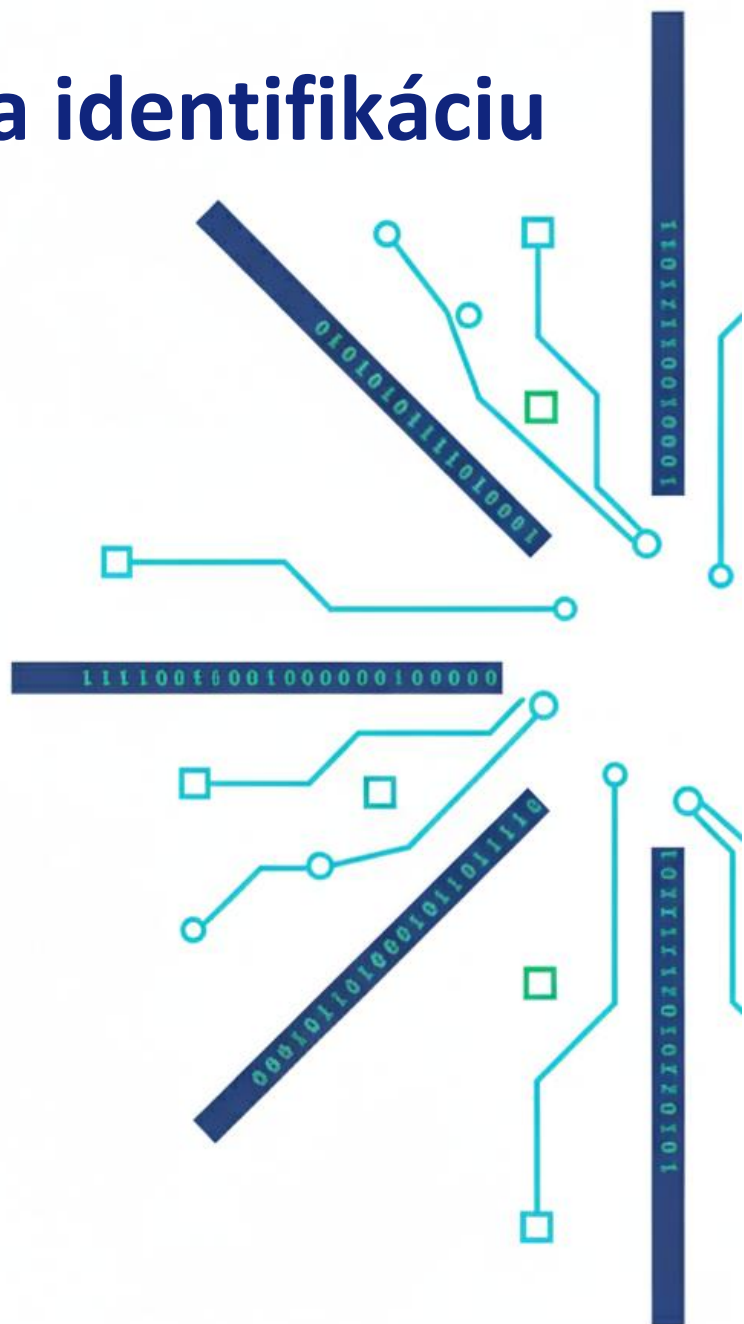


D19 – Metóda na identifikáciu digitálnych stôp



Projekt Automatizácia digitálnej forenznej analýzy a odpovede na incident (ADFIR) financovaný Európskou úniou – Next GenerationEU prostredníctvom Plánu obnovy a odolnosti Slovenskej republiky pod číslom projektu č. 09I05-03-V02-00079.

OBSAH

Popis projektu	3
Úvod	4
1 Teoretický úvod	5
1.1 Detekcia anomálií	6
1.2 Detekcia anomálií bez učiteľa (Unsupervised)	6
1.3 Typy anomálií.....	6
1.3.1 Bodové anomálie (point anomalies)	6
1.3.2 Kontextové anomálie (Contextual anomalies)	7
1.3.3 Kolektívne anomálie (Collective anomalies)	7
1.4 Kategorizácia metód pre detekciu anomálií.....	7
2 Aplikácia vybraných metód	9
2.1 Detekcia anomálií pomocou metód strojového učenia bez učiteľa.....	9
2.2 Dátové súbory a predspracovanie	9
2.3 Použité metódy detekcie anomálií.....	10
2.3.1 Principal Component Analysis (PCA)	11
2.3.2 Local Outlier Factor (LOF).....	11
2.3.3 Subspace Outlier Detection (SOD)	11
2.3.4 Empirical-Cumulative-distribution-based Outlier Detection (ECOD)	12
2.3.5 Copula-Based Outlier Detection (COPOD)	12
2.3.6 Isolation Forest.....	12
2.3.7 Isolation-based Nearest Neighbour Ensemble (INNE)	12
2.3.8 Lightweight On-line Detector of Anomalies (LODA)	12
2.4 Nastavenie experimentov a porovnanie metód detekcie anomálií	13
2.4.1 Nastavenie experimentov	13
2.4.2 Vyhodnotenie úspešnosti detekcie	14
2.4.3 Porovnanie výpočtovej náročnosti.....	14
2.4.4 Interpretácia výsledkov	15
Zhrnutie	19
Bibliografia.....	21
Prílohy	22

Popis projektu

Projekt **Automatizácia digitálnej forenznej analýzy a odpovede na incident** (ďalej len „ADFIR“) je financovaný Európskou úniou – Next GenerationEU prostredníctvom Plánu obnovy a odolnosti Slovenskej republiky pod číslom projektu č. 09I05-03-V02-00079. Tento projekt sa zaoberá jednou z kľúčových výziev v oblasti kybernetickej bezpečnosti a informačnej bezpečnosti – ako spracovať obrovské množstvo digitálnych dôkazov, ktoré vznikajú počas incidentov kybernetickej bezpečnosti alebo forenznych vyšetrovaní. V súčasnosti je tento proces veľmi náročný z hľadiska ľudských zdrojov a času. Automatizácia pomocou metód strojového učenia môže preto výrazne **zlepšiť kvalitu digitálnej forenznej analýzy** a skrátiť čas potrebný na jej vykonanie. Celkovo to umožňuje bezpečnostným tímom efektívnejšie reagovať na kybernetické hrozby. Hlavné prínosy tohto projektu sú:

- **Rýchlejšie riešenie incidentov v oblasti kybernetickej bezpečnosti.** Projekt ADFIR zavádza automatizované prístupy k zberu, spracovaniu a analýze digitálnych stôp. Vďaka tomu môžu bezpečnostné tímy rýchlejšie identifikovať príčiny incidentov a prijať účinné opatrenia na ich riešenie.
- **Zníženie pracovnej záťaže forenznych analytikov.** Rutinné a časovo náročné úlohy spojené so spracovaním digitálnych stôp budú nahradené automatizovanými metódami. To umožní analytikom sústrediť sa na zložitejšie prípady a strategické rozhodovanie.
- **Vyššia kvalita a konzistentnosť výstupov.** Použitie jednotných metodík a nástrojov zaručuje, že spracované digitálne stopy budú presnejšie, konzistentnejšie a ľahšie overiteľné. To výrazne znižuje riziko chýb spôsobených ľudskými faktormi.
- **Možné využitie v trestnom konaní.** Výstupy projektu budú vyvinuté v súlade s právnymi požiadavkami a normami, čo umožní, aby digitálne stopy boli akceptované ako relevantné dôkazy pre vyšetrovanie a súdne konania.

