accounts/<User ID>/urlscanner/v2/search?size=1&q=page.domain:only-fans.uk HTTP/1.1
Host: api.cloudflare.com
Authorization: Bearer <API key>
```

Odpověď


```
{
    "results": [
        {
            "task": {
                "uuid": "e788e5c1-6a1b-4d29-a262-a2b0b741548b",
                "visibility": "public",
                "time": "2025-04-27T15:05:19.255000+00:00",
                "url": "https://only-fans.uk/Lalalalalalalala",
                "success": true
            },
            "page": {
                "url": "https://only-fans.uk/Lalalalalalalala",
                "country": "US",
                "asn": "AS13335",
                "ip": "104.21.41.236"
            },
            "stats": {
                "uniqIPs": 2,
                "uniqCountries": 1,
                "requests": 7,
                "dataLength": 198098
            },
            "verdicts": {
                "malicious": true
            },
            "_id": "e788e5c1-6a1b-4d29-a262-a2b0b741548b",
            "result": "https://radar.cloudflare.com/scan/e788e5c1-6a1b-4d29-a262-a2b0b741548b"
        }
    ]
}
```


criminalip.io

Zvolená maligní IPv4 adresa: `162.241.85.85`

Dotaz

```
GET /v1/asset/ip/report?ip=162.241.85.85?full=true HTTP/1.1
Host: api.criminalip.io
Accept: application/json
x-api-key: <API key>
```

Odpověď

Odpověď byla pro přehlednost upravena odstraněním dlouhých, ale ne zas tak podstatných částí. Odstraněné části jsou vyznačeny takto: `[...]`.

```
{
    "ip": "162.241.85.85",
    "issues": {
        "is_vpn": false,
        "is_cloud": false,
        "is_tor": false,
        "is_proxy": false,
        "is_hosting": false,
        "is_mobile": false,
        "is_darkweb": false,
        "is_scanner": false,
        "is_snort": false
    },
    "score": {
        "inbound": "Critical",
        "outbound": "Dangerous"
    },
    "user_search_count": 1,
    "domain": {
        "count": 3,
        "data": [
            [...]
        ]
    },
    "whois": {
        "count": 0,
        "data": []
    },
    "hostname": {
        "count": 0,
        "data": []
    },
    "ids": {
        "count": 0,
        "data": []
    },
    "vpn": {
        "count": 0,
        "data": []
    },
    "webcam": {
        "count": 0,
        "data": []
    },
    "honeypot": {
        "count": 0,
        "data": []
    },
    "ip_category": {
        "count": 0,
        "data": []
    },
    "port": {
        "count": 0,
        "data": []
    },
    "vulnerability": {
        "count": 2,
        "data": [
            [...]
        ]
    },
    "mobile": {
        "count": 0,
        "data": []
    },
    "status": 200
}
```

fortiguard.com

Zvolené maligní doménové jméno: `public.trzor.us`

Dotaz

```
POST /learnmore/check-blocklist HTTP/1.1
Host: www.fortiguard.com
Content-Type: application/json
Content-Length: 29

{
    "url": "public.trzor.us"
}
```

Odpověď

```
{
    "spam": true
}
```

safebrowsing.google.com

Zvolená maligní URL: `https://testsafebrowsing.appspot.com/s/phishing.html`

Dotaz

```
POST /v4/threatMatches:find?key=<API key> HTTP/1.1
Host: safebrowsing.googleapis.com
Content-Type: application/json
Content-Length: 464

{
    "client": {
        "clientId": "domrad-bp",
        "clientVersion": "1.0"
    },
    "threatInfo": {
        "threatTypes": ["THREAT_TYPE_UNSPECIFIED", "MALWARE", "SOCIAL_ENGINEERING", "UNWANTED_SOFTWARE
            ", "POTENTIALLY_HARMFUL_APPLICATION"],
        "platformTypes": ["ALL_PLATFORMS"],
        "threatEntryTypes": ["URL"],
        "threatEntries": [
            {"url": "https://testsafebrowsing.appspot.com/s/phishing.html"}
        ]
    }
}
```

Odpověď

