D16 – Model na extrakciu digitálnych stôp do grafovej reprezentácie

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

Osnova

1 Popis projektu

Projekt **Automatizácia digitálnej forenzej 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. 09-I05-03-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 forenzných 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:

- **Urýchlené riešenie incidentov v oblasti kybernetickej bezpečnosti**. Projekt ADFIR zavádza automatizované prístupy k zberu, spracovaniu a analýze digitálnych stôp. V dôsledku toho môžu bezpečnostné tímy rýchlejšie identifikovať príčiny incidentov a prijať účinné opatrenia na ich riešenie.
- **Zníženie pracovnej záťaže forenzných 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 boli digitálne stopy akceptované ako relevantné dôkazy pre vyšetrovanie a súdne konania.

2 Úvod

S rastúcim objemom digitálnych dát a zvyšujúcou sa komplexnosťou kybernetických incidentov narastá aj potreba efektívnych metód na identifikáciu, extrakciu a interpretáciu digitálnych stôp. Tradičné prístupy k forenznej analýze často pracujú s veľkým množstvom heterogénnych artefaktov, ako sú logy, metadáta, súborové štruktúry, sieťové záznamy či pamäťové výpisy, pričom ich vzájomné vzťahy nemusia byť na prvý pohľad zrejmé. Grafová reprezentácia predstavuje vhodný spôsob modelovania týchto entít a ich väzieb, pretože umožňuje zachytiť nielen samotné objekty forenznej analýzy, ale aj ich kontext, časové súvislosti a interakcie medzi nimi.

V prostredí digitálnej forenznej analýzy je práve vzťahovosť údajov jedným z kľúčových predpokladov úspešnej rekonštrukcie udalostí. Grafové modely umožňujú reprezentovať digitálne stopy ako množinu uzlov a hrán, kde uzly predstavujú napríklad zariadenia, používateľov, procesy, súbory alebo komunikačné entity a hrany vyjadrujú ich vzájomné väzby, interakcie alebo kauzálne nadväznosti. Takáto forma reprezentácie podporuje nielen vizualizáciu dôkazného materiálu, ale aj aplikáciu analytických metód založených na teórii grafov, automatizovanom uvažovaní a identifikácii anomálií v skúmanom prostredí.

Model na extrakciu digitálnych stôp do grafovej reprezentácie preto možno chápať ako prostriedok na systematickú transformáciu forenzných artefaktov z nesúrodých dátových zdrojov do jednotnej štruktúrovanej podoby. Takto vytvorený model môže výrazne prispieť k lepšiemu pochopeniu priebehu incidentu, k efektívnejšiemu vyhľadávaniu relevantných súvislostí a k presvedčivejšej prezentácii výsledkov vyšetrovania.

3 Zdroje digitálnych stôp

V digitálnej forenznej analýze sa často stretávame s dvoma prístupmi: **offline** (postmortem) a **online** analýzou. Offline analýza prebieha až po incidente, keď je systém vypnutý alebo izolovaný, a odborník skúma najmä uložené dáta, diskové obrazy, logy a ďalšie trvalé artefakty. Online forenzná analýza sa naopak robí na bežiacom systéme počas incidentu, aby bolo možné zachytiť aj volatilné údaje, napríklad obsah pamäte RAM, aktívne procesy alebo otvorené sieťové spojenia.

Rozdiel medzi nimi je najmä v tom, **čo sa dá zachytiť** a **aké riziko hrozí integrite digitálnych stôp**. Pri postmortem prístupe je výhodou stabilné prostredie a menšie riziko zmeny dát, preto sa dobre hodí na detailné vyšetrovanie po incidente. Online analýza je zase nenahraditeľná v situáciách, keď môžu dôležité dôkazy zmiznúť po vypnutí zariadenia alebo po reštarte systému.

3.1 Databázy zo simulovaných útokov a techník útočníkov

Jedným z prístupov k vytváraniu dátových súborov je **cielená simulácia útokov a techník útočníkov** v kontrolovanom a izolovanom prostredí. V tomto prípade sa útoky vykonávajú vedome a systematicky. Útoky sa často vykonávajú na základe známych rámcov, ako je MITRE ATT&CK, s cieľom generovať forenzné artefakty zodpovedajúce konkrétnym technikám a fázam útoku.