Úvod

V oblasti digitálnej forenznej analýzy predstavuje efektívne spracovanie veľkého množstva heterogénnych dát zásadnú výzvu, najmä pri rekonštrukcii udalostí a identifikácii kauzálnych súvislostí medzi jednotlivými artefaktmi. Moderné prístupy preto čoraz viac využívajú koncepty hľadania anomálií dát, ktoré umožňujú transformovať izolované forenzne stopy do analytikmi využiteľného modelu.

Tento dokument sa zameriava na návrh a implementáciu metódy identifikácie forenznych stôp v rámci projektu ADFIR (Úloha KPB4.1 – Identifikácia vhodných metód na identifikáciu digitálnych dôkazov súvisiacich s analyzovaným prípadom). Cieľom je vytvoriť model, ktorý umožní systematické spracovanie artefaktov, redukciu množstva záznamov, ich koreláciu naprieč rôznymi zdrojmi a následnú interpretáciu v kontexte identifikovaných scenárov správania systému alebo používateľa.

Kľúčovou časťou práce je spracovanie dát a výber vhodných modelov strojového učenia na identifikáciu digitálnych stôp v digitálnej forenznej analýze. Tento prístup umožňuje vyhľadávať digitálne stopy pomocou artefaktov a zároveň znížiť množstvo jednotlivých záznamov, ktoré musí analytik spracovať.

Navrhovaný model tak poskytuje systematický spôsob, ako prejsť od fragmentovaných dát k integrovanému pohľadu na digitálne stopy, pričom podporuje nielen rekonštrukciu udalostí, ale aj identifikáciu vzorcov správania relevantných pre foreznú analýzu.

1 Teoretický úvod

V kontexte digitálnej forenznej analýzy predstavuje identifikácia digitálnych stôp komplexný proces, ktorého cieľom je odhalenie relevantných vzorov, vzťahov a odchýlok v rozsiahlych dátových množinách. Keďže charakter a štruktúra digitálnych stôp sa môžu výrazne líšiť v závislosti od zdroja dát a analyzovaného prostredia, neexistuje univerzálna metóda vhodná pre všetky typy analýz. Z tohto dôvodu sú v rámci tohto dokumentu predstavené viaceré metódy vhodné na identifikáciu digitálnych stôp z rôznych perspektív.

Významnú časť výstupu tvorí identifikácia anomálií. Pred samotnou detekciou je však nevyhnutné jednoznačne definovať pojem anomálie v kontexte analyzovaných dát. Správanie, ktoré sa javí ako neobvyklé z pohľadu globálneho datasetu, môže predstavovať legítimnú aktivitu konkrétného používateľa alebo systému, zatiaľ čo naopak bežné správanie môže v určitom kontexte indikovať bezpečnostný incident. Z tohto dôvodu nie je možné anomáliu definovať iba ako štatistickú odchýlku od bežného správania, ale je potrebné zohľadniť aj charakter analyzovaného prostredia a význam jednotlivých digitálnych stôp.

Keďže manuálne označenie všetkých dostupných dátových sád by bolo časovo aj prakticky nerealizovateľné, boli využité metódy detekcie anomálií typu unsupervised learning (učenie bez učiteľa), ktoré umožňujú identifikovať odchýlky bez potreby vopred definovaných tried. Tieto metódy sú vhodné najmä v prostredí digitálnej forenznej analýzy, kde dostupné dáta často obsahujú prevažne legítimne udalosti a iba malé množstvo potenciálne významných alebo škodlivých aktivít.

V rámci analýzy bolo skúmaných viacero prístupov k detekcii anomálií, ktoré reprezentujú rozdielne princípy identifikácie odchýlok v dátach. Hodnotenú boli lineárne modely založené na transformácii dát do vhodnejšieho priestoru, metódy založené na lokálnej podobnosti objektov, pravdepodobnostné modely využívajúce distribučné vlastnosti dát a ensemble metódy kombinujúce viacero detekčných prístupov. Cieľom bolo overiť schopnosť jednotlivých metód identifikovať neštandardné udalosti v komplexných a heterogénnych dátach a zároveň porovnať ich vhodnosť pre potreby digitálnej forenznej analýzy.

Z tohto dôvodu bolo kľúčové dôkladné predspracovanie dát, podrobne opísané vo výstupe ADFIR-D13-KPB2-02 Metóda na tvorbu datasetu, ktoré zabezpečilo zachovanie heterogenity, kvality a komplexnosti analyzovaných údajov. Predspracovanie zahŕňalo úpravu vstupných dát do formy vhodnej pre jednotlivé algoritmy, odstránenie nerelevantných informácií a vytvorenie reprezentácie umožňujúcej efektívnu analýzu jednotlivých vlastností digitálnych stôp.

Významnou súčasťou využitia metód strojového učenia v oblasti digitálnej forenznej analýzy nie je iba samotná schopnosť identifikovať odchýlky v dátach, ale aj možnosť interpretovať získané výsledky. V prostredí digitálnej forenznej analýzy predstavuje vysvetliteľnosť (explainability) jeden z kľúčových aspektov, keďže výsledky automatizovaných metód musia byť zrozumiteľné a overiteľné forezným analytikom. Model alebo metóda, ktorá identifikuje

podozrivé správanie bez poskytnutia informácie o dôvodoch svojho rozhodnutia, môže mať v procese vyšetrovania obmedzenú využiteľnosť. Z tohto dôvodu sa v rámci navrhnutého prístupu venujeme aj analýze vlastností a charakteristík, ktoré najvýznamnejšie prispievajú k označeniu udalostí ako potenciálne anomálnych.

1.1 Detekcia anomálií

Detekcia anomálií je kritickou súčasťou digitálnej forenznej analýzy. Jej hlavným cieľom je identifikovať vzorce, udalosti alebo záznamy v dátach, ktoré nezodpovedajú očakávanému (normálnemu) správaniu systému. Tieto neštandardné javy sa označujú ako anomálie alebo odľahlé hodnoty (outliers) a vo forenznom kontexte často reprezentujú stopy po kybernetickom incidente, neoprávnenom prístupe či manipulácii s dátami [1].

1.2 Detekcia anomálií bez učiteľa (Unsupervised)

V reálnych forenznych scenároch málokedy máme k dispozícii vopred označené (labeled) dáta, ktoré by presne definovali, čo je útok a čo normálna prevádzka. Preto sa primárne využíva **unsupervised anomaly detection** [2].