```
{
    "matches": [
        {
            "threatType": "SOCIAL_ENGINEERING",
            "platformType": "ALL_PLATFORMS",
            "threat": {
                "url": "https://testsafebrowsing.appspot.com/s/phishing.html"
            },
```

```
            "cacheDuration": "300s",
            "threatEntryType": "URL"
        }
    ]
}
```


greynoise.io

Zvolená maligní IPv4 adresa: `188.114.97.3`

Dotaz

```
GET /v3/community/188.114.97.3 HTTP/1.1
Host: api.greynoise.io
Accept: application/json
key: <API key>
```

Odpověď


```
{
    "ip": "188.114.97.3",
    "noise": true,
    "riot": true,
    "classification": "unknown",
    "name": "unknown",
    "link": "https://viz.greynoise.io/ip/188.114.97.3",
    "last_seen": "2025-04-27",
    "message": "Success"
}
```


hybrid-analysis.com

Zvolená maligní IPv4 adresa: `3.33.130.190`

Dotaz

```
POST /api/v2/search/terms HTTP/1.1
Host: www.hybrid-analysis.com
api-key: <API key>
Content-Type: application/x-www-form-urlencoded
Content-Length: 17

host=104.21.1.162
```

Odpověď


```
{
    "search_terms": [
        {
            "id": "host",
            "value": "104.21.1.162"
        }
    ],
    "count": 2,
    "result": [
        {
```

```
            "verdict": "suspicious",
            "av_detect": "28",
            "threat_score": 68,
            "vx_family": null,
            "job_id": "67f432764e79c3b9650704c0",
            "sha256": "ff41bc558de958cdeff0d3685543f59a39f7ea87449530857a400549aca92aa7",
            "environment_id": 160,
            "analysis_start_time": "2025-04-07 20:15:50",
            "submit_name": "https://carlwasnagi2-hub.netlify.app/",
            "environment_description": "Windows 10 64 bit",
            "size": 61,
            "type": null,
            "type_short": "url"
        },
        {
            "verdict": "malicious",
            "av_detect": "25",
            "threat_score": 100,
            "vx_family": null,
            "job_id": "6786ac4647a471891e077763",
            "sha256": "efa9cb1ece1fe4765103c157c8c8003c653661edc8ef7098b7fd9f19ed650e3a",
            "environment_id": 160,
            "analysis_start_time": "2025-01-14 18:26:14",
            "submit_name": "https://www.lolex.net/",
            "environment_description": "Windows 10 64 bit",
            "size": 46,
            "type": null,
            "type_short": "url"
        }
    ]
}
```


nerd.cesnet.cz

Zvolená maligní IPv4 adresa: `88.214.48.19`

Dotaz

```
GET /nerd/api/v1/ip/88.214.48.19 HTTP/1.1
Host: nerd.cesnet.cz
Accept: application/json
Authorization: <API key>
```

Odpověď


```
{
    "asn": [
        213355
    ],
    "bgppref": "88.214.48.0/24",
    "bl": [
        "blocklist_de-ssh",
        "uceprotect",
        "abuseipdb",
        "turris_greylist",
        "spamhaus-sbl",
        "spamhaus-drop",
        "dshield"
    ],
    "fmp": {
        "general": 0.0
    },
```

```
    "geo": {
        "ctry": "US"
    },
    "hostname": "",
    "ip": "88.214.48.19",
    "ipblock": "88.214.48.0 - 88.214.51.255",
    "rep": 0.9764136053266979,
    "tags": [
        {
            "c": 1,
            "n": "attemptlogin"
        },
        {
            "c": 1,
            "n": "reconscanning"
        }
    ]
}
```


opentip.kaspersky.com

Zvolené maligní doménové jméno: `redteamminepool.ug`

Dotaz

```
GET /api/v1/search/domain?request=redteamminepool.ug HTTP/1.1
Host: opentip.kaspersky.com
x-api-key: <API key>
```

Odpověď