Hlavnou výhodou tohto prístupu je **kontrola nad scenárom**. Výskumník alebo vývojár presne vie, ktoré techniky boli použité, v akom poradí a na aký účel. To umožňuje presné označovanie údajov, vytváranie vyvážených dátových súborov a systematické testovanie schopnosti modelov detegovať konkrétne typy správania. Simulované dátové súbory sú vhodné aj na vytváranie referenčných údajov pre experimentálne porovnania a validačné štúdie.

Na druhej strane majú simulované útoky aj svoje obmedzenia. Aj pri vysokej úrovni odbornosti môže byť simulácia **zjednodušená** a nemusí zachytiť všetky nepredvídateľné aspekty skutočných incidentov, ako sú chyby útočníkov, kombinované útoky viacerých aktérov alebo dlhodobé kampane s nízkou intenzitou. Simulované útoky však predstavujú dôležitý kompromis medzi realizmom a kontrolovateľnosťou údajov.

3.2 Databázy zo súťaží CTF

Ďalšia kategória pozostáva z **dátových súborov zo súťaží CTF (*Capture the Flag*)**, napr. nami použité [7-10]**)**, ktoré sa už dlho používajú vo vzdelávaní a odbornej príprave v oblasti kybernetickej bezpečnosti. Úlohy CTF sú zvyčajne navrhnuté tak, aby simulovali konkrétny bezpečnostný incident alebo jeho časť, a účastníci odpovedajú na vopred definované otázky analýzou digitálnych stôp.

Používanie dátových súborov CTF má určité **vnútorné obmedzenia**. Scenáre CTF sú často umelo vytvorené, zamerané na konkrétny typ útoku a zvyčajne ich vykonáva jediný útočník. Takýto model nemusí plne odrážať realitu, kde môže byť zapojených viacero útočníkov, môžu sa odohrávať viacfázové útoky a výsledok vyšetrovania nie je vopred známy. Okrem toho údaje CTF implicitne predpokladajú existenciu riešenia, zatiaľ čo bezpečnostné incidenty v reálnom svete sa vyznačujú vysokou mierou neistoty a nejednoznačnosti.

Napriek týmto obmedzeniam majú súťaže CTF **významné paralely so skutočnými bezpečnostnými incidentmi**. Aj v scenároch CTF útočníci používajú špecializované nástroje, techniky a taktiky, ktoré sú často identické alebo veľmi podobné tým, ktoré sa používajú v praxi. Proces skúmania údajov a digitálnych stôp, korelácie artefaktov a rekonštrukcie udalostí je porovnateľný so skutočnou digitálnou forenznou analýzou.

Hoci sú digitálne stopy v úlohách CTF umelo vytvorené, samotný **proces analýzy údajov a extrakcie forenzných artefaktov** je do veľkej miery rovnaký ako pri vyšetrovaní reálnych incidentov. Okrem toho štruktúrovaná povaha úloh CTF, často založená na otázkach a odpovediach, môže podporovať systematický analytický prístup, ktorý je prenosný do reálneho sveta a môže viesť k rýchlejšej identifikácii a riešeniu problémov.

3.3 Sémantická analýza

Sémantická analýza je proces spracovania vstupných dát do ľahko interpretovateľnej časovej osi ďalej použiteľná na zostrojenie grafovej reprezentácie. Pozostáva primárne z dvoch krokov:

1. predspracovanie a
2. označovanie vstupných dát.

Ako vstup do prvého kroku používame štandardne vytvorenú Plaso super časovú os [5], z ktorej je následne vybratých 6 kľúčových atribútov - *datetime*, *sourcetype*, *type*, *MACB*, *user*, *host* - ktoré sú použité v surovej "nespracovanej" forme a atribút *desc*, ktorý nesie najdôležitejšiu časť samotnej udalosti a orezáva sa na 200 znakov. Následne sa takto predspracovaná časová os rozdeľuje do okien po 400 riadkoch a pokračuje do 2. kroku - označovania dát.