Tento prístup funguje bez akýchkoľvek predchádzajúcich vedomostí o anomáliách. Vychádza z dvoch základných predpokladov:

1. Drvivá väčšina dát v datasete reprezentuje normálne, legitímne správanie systému.
2. Anomálie sú štatisticky veľmi zriedkavé a od normálnych dát sa výrazne líšia svojimi vlastnosťami [1], [2].

Algoritmy bez učiteľa analyzujú vnútornú štruktúru dát a hľadajú skupiny s vysokou hustotou (normálny stav). Všetko, čo leží mimo týchto hustých skupín, prípadne prislúcha k príliš malým a izolovaným zhlukom, je označené za anomáliu [2].

1.3 Typy anomálií

Vo forenznej praxi a analýze digitálnych stôp rozlišujeme tri základné kategórie anomálií [1]:

1.3.1 Bodové anomálie (point anomalies)

Ide o najjednoduchší typ anomálie. Nastáva vtedy, keď sa **jedna konkrétna udalosť alebo dátový bod** zásadne líši od všetkých ostatných dát v celej dátovej sade.

- *Príklad:* Súbor s extrémne veľkou veľkosťou v priečinku, kde bývajú len malé textové logy, alebo jednorazový pokus o prenos obrovského objemu dát zo servera.

1.3.2 Kontextové anomálie (Contextual anomalies)

Dátový bod sa považuje za anomálny len **vzhľadom na špecifický kontext**, hoci v inom kontexte by bol úplne normálny. Pri identifikácii je kľúčové hodnotiť nielen samotnú hodnotu (behaviorálny atribút), ale aj kontext (napr. čas, poloha).

- *Príklad:* Prihlásenie používateľa do systému o 14:00 počas pracovného dňa je normálne. Prihlásenie rovnakého používateľa z rovnakej IP adresy, ale o 03:00 ráno na Štedrý deň, vykazuje silné kontextové anomálne črty.

1.3.3 Kolektívne anomálie (Collective anomalies)

Nastávajú vtedy, keď **sekvencia alebo skupina udalostí** spoločne vykazuje anomálne správanie, hoci každá jedna udalosť sama o sebe by bola braná ako úplne bežná. Anomália vzniká až ich vzájomným prepojením a štruktúrou v čase či priestore.

- *Príklad:* Jeden neúspešný pokus o prihlásenie je bežná chyba používateľa (nie je to anomália). Avšak 500 neúspešných pokusov o prihlásenie v priebehu 30 sekúnd z rôznych účtov tvorí kolektívnu anomáliu, ktorá jasne indikuje útok hrubou silou (brute-force).

1.4 Kategorizácia metód pre detekciu anomálií

Metódy unsupervised detekcie anomálií môžeme na základe ich matematického a algoritmického princípu rozdeliť do štyroch hlavných kategórií [3],[4]:

- **Lineárne modely pre detekciu odľahlých hodnôt (Linear Models):** Tieto prístupy predpokladajú, že normálne dáta ležia v nízkorozmernom lineárnom podpriestore. Anomálie sa identifikujú na základe ich vzdialenosti od tohto podpriestoru [5].
 - *Príklady:* Analýza hlavných komponentov (PCA), One-Class Support Vector Machines využívajúce stochastický gradientný zostup (SGDOCSVM).
- **Modely založené na blízkosti (Proximity-Based Models):** Tieto metódy analyzujú priestorové rozdelenie dát. Predpokladajú, že anomálie sa nachádzajú v izolovaných oblastiach s nízkou hustotou, ďaleko od svojich najbližších susedov.
 - *Príklady:* Local Outlier Factor (LOF) [6], Subspace Outlier Detection (SOD) [7], Rotation-based Outlier Detection (ROD) [8].
- **Probabilistické (pravdepodobnostné) modely:** Tieto prístupy modelujú anomálie na základe odhadu distribúcie hustoty pravdepodobnosti dát. Body s extrémne nízkou pravdepodobnosťou výskytu v danej distribúcii sú označené ako anomálie.
 - *Príklady:* Empirical-Cumulative-distribution-based Outlier Detection (ECOD) [9], Copula Based Outlier Detection (COPD) [10].

- **Združené modely a kombinované rámce (Outlier Ensembles):** Využívajú kombináciu viacerých modelov alebo techník (tzv. súbory/ensembles) na dosiahnutie robustnejších výsledkov a zníženie miery falošne pozitívnych detekcií. Často fungujú na princípe izolácie anomálií od zvyšku dát.
 - *Príklady:* Isolation Forest (Izolačný les), Isolation-based Anomaly Detection Using Nearest-Neighbor Ensembles (INNE) [11], Lightweight on-line detector of anomalies (LODA) [12].

2 Aplikácia vybraných metód

V tejto kapitole sú predstavené praktické kroky aplikácie vybraných analytických metód na pripravené forenzné datasety. Cieľom experimentálnej časti je overiť využiteľnosť metód popísaných v predchádzajúcich kapitolách pri identifikácii významných vzorov v digitálnych stopách, detekcii odchýlok a následnej interpretácii získaných výsledkov.

2.1 Detekcia anomálií pomocou metód strojového učenia bez učiteľa

Významnou súčasťou identifikácie digitálnych stôp v oblasti digitálnej forenznej analýzy je detekcia neobvyklého alebo podozrivého správania v rozsiahlych dátových množinách. Takéto správanie však nie je možné vždy jednoznačne definovať, keďže udalosť, ktorá sa javí ako anomálna z pohľadu celého datasetu, môže predstavovať legítimnú aktivitu konkrétneho systému alebo používateľa. Z tohto dôvodu bola využitá kategória metód učenia bez učiteľa (unsupervised learning), ktoré umožňujú identifikovať odľahlé objekty bez potreby manuálne označených dát.

Detekcia anomálií (outlier detection) predstavuje proces identifikácie pozorovaní, ktoré sa významne odlišujú od ostatných objektov v analyzovanom priestore. Na rozdiel od klasifikačných metód, kde je potrebné vopred definovať jednotlivé triedy, metódy detekcie anomálií pracujú iba s charakteristikami samotných dát a hľadajú objekty s neštandardným správaním.

V rámci tejto práce boli skúmané viaceré prístupy k detekcii anomálií, ktoré je možné rozdeliť do štyroch hlavných kategórií:

- **lineárne metódy detekcie anomálií,**
- **metódy založené na podobnosti a vzdialenosti objektov,**
- **pravdepodobnostné metódy,**
- **ensemble metódy založené na kombinácii viacerých detektorov.**

2.2 Dátové súbory a predspracovanie

Pre experimenty s detekciou anomálií boli použité rovnaké dátové súbory ako pri metóde hierarchického zhľukovania. Analyzované datasety pozostávali z digitálnych stôp získaných z operačného systému Windows, ktoré boli transformované do štruktúrovanej reprezentácie pomocou procesu predspracovania dát opísaného v predchádzajúcej kapitole, resp. vo výstupe (ADFIR-D13-KPB2-02-Metóda na tvorbu datasetu (SK)).