```
{
    "Zone": "Red",
    "DomainGeneralInfo": {
        "FilesCount": 10,
        "UrlsCount": 10,
        "HitsCount": 1000,
        "Domain": "redteamminepool.ug",
        "Ipv4Count": 5,
        "Categories": [
            "CATEGORY_MALWARE"
        ],
        "CategoriesWithZone": [
            {
                "Name": "CATEGORY_MALWARE",
                "Zone": "Red"
            }
        ]
    },
    "DomainWhoIsInfo": {
        "DomainName": "redteamminepool.ug",
        "Created": "2021-05-17T21:00:00Z",
        "Expires": "2022-05-17T21:00:00Z",
        "NameServers": [
            "ns101.cloudns.net",
            "ns102.cloudns.net",
            "ns103.cloudns.net",
            "ns104.cloudns.net"
        ],
        "Contacts": [
            {
                "ContactType": "registrant",
```

```
                "Name": "N/A",
                "Organization": "N/A",
                "Address": "N/A",
                "City": "N/A",
                "State": "N/A",
                "PostalCode": "N/A",
                "CountryCode": "N/A",
                "Email": "admin@redteamminepool.ug"
            },
            {
                "ContactType": "administrative",
                "Name": "N/A",
                "Organization": "N/A",
                "Address": "N/A",
                "City": "N/A",
                "State": "N/A",
                "PostalCode": "N/A",
                "CountryCode": "N/A",
                "Email": "admin@redteamminepool.ug"
            },
            {
                "ContactType": "technical",
                "Name": "N/A",
                "Organization": "N/A",
                "Address": "N/A",
                "City": "N/A",
                "State": "N/A",
                "PostalCode": "N/A",
                "CountryCode": "N/A",
                "Email": "admin@redteamminepool.ug"
            }
        ],
        "Registrar": {},
        "DomainStatus": [
            "ACTIVE"
        ]
    }
}
```


pulsedive.com

Zvolené maligní doménové jméno: `post-gieu3.top`

Dotaz

```
GET /api/info.php?key=<API key>&indicator=post-gieu3.top&pretty=1 HTTP/1.1
Host: pulsedive.com
Accept: application/json
```

Odpověď


```
{
    "qid": null,
    "iid": 58480864,
    "indicator": "post-gieu3.top",
    "type": "domain",
    "risk": "high",
    "risk_recommended": "high",
    "manualrisk": 0,
    "retired": "No recent activity",
    "stamp_added": "2024-04-18 06:40:28",
    "stamp_updated": "2024-09-03 20:06:46",
```

```
    "stamp_seen": "2024-06-02 12:19:08",
    "stamp_probed": null,
    "stamp_retired": "2024-09-03 20:06:46",
    "recent": 0,
    "submissions": 0,
    "umbrella_rank": null,
    "umbrella_domain": null,
    "riskfactors": [
        {
            "rfid": 6,
            "description": "does not resolve to an IP",
            "risk": "medium"
        },
        {
            "rfid": 51,
            "description": "uncommon TLD",
            "risk": "medium"
        }
    ],
    "redirects": {
        "from": [],
        "to": []
    },
    "threats": [],
    "feeds": [],
    "comments": [],
    "attributes": [],
    "properties": []
}
```


threatfox.abuse.ch

Zvolené maligní doménové jméno: `ns1.dmakk.cn`

Dotaz

```
POST /api/v1/ HTTP/1.1
Host: threatfox-api.abuse.ch
Auth-Key: <API key>
Content-Type: application/json
Content-Length: 94

{
    "query": "search_ioc",
    "search_term": "ns1.dmakk.cn",
    "exact_match": false
}
```

Odpověď


```
{
    "query_status": "ok",
    "data": [
        {
            "id": "1514191",
            "ioc": "ns1.dmakk.cn",
            "threat_type": "botnet_cc",
            "threat_type_desc": "Indicator that identifies a botnet command&control server (C&C)",
            "ioc_type": "domain",
            "ioc_type_desc": "Domain that is used for botnet Command&control (C&C)",
            "malware": "win.cobalt_strike",
            "malware_printable": "Cobalt Strike",
```

```
            "malware_alias": "Agentemis,BEACON,CobaltStrike,cobeacon",
            "malware_malpedia": "https://malpedia.caad.fkie.fraunhofer.de/details/win.cobalt_strike",
            "confidence_level": 75,
            "first_seen": "2025-04-30 20:54:47 UTC",
            "last_seen": null,
            "reference": null,
            "reporter": "abuse_ch",
            "tags": [
                "CobaltStrike",
                "drb-ra"
            ],
            "malware_samples": []
        }
    ]
}
```


virustotal.com

Zvolené maligní doménové jméno: `ns1.dmakk.cn`

Dotaz