Základnou jednotkou analýzy bolo časové okno reprezentujúce stav systému počas definovaného časového intervalu. Jednotlivé udalosti operačného systému boli agregované

pomocou vybraných agregáčnych funkcií, pričom výsledkom bola matica atribútov opisujúca charakteristické vlastnosti jednotlivých časových okien.

Každé časové okno bolo reprezentované množinou atribútov získaných z forenzných artefaktov extrahovaných pomocou nástroja Plaso. Atribúty obsahovali informácie o rôznych typoch systémovej aktivity, napríklad udalostiach súborového systému, systémových registroch, používateľských aktivitách alebo vykonaných operáciách.

Keďže dostupné datasety neobsahovali kompletne manuálne označenie všetkých anomálnych udalostí, nebolo možné použiť klasické metódy učenia s učiteľom. Z tohto dôvodu boli zvolené výhradne metódy učenia bez učiteľa, ktoré umožňujú identifikovať potenciálne podozrivé objekty priamo na základe charakteristík dát.

2.3 Použité metódy detekcie anomálií

Na implementáciu jednotlivých algoritmov boli využité knižnice **scikit-learn** [13] a **PyOD** [3]. Knižnica scikit-learn poskytuje implementácie viacerých štandardných metód strojového učenia vrátane redukcie dimenzionality a detekcie odľahlých hodnôt. PyOD (*Python Outlier Detection*) je špecializovaná knižnica určená na detekciu anomálií v multidimenzionálnych dátach a obsahuje široké spektrum dostupných detektorov.

V rámci experimentov boli testované nasledovné metódy:

Metóda	Kategória	Hlavný princíp
PCA	Lineárne metódy	Identifikácia odchýlok v priestore hlavných komponentov
LOF	Proximity-based metódy	Porovnanie lokálnej hustoty objektu s jeho okolím
ECOD	Pravdepodobnostné metódy	Analýza empirických distribučných funkcií
COPOD	Pravdepodobnostné metódy	Detekcia na základe kopúlovej distribúcie
Isolation Forest	Ensemble metódy	Izolácia objektov pomocou náhodného rozdelenia priestoru
INNE	Ensemble metódy	Detekcia založená na izolácii pomocou najbližších susedov
LODA	Ensemble metódy	Online detekcia pomocou kombinácie jednoduchých histogramových modelov

Tabuľka 1: Metódy detekcie anomálií

Výber metód bol realizovaný s cieľom pokryť hlavné kategórie unsupervised detekcie anomálií. Do experimentov boli zaradené lineárne, prístupy založené na hustote, pravdepodobnostné aj ensemble prístupy, aby bolo možné porovnať ich správanie na rovnakých forenzných datasetoch a identifikovať vhodné vlastnosti jednotlivých kategórií pre digitálnu forenznú analýzu.

2.3.1 Principal Component Analysis (PCA)

Metóda Principal Component Analysis (PCA) predstavuje lineárnu transformačnú metódu, ktorá redukuje pôvodný priestor atribútov do menšieho počtu hlavných komponentov. Tieto komponenty zachovávajú najvýznamnejšiu variabilitu pôvodných dát.

Pri detekcii anomálií PCA predpokladá, že bežné objekty sa nachádzajú v priestore definovanom hlavnými komponentmi, zatiaľ čo anomálne objekty vykazujú väčšiu odchýlku od tejto reprezentácie. Veľkosť tejto odchýlky môže byť následne využitá ako skóre anomálie.

Výhodou PCA je nízka výpočtová náročnosť a vhodnosť pre vysokodimenzionálne dáta. Nevýhodou je predpoklad lineárnych vzťahov medzi atribútmi, čo nemusí byť vždy splnené pri komplexných forenzných dátach.

2.3.2 Local Outlier Factor (LOF)

Local Outlier Factor (LOF) je neparametrická metóda založená na analýze lokálnej hustoty objektov. Každý objekt je porovnávaný so svojím okolím definovaným množinou najbližších susedov.

Anomálne objekty majú typicky výrazne nižšiu lokálnu hustotu ako ich susedia, čo vedie k vyššej hodnote LOF skóre. Metóda je vhodná najmä v prípadoch, keď sa normálne objekty nachádzajú v rôznych hustotných oblastiach.

Pri experimentoch boli sledované parametre:

- počet susedov,
- použitá metrika vzdialenosti,
- predpokladaný podiel anomálií (*contamination*).

2.3.3 Subspace Outlier Detection (SOD)

Metóda Subspace Outlier Detection identifikuje anomálie v podpriestoroch dát. Na rozdiel od metód pracujúcich s kompletným priestorom atribútov predpokladá, že anomália nemusí byť viditeľná vo všetkých dimenziách, ale iba v určitej podmnožine atribútov.

Tento prístup je vhodný najmä pre vysokodimenzionálne datasety, kde môže byť významná iba časť dostupných charakteristík.

2.3.4 Empirical-Cumulative-distribution-based Outlier Detection (ECOD)

ECOD je bezparametrická metóda založená na empirických kumulatívnych distribučných funkciách. Metóda nevyžaduje učenie modelu ani zložité nastavovanie parametrov, ale hodnotí, nakoľko je výskyt určitej hodnoty atribútu pravdepodobný vzhľadom na celé rozdelenie dát.

Výhodou ECOD je vysoká interpretovateľnosť, keďže výsledné skóre priamo súvisí s odchýlkou jednotlivých atribútov.

2.3.5 Copula-Based Outlier Detection (COPOD)

COPOD je deterministická metóda využívajúca štatistické vlastnosti kopúl a empirických distribučných funkcií. Podobne ako ECOD nevyžaduje fázu učenia ani optimalizáciu hyperparametrov.

Výhodou tejto metódy je stabilita výsledkov a nízka výpočtová náročnosť, čo umožňuje jej použitie aj pri rozsiahlych dátových množinách.

2.3.6 Isolation Forest

Isolation Forest je jedna z najpoužívanějších metód detekcie anomálií vo vysokodimenzionálnych dátach. Je založená na princípe, že anomálne objekty je možné izolovať jednoduchšie ako bežné objekty.

Metóda vytvára množstvo náhodných rozhodovacích stromov, pričom objekty, ktoré sú izolované pomocou kratších ciest v stromoch, získavajú vyššie skóre anomálie.

2.3.7 Isolation-based Nearest Neighbour Ensemble (INNE)

INNE rozširuje princíp izolácie objektov využitím informácie o najbližších susedoch. V porovnaní s tradičnými metódami založenými na vzdialenosti je vhodnejšia pre veľké a vysokodimenzionálne datasety, keďže dosahuje nižšiu výpočtovú náročnosť.

2.3.8 Lightweight On-line Detector of Anomalies (LODA)

LODA predstavuje ensemble metódu založenú na kombinácii viacerých jednoduchých histogramových modelov. Je navrhnutá najmä pre rýchlu analýzu veľkých dátových tokov a dokáže pracovať aj s neúplnými dátami.

Výhodou metódy je nízka pamäťová a časová náročnosť, čo ju robí vhodnou pre rozsiahle analytické úlohy.

2.4 Nastavenie experimentov a porovnanie metód detekcie anomálií

Na základe uvedených kategórií boli jednotlivé metódy ďalej porovnávané z hľadiska ich parametrov, výpočtovej náročnosti a dosiahnutých výsledkov. Kompletný prehľad experimentálne testovaných nastavení je uvedený v prílohe A, pričom hlavné charakteristiky jednotlivých metód sú zhrnuté v samostatnej tabuľke.

2.4.1 Nastavenie experimentov

Na overenie navrhnutého prístupu boli realizované experimenty na pripravených forenzných datasetoch obsahujúcich digitálne stopy získané z operačného systému Windows. Jednotlivé algoritmy boli testované s rôznymi konfiguráciami hyperparametrov, pričom osobitná pozornosť bola venovaná parametru *contamination*, ktorý reprezentuje predpokladaný podiel anomálnych objektov v datasete. Pre metódy založené na lokálnej podobnosti bola sledovaná aj citlivosť na počet susedov a použitú metriku vzdialenosti. Pri ensemble metódach bol analyzovaný vplyv počtu vytvorených modelov na stabilitu výsledkov.

Cieľom experimentov bolo identifikovať nastavenia, ktoré umožňujú efektívne odlišovať potenciálne podozrivé udalosti od bežného správania systému a zároveň zachovať primeranú výpočtovú náročnosť spracovania.

Jedným zo sledovaných parametrov bol **contamination**, ktorý predstavuje predpokladaný podiel anomálnych objektov v analyzovanom datasete. Tento parameter je pri detekcii anomálií významný najmä v oblasti digitálnej forenznej analýzy, kde bezpečnostne relevantné udalosti tvoria iba malú časť všetkých dostupných záznamov. Nesprávne nastavenie tohto parametra môže viesť buď k nadmernému označeniu legitímnych udalostí ako anomálnych, alebo naopak k prehliadnutiu podozrivých aktivít.

Pri jednotlivých algoritmoch boli okrem spoločného parametra *contamination* sledované aj špecifické hyperparametre ovplyvňujúce správanie konkrétnej metódy. Pri metóde **Local Outlier Factor** (LOF) bol analyzovaný najmä počet susedov - **n_neighbors** a použitá **metrika vzdialenosti**, ktoré určujú spôsob hodnotenia lokálneho okolia objektov. Pri metódach založených na izolácii, ako **Isolation Forest** a **INNE**, bol sledovaný počet vytvorených modelov **n_estimators**, ktorý ovplyvňuje stabilitu výsledného skóre anomálie. Pri metódach pracujúcich s podpriestormi, napríklad **Subspace Outlier Detection** (SOD), boli hodnotené parametre určujúce výber referenčných objektov a citlivosť detekcie v jednotlivých dimenziách.

2.4.2 Vyhodnotenie úspešnosti detekcie

Úspešnosť jednotlivých metód bola vyhodnocovaná porovnaním identifikovaných anomálnych objektov s referenčnými udalosťami získanými manuálnou analýzou datasetov. Vzhľadom na charakter digitálnych forenzných dát nebolo cieľom vytvoriť klasifikátor bezpečnostných incidentov, ale identifikovať časové okná a udalosti vykazujúce neštandardné správanie.

Výsledky ukázali, že jednotlivé algoritmy dosahovali rozdielnu úspešnosť v závislosti od charakteru dát. Kým niektoré metódy identifikovali väčší počet potenciálnych anomálií, iné poskytovali konzervatívnejšie výsledky s nižším počtom falošne pozitívnych označení.

Súčasťou porovnania bola aj časová zložitosť algoritmov, ktorá poskytuje teoretický odhad ich škálovateľnosti vzhľadom na veľkosť vstupných dát. V tabuľke je časová zložitosť vyjadrená pomocou *O-notácie*, pričom premenná n označuje počet analyzovaných objektov a d počet atribútov dátovej reprezentácie. Tento ukazovateľ je významný najmä pri forenzných datasetoch, ktoré môžu obsahovať veľké množstvo udalostí a vyžadujú metódy schopné spracovania rozsiahlych dát v prijateľnom čase.

Okrem teoretickej výpočtovej náročnosti bol zaznamenaný aj reálny čas behu jednotlivých algoritmov pri spracovaní analyzovaného datasetu. Namerané hodnoty umožňujú porovnať praktickú využiteľnosť jednotlivých metód, keďže algoritmus s nízkou teoretickou zložitou nemusí byť vždy najrýchlejší v konkrétnej implementácii. Výsledný čas ovplyvňuje nielen samotný algoritmus, ale aj použitá knižnica, optimalizácie implementácie, veľkosť dát a počet testovaných parametrov.

2.4.3 Porovnanie výpočtovej náročnosti

Významným faktorom pri výbere metódy detekcie anomálií je aj čas potrebný na spracovanie rozsiahlych dátových množín. Z tohto dôvodu bola sledovaná teoretická časová zložitosť algoritmov aj reálny čas behu experimentov.

Pravdepodobnostné metódy ECOD a COPOD vykazovali nízku výpočtovú náročnosť a stabilné správanie bez potreby rozsiahlej optimalizácie parametrov. Naopak, metódy založené na lokálnej podobnosti objektov alebo vytváraní ensemble modelov vyžadovali vyššie výpočtové zdroje, avšak dokázali zachytiť aj komplexnejšie formy odchýlok v dátach.

Úspešnosť jednotlivých metód bola hodnotená na základe počtu správne identifikovaných anomálnych objektov v porovnaní s referenčnými hodnotami získanými manuálnym označením. Vzhľadom na charakter analyzovaných dát nebolo možné jednoznačne označiť všetky potenciálne anomálie, preto boli výsledky interpretované ako schopnosť jednotlivých

metód nájst' známu množinu podozrivých objektov. V tabuľke je preto uvedený počet korektne identifikovaných outlierov a priemerný počet správne detegovaných objektov pri opakovaných experimentoch s rôznymi konfiguráciami parametrov.