```
GET /api/v3/domains/right.bestresulttostart.com HTTP/1.1
Host: www.virustotal.com
Accept: application/json
x-apikey: <API key>
```

Odpověď

Odpověď byla pro přehlednost upravena odstraněním dlouhých, ale ne zas tak podstatných, částí. Odstraněné části jsou vyznačeny takto: `[...]`


```
{
    "data": {
        "id": "right.bestresulttostart.com",
        "type": "domain",
        "links": {
            "self": "https://www.virustotal.com/api/v3/domains/right.bestresulttostart.com"
        },
        "attributes": {
            "last_modification_date": 1746033887,
            "categories": {
                "alphaMountain.ai": "Malicious (alphaMountain.ai)"
            },
            "expiration_date": 1772636077,
            "last_dns_records_date": 1743933812,
            "last_analysis_date": 1745136168,
            "last_update_date": 1741162926,
            "jarm": "27d40d40d00040d1dc42d43d00041d6183ff1bfae51ebd88d70384363d525c",
            "last_analysis_results": {
                [...]
            },
            "creation_date": 1709564090,
            "last_dns_records": [
                [...]
            ],
            "total_votes": {
                "harmless": 0,
                "malicious": 1
            },
            "popularity_ranks": {
                "Cisco Umbrella": {
```