2.4.4 Interpretácia výsledkov

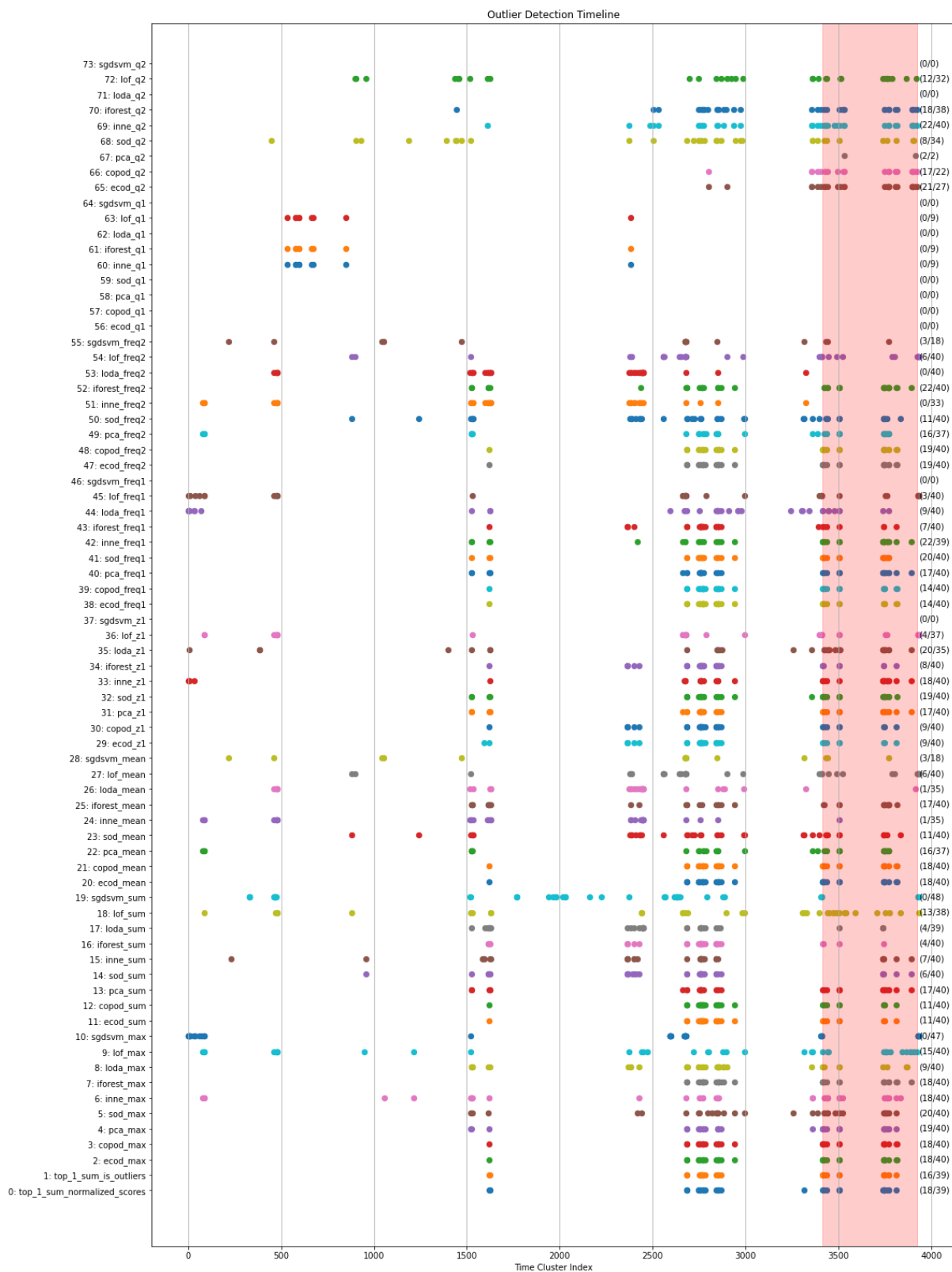
Detegované anomálie predstavovali udalosti alebo časové úseky, ktoré sa svojimi charakteristikami odlišovali od bežného správania systému. Samotná identifikácia anomálie však automaticky neznamena prítomnosť bezpečnostného incidentu. Časť označených objektov zodpovedala legitímnym, avšak menej častým aktivitám používateľov alebo operačného systému.

Výsledky preto slúžia predovšetkým ako podporný mechanizmus pre digitálneho forenzného analytika, ktorý môže následne zamerať pozornosť na menšiu podmnožinu potenciálne významných udalostí. Kombinácia viacerých metód detekcie anomálií sa ukázala ako vhodný prístup, keďže jednotlivé algoritmy identifikovali odlišné typy anomálneho správania a poskytovali komplementárne výsledky.

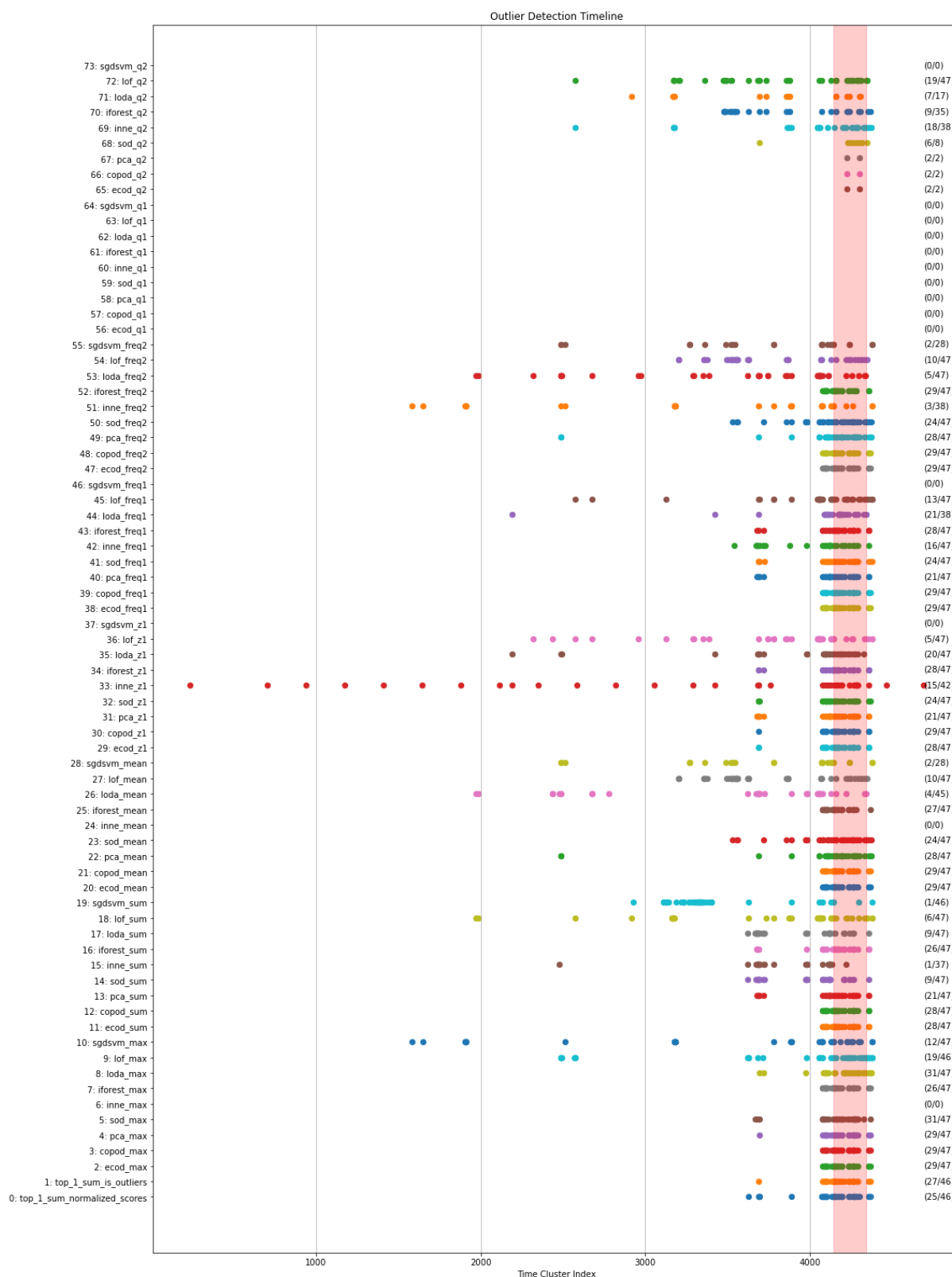
Pri interpretácii výsledkov je potrebné zohľadniť špecifiká digitálnej foreznej analýzy. Detekcia anomálie nepredstavuje automatické potvrdenie škodlivej aktivity, ale označenie udalosti, ktorá sa svojimi charakteristikami odlišuje od ostatného správania v datasete. Výsledky jednotlivých algoritmov preto slúžia najmä ako podpora pre forenzného analytika pri redukcii množstva dát a identifikácii oblastí vyžadujúcich podrobnejšie manuálne preskúmanie.