```
                "rank": 513871,
                "timestamp": 1712681888
            }
        },
        "registrar": "NICENIC INTERNATIONAL GROUP CO., LIMITED",
        "tld": "com",
        "last_https_certificate_date": 1743933812,
        "tags": [],
        "reputation": -55,
        "last_analysis_stats": {
            "malicious": 11,
            "suspicious": 0,
            "undetected": 30,
            "harmless": 53,
            "timeout": 0
        },
        "last_https_certificate": {
            [...]
        }
    }
  }
}
```

Příloha B

Seznam příznaků

V této příloze jsou uvedeny vybrané příznaky získaných dat z reputačních systémů, které jsou využity pro klasifikaci doménových jmen. Příznaky jsou rozděleny na příznaky získaných o doménovém jméně a na příznaky získaných z přidružených IP adres.

Příznaky o doménovém jméně založené na reputačních systémech (rep_dn_)	
Název	**Popis**
virustotal_reputation	Přidělená reputace
virustotal_malicious	Počet označení domény za `malicious`
virustotal_suspicious	Počet označení domény za `suspicious`
virustotal_undetected	Počet označení domény za `undetected`
virustotal_harmless	Počet označení domény za `harmless`
opentip_kaspersky_zone	Přidělená zóna převedená na číslo
opentip_kaspersky_category_green_zone_cnt	Počet `Green` kategorií
opentip_kaspersky_category_grey_zone_cnt	Počet `Grey` kategorií
opentip_kaspersky_category_yellow_zone_cnt	Počet `Yellow` kategorií
opentip_kaspersky_category_orange_zone_cnt	Počet `Orange` kategorií
opentip_kaspersky_category_red_zone_cnt	Počet `Red` kategorií
opentip_kaspersky_is_category_adware	Booleovská hodnota značící, zda je doména označená jako `ADWARE`
opentip_kaspersky_is_category_phishing	Booleovská hodnota značící, zda je doména označená jako `PHISHING`
opentip_kaspersky_is_category_malware	Booleovská hodnota značící, zda je doména označená jako `MALWARE`
opentip_kaspersky_is_category_compromised	Booleovská hodnota značící, zda je doména označená jako `COMPROMISED`
opentip_kaspersky_is_category_botnet_cnc	Booleovská hodnota značící, zda je doména označená jako `BOTNET_CNC`
hybrid_analysis_malicious_cnt	Počet označení domén jako `malicious`
hybrid_analysis_suspicious_cnt	Počet označení domén jako `suspicious`
hybrid_analysis_no_threat_cnt	Počet označení domén jako `no threat`
hybrid_analysis_whitelisted_cnt	Počet označení domén jako `whitelisted`
hybrid_analysis_worst_score	Nejhorší skóre přidělené doméně
hybrid_analysis_best_score	Nejlepší skóre přidělené doméně
hybrid_analysis_avg_score	Průměrné skóre všech přidělených skóre doméně
hybrid_analysis_null_score_cnt	Počet nahlášení, u kterých nebylo uvedeno skóre
google_safe_browsing_unspecified_cnt	Počet nahlášení, u kterých není uvedena přesná hrozba
google_safe_browsing_malware_cnt	Počet nahlášení s kategorií `malware`
google_safe_browsing_social_engineering_cnt	Počet nahlášení s kategorií `social engineering`
google_safe_browsing_unwanted_software_cnt	Počet nahlášení s kategorií `unwanted software`

pokračování na další straně

Název	Popis
google_safe_browsing_potentially_harmful_cnt	Počet nahlášení s kategorií `potentially harmful`
urlvoid_detection_cnt	Počet detekcí
cloudflare_malicious	Zda se jedná o maligní doménu (či žádné informace) mapované na číslo
fortiguard_spam	Booleovská hodnota značící, zda je doména označená za spam