Výsledky experimentov ukázali, že jednotlivé algoritmy identifikujú odlišné typy anomálneho správania. Metódy založené na lokálnej hustote objektov boli citlivejšie na lokálne odchýlky, zatiaľ čo pravdepodobnostné a ensemble prístupy vykazovali stabilnejšie výsledky pri spracovaní rozsiahlych datasetov. Z pohľadu digitálnej foreznej analýzy sa preto javí ako výhodné kombinovať viacero detekčných prístupov, ktoré sa vzájomne dopĺňajú a umožňujú zachytiť širšie spektrum potenciálne významných udalostí.

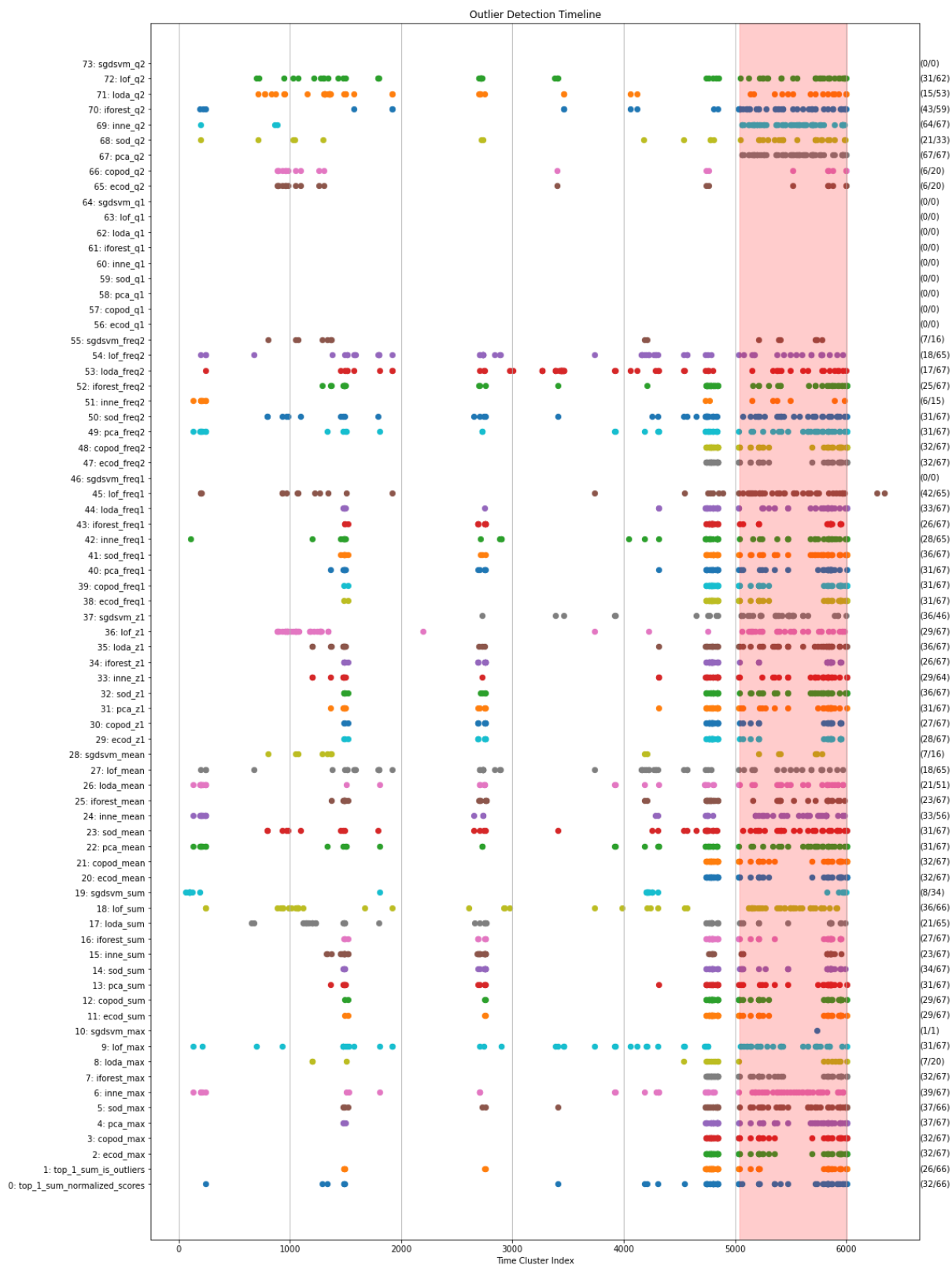
Časovú súvislosť medzi detegovanými anomáliami a bezpečnostným incidentom ilustrujú **Obrázok 1, Obrázok 2, Obrázok 3**. Vizualizácia zobrazuje anomálie identifikované jednotlivými použitými metódami, pričom červený pás označuje časové obdobie bezpečnostného incidentu. Z obrázka je zrejmé, že významná časť detegovaných anomálií sa časovo prekrýva s obdobím incidentu. Tento výsledok naznačuje, že použité metódy dokázali identifikovať časové úseky obsahujúce správanie súvisiace s incidentom, pričom zároveň poskytli foreznému analytikovi užšiu množinu udalostí vhodných na ďalšie preskúmanie.



Obrázok 1: Výsledky metód pre Stolen Szechuan Sauce DC



Obrázok 2: Výsledky metód pre Stolen Szechuan Sauce Desktop



Obrázok 3: Výsledky metód pre Nist Data Leakage Case

Zhrnutie

V rámci tohto dokumentu bol navrhnutý a experimentálne overený prístup k identifikácii potenciálne významných digitálnych stôp pomocou metód detekcie anomálií založených na učení bez učiteľa (unsupervised learning). Hlavným cieľom bolo preskúmať možnosti automatizovaného vyhľadávania neštandardného správania v rozsiahlych forenzných datasetoch, v ktorých nie je dostupné kompletne manuálne označenie všetkých bezpečnostne relevantných udalostí.

V prvej fáze bola vykonaná transformácia heterogénnych digitálnych stôp získaných z operačného systému Windows do reprezentácie vhodnej pre algoritmickú analýzu. Keďže jednotlivé forenzné udalosti samostatne často neposkytujú dostatočný kontext na rozhodnutie o ich významnosti, dáta boli agregované do časových okien reprezentujúcich správanie systému počas definovaných intervalov. Takýto prístup umožnil analyzovať nielen jednotlivé udalosti, ale aj širšie vzory správania a vzájomné súvislosti medzi jednotlivými atribútmi.

V rámci experimentov bolo porovnaných viacero kategórií metód detekcie anomálií reprezentujúcich rozdielne analytické princípy. Testované boli lineárne metódy založené na transformácii dát, metódy využívajúce lokálnu podobnosť objektov, pravdepodobnostné prístupy založené na distribučných vlastnostiach dát a ensemble metódy využívajúce kombináciu viacerých detekčných modelov. Takéto porovnanie umožnilo posúdiť rozdiely medzi jednotlivými prístupmi z pohľadu presnosti detekcie, výpočtovej náročnosti a praktickej využiteľnosti pri analýze forenzných dát.

Výsledky experimentov ukázali, že jednotlivé metódy majú rozdielne vlastnosti a ich vhodnosť závisí od charakteru analyzovaných dát. Jednoduchšie a výpočtovo menej náročné metódy umožňujú rýchlu prvotnú analýzu rozsiahlych datasetov, zatiaľ čo komplexnejšie metódy dokážu zachytiť zložitejšie vzory správania. Z tohto dôvodu nie je možné určiť univerzálne najlepšiu metódu pre všetky forenzné scenáre, ale vhodnejším prístupom je kombinácia viacerých analytických techník podľa požiadaviek konkrétneho vyšetrovania.

Z pohľadu digitálnej forenznej analýzy predstavuje detekcia anomálií významný nástroj najmä pri práci s rozsiahlymi časovými osami a veľkým množstvom systémových artefaktov. Výstupom algoritmov nie je priame potvrdenie bezpečnostného incidentu, ale identifikácia udalostí alebo časových úsekov, ktoré sa odlišujú od očakávaného správania a zaslúžia si ďalšiu analýzu. Takýto prístup umožňuje foreznému analytikovi zamerať pozornosť na najrelevantnejšie časti dát, čím dochádza k výraznej redukcii manuálne spracovávaného objemu informácií.

Navrhnutý prístup preto predstavuje podporný analytický nástroj, ktorý kombinuje možnosti moderných metód strojového učenia s požiadavkami digitálnej forenznej praxe. Jeho hlavný prínos spočíva v schopnosti automatizovane vyhľadávať potenciálne významné odchýlky

v komplexných dátach, čím podporuje efektívnejšiu rekonštrukciu udalostí, identifikáciu podozrivého správania a následné rozhodovanie forenzného analytika.

Bibliografia

- [1] CHANDOLA, Varun, Arindam BANERJEE a Vipin KUMAR. Anomaly Detection: A Survey. *ACM Computing Surveys*. 2009, 41(3), 15:1–15:58. ISSN 0360-0300. DOI: 10.1145/1541880.1541882
- [2] AGGARWAL, Charu C. *Outlier Analysis*. 2nd ed. Cham: Springer, 2017. ISBN 978-3-319-47577-3. DOI: 10.1007/978-3-319-47577-6
- [3] ZHAO, Yue, Zainab NASRULLAH a Zheng LI. PyOD: A Python Toolbox for Scalable Outlier Detection. *Journal of Machine Learning Research*. 2019, 20(96), 1–7. ISSN 1533-7928.
- [4] S.J. Stolfo, F. Apap, E. Eskin, K. Heller, S. Hershkop, A. Honig, K. Svore, A comparative evaluation of two algorithms for windows registry anomaly detection. *Journal of Computer Security* 13(4), 659–693 (2005). <https://doi.org/10.3233/JCS-2005-13403>
- [5] Shyu, Mei-Ling & Chen, Shu-Ching & Sarinnapakorn, Kanoksri & Chang, Liwu. (2003). A Novel Anomaly Detection Scheme Based on Principal Component Classifier. *Proceedings of International Conference on Data Mining*.
- [6] Markus M. Breunig, Hans-Peter Kriegel, Raymond T. Ng, and Jörg Sander. 2000. LOF: identifying density-based local outliers. *SIGMOD Rec.* 29, 2 (June 2000), 93–104. <https://doi.org/10.1145/335191.335388>
- [7] KRIEGEL, Hans-Peter, Peer KRÖGER, Erich SCHUBERT a Arthur ZIMEK. Outlier Detection in Arbitrarily Oriented Subspaces. In: *IEEE 12th International Conference on Data Mining*. 2012, s. 379–388. DOI: 10.1109/ICDM.2012.140.
- [8] Y. Almardeny, N. Boujnah and F. Cleary, "A Novel Outlier Detection Method for Multivariate Data," in *IEEE Transactions on Knowledge and Data Engineering*, vol. 34, no. 9, pp. 4052-4062, 1 Sept. 2022, doi: 10.1109/TKDE.2020.3036524.
- [9] Li, Zheng & Zhao, Yue & Hu, Xiyang & Botta, Nicola & Ionescu, Cezar & Chen, George. (2023). ECOD: Unsupervised Outlier Detection Using Empirical Cumulative Distribution Functions. *IEEE Transactions on Knowledge and Data Engineering*. 35. 12181-12193. 10.1109/TKDE.2022.3159580.
- [10] Z. Li, Y. Zhao, N. Botta, C. Ionescu and X. Hu, "COPOD: Copula-Based Outlier Detection," *2020 IEEE International Conference on Data Mining (ICDM)*, Sorrento, Italy, 2020, pp. 1118-1123, doi: 10.1109/ICDM50108.2020.00135.
- [11] Bandaragoda TR, Ting KM, Albrecht D, Liu FT, Zhu Y, Wells JR. *Isolation-based anomaly detection using nearest-neighbor ensembles. Computational Intelligence*. 2018; 34: 968–998. <https://doi.org/10.1111/coin.12156>
- [12] Pevný, T. Loda: Lightweight on-line detector of anomalies. *Mach Learn* **102**, 275–304 (2016). <https://doi.org/10.1007/s10994-015-5521-0>
- [13] F. Pedregosa *et al.*, "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011

Prílohy

ADFIR-D19-KPB4-01-02-Appendix A.xlsx