threatfox_malware_confidence_level	Míra jistoty v případě, že doména byla označen za malware, jinak -1
pulsedive_risk	Hodnota risku mapovaná na číslo
pulsedive_risk_recommended	Hodnota doporučeného risku mapovaná na číslo
pulsedive_manual_risk	Hodnota manuálního risku

Tabulka B.1: Seznam příznaků o doménovém jméně založených na datech z reputačních systémů pro klasifikaci

Příznaky o IP adresách založené na reputačních systémech (rep_ip_)	
Název	**Popis**
virustotal_avg_reputation	Průměrná přidělená reputace
virustotal_avg_malicious	Průměrný počet označení IP adresy za `malicious`
virustotal_avg_suspicious	Průměrný počet označení IP adresy za `suspicious`
virustotal_avg_undetected	Průměrný počet označení IP adresy za `undetected`
virustotal_avg_harmless	Průměrný počet označení IP adresy za `harmless`
opentip_kaspersky_green_zone_cnt	Počet zón IP adres označených jako `Green`
opentip_kaspersky_grey_zone_cnt	Počet zón IP adres označených jako `Grey`
opentip_kaspersky_yellow_zone_cnt	Počet zón IP adres označených jako `Yellow`
opentip_kaspersky_orange_zone_cnt	Počet zón IP adres označených jako `Orange`
opentip_kaspersky_red_zone_cnt	Počet zón IP adres označených jako `Red`
opentip_kaspersky_category_green_zone_cnt	Celkový počet `Green` kategorií
opentip_kaspersky_category_grey_zone_cnt	Celkový počet `Grey` kategorií
opentip_kaspersky_category_yellow_zone_cnt	Celkový počet `Yellow` kategorií
opentip_kaspersky_category_orange_zone_cnt	Celkový počet `Orange` kategorií
opentip_kaspersky_category_red_zone_cnt	Celkový počet `Red` kategorií

pokračování na další straně

Název	Popis
opentip_kaspersky_category_adware_cnt	Celkový počet IP adres označených jako `ADWARE`
opentip_kaspersky_category_phishing_cnt	Celkový počet IP adres označených jako `PHISHING`
opentip_kaspersky_category_malware_cnt	Celkový počet IP adres označených jako `MALWARE`
opentip_kaspersky_category_compromised_cnt	Celkový počet IP adres označených jako `COMPROMISED`
opentip_kaspersky_category_botnet_cnc_cnt	Celkový počet IP adres označených jako `BOTNET_CNC`
hybrid_analysis_malicious_cnt	Celkový počet IP adres označených jako `malicious`
hybrid_analysis_suspicious_cnt	Celkový počet IP adres označených jako `suspicious`
hybrid_analysis_no_threat_cnt	Celkový počet IP adres označených jako `no threat`
hybrid_analysis_whitelisted_cnt	Celkový počet IP adres označených jako `whitelisted`
hybrid_analysis_worst_score	Nejhorší skóre všech IP adres
hybrid_analysis_best_score	Nejlepší skóre všech IP adres
hybrid_analysis_avg_score	Průměrné skóre všech přidělených skóre všech IP adres
hybrid_analysis_null_score_cnt	Počet nahlášení, u kterých nebylo uvedeno skóre
google_safe_browsing_total_unspecified_cnt	Celkový počet nahlášení IP adres, u kterých není uvedena přesná hrozba
google_safe_browsing_total_malware_cnt	Celkový počet nahlášení IP adres s kategorií `malware`
google_safe_browsing_total_social_engineering_cnt	Celkový počet nahlášení IP adres s kategorií `social engineering`
google_safe_browsing_total_unwanted_software_cnt	Celkový počet nahlášení IP adres s kategorií `unwanted software`
google_safe_browsing_total_potentially_harmful_cnt	Celkový počet nahlášení IP adres s kategorií `potentially harmful`
nerd_max_score	Nejhorší skóre IP adresy
project_honeypot_malicious_ratio	Poměr IP adres označených jako `malicious` k počtu IP adres
cloudflare_radar_malicious_ratio	Poměr IP adres označených jako `malicious` k počtu IP adres
abuseipdb_worst_score	Nejhorší skóre všech IP adres
abuseipdb_avg_score	Průměrné skóre všech IP adres
abuseipdb_whitelisted_ratio	Poměr IP adres označených jako `whitelisted` k počtu IP adres
abuseipdb_tor_ratio	Poměr IP adres označených jako `tor` k počtu IP adres
abuseipdb_total_reports	Celkový počet nahlášení všech IP adres
abuseipdb_avg_reports	Průměrná hodnota nahlášení na IP adresu
fortiguard_spam_ratio	Poměr IP adres označených jako `spam` k počtu IP adres
greynoise_noise_ratio	Poměr IP adres označených jako `noise` k počtu IP adres
greynoise_riot_ratio	Poměr IP adres označených jako `riot` k počtu IP adres

pokračování na další straně

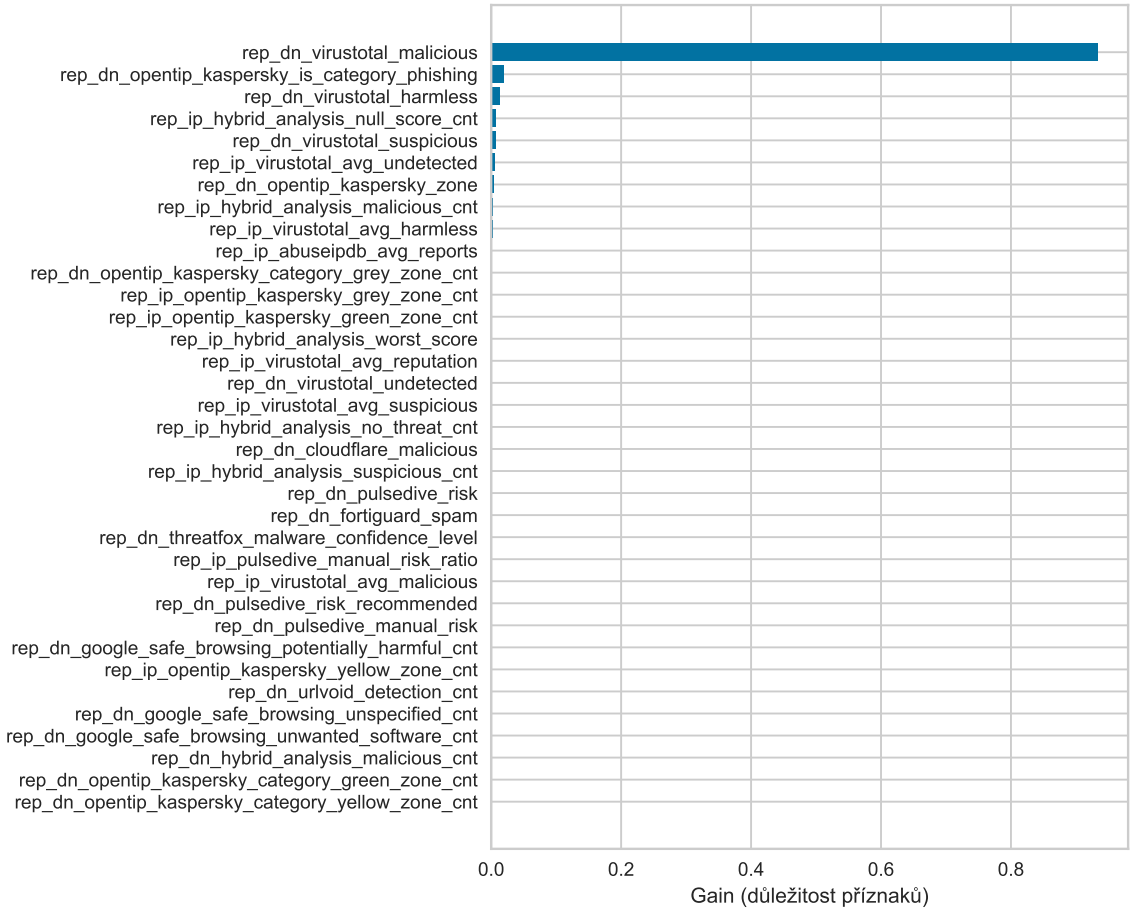
Název	**Popis**
greynoise_total_benign	Počet IP adres označených jako `benign`
greynoise_total_malicious	Počet IP adres označených jako `malicious`
criminalip_score_safe_cnt	Počet IP adres se skórem `safe`
criminalip_score_low_cnt	Počet IP adres se skórem `low`
criminalip_score_moderate_cnt	Počet IP adres se skórem `moderate`
criminalip_score_dangerous_cnt	Počet IP adres se skórem `dangerous`
criminalip_score_critical_cnt	Počet IP adres se skórem `critical`
threatfox_malware_ratio	Poměr IP adres s detekovaným malwarem k počtu IP adres
threatfox_avg_confidence_level	Průměrná jistota detekovaného malwaru (-1 pokud není uvedeno)
pulsedive_risk_none_cnt	Počet IP adres se skórem `none` (`risk` + `risk_recommended`)
pulsedive_risk_low_cnt	Počet IP adres se skórem `low` (`risk` + `risk_recommended`)
pulsedive_risk_medium_cnt	Počet IP adres se skórem `medium` (`risk` + `risk_recommended`)
pulsedive_risk_high_cnt	Počet IP adres se skórem `high` (`risk` + `risk_recommended`)
pulsedive_risk_critical_cnt	Počet IP adres se skórem `critical` (`risk` + `risk_recommended`)
pulsedive_risk_retired_cnt	Počet IP adres se skórem `retired` (`risk` + `risk_recommended`)
pulsedive_risk_unknown_cnt	Počet IP adres se skórem `unknown` (`risk` + `risk_recommended`)
pulsedive_manual_risk_ratio	Poměr IP adres s manuálním riskem 1 k počtu IP adres

Tabulka B.2: Seznam příznaků o IP adresách založených na datech z reputačních systémů pro klasifikaci

Příloha C

Důležitost příznaků modelu Decision Tree

Zde jsou zobrazeny důležitosti příznaků dat z reputačních systémů u modelu `Decision Tree`. Modely jsou silně závislé na jediném příznaku. Většina ostatních příznaků má velice zanedbatelnou důležitost, že v grafu C.1 a v grafu C.2 nejde přínos okem zaznamenat.



Obrázek C.1: Důležitost příznaků modelu Decision Tree pro phishingové domény

Obrázek C.2: Důležitost příznaků modelu Decision Tree pro malwarové domény

Příloha D

Obsah paměťového média

```
/
├── bp_latex/ ..... Zdrojové soubory textu této práce
├── dataset/ ..... Datové sady
├── domainradar/ ..... Nástroj DomainRadar s rozšířením
│   ├── colext/
│   ├── domainradar/
│   ├── infra/
│   ├── input/
│   ├── mock_api/
│   ├── webui/
│   └── soubory.md . Výpis upravených a vytvořených souborů v nástroji DomainRadar
├── notebooks/ ..... Jupyter notebooky
├── scripts/ ..... Pomocné skripty
├── bp.pdf ..... Tato práce ve formátu PDF
├── bp_tisk.pdf ..... Tato práce ve formátu PDF určená pro tisk
└── README.md ..... Popis projektu a jeho spuštění
```