Označovanie dát je vykonávané pomocou inferencie veľkého jazykového modelu Gemma 4[1] s 27 miliardami parametrov lokálne na koncovom zariadení nad predspracovanou časovou osou. Samotné výstupy sú následne ošetrené vzhľadom ku korektnému tvaru dát (ako pre grafické zobrazenie, tak aj ďalšiu analýzu). Každý riadok výstupu je potenciálne označiteľný ako podozrivý, riadky takto označené sú uložené a špecificky dooznačené následnými atribútmi - *attack_type*, *mitre_tactics*, *mitre_techniques*, *killchain_stage* a sprístupnené na export ako vo formáte JSON, tak CSV.

4 Graf ako nástroj na modelovanie forenzných dát

4.1 Teória grafov

Teória grafov je oblasť diskrétnej matematiky, ktorá sa zaoberá štúdiom grafov, ich vlastností a vzťahov medzi ich prvkami. Grafy slúžia ako veľmi užitočný nástroj na modelovanie rôznych situácií z praxe, napríklad dopravných sietí, počítačových sietí, sociálnych vzťahov alebo rôznych typov prepojení medzi objektmi. V nasledujúcom texte uvádzame základné pojmy a vlastnosti z tejto oblasti (viac informácii v [1]).

Graf sa zvyčajne označuje ako *G=(V, E)*, kde *V* je množina **vrcholov** a *E* je množina **hrán**. Vrcholy predstavujú jednotlivé objekty alebo body a hrany znázorňujú spojenia medzi nimi. Ak sú hrany **neorientované**, spojenie medzi dvoma vrcholmi nemá smer (hrana je obojsmerná). V prípade **orientovaného** grafu (digrafu) má každá hrana určený smer od jedného vrcholu k druhému.

[1] Model Gemma 4, popis modelu je dostupný na webovej stránke: https://deepmind.google/models/gemma/gemma-4/

Okrem toho sa môžeme stretnúť aj s **ohodnotenými** grafmi, v ktorých majú hrany alebo vrcholy priradenú určitú hodnotu, napríklad dĺžku, cenu alebo kapacitu. Špeciálnym prípadom je **multigraf**, kde sa môžu vyskytnúť aj **viacnásobné hrany** (medzi jednou dvojicou vrcholov) a **slučky** (hrana z jedného do toho istého vrchola). Zovšeobecnením grafu je **hypergraf**, ktorý umožnuje hrany medzi množinou vrcholov (nielen dvojicou).

Stupeň vrchola udáva, koľko hrán je s daným vrcholom spojených. V prípade orientovaných grafov máme **vstupný** aj **výstupný stupeň** vrchola. Dva vrcholy sú **susedné**, ak ich spája hrana. Pojem **incidencia** vyjadruje vzťah medzi vrcholom a hranou, ak je vrchol jej koncovým bodom. **Sled** je alternujúca postupnosť incidentných vrcholov a hrán (začínajúca aj končiaca vo vrchole). **Ťah** je sled, v ktorom sa neopakuje žiadna hrana. **Cesta** je sled, v ktorom sa neopakuje žiaden vrchol (a teda ani hrana). **Kružnica** v teórii grafov je uzavretý ťah, v ktorom sa prvý a posledný vrchol zhodujú a ostatné vrcholy sa neopakujú. **Podgraf** grafu je časť pôvodného grafu, ktorá vznikne výberom niektorých vrcholov a hrán.

Graf je **súvislý**, ak medzi každými dvoma vrcholmi existuje aspoň jedna cesta. Ak táto podmienka neplatí, graf je **nesúvislý** a skladá sa z viacerých **komponentov súvislosti**. V orientovaných grafoch sa skúma aj **silná súvislosť**, pri ktorej musí existovať orientovaná cesta medzi každými dvoma vrcholmi v oboch smeroch. **Artikulácia** je vrchol, ktorého odstránenie zvýši počet komponentov súvislosti grafu. **Most** je hrana, ktorej odstránenie zvýši počet komponentov súvislosti grafu.

Úplný graf je taký graf, v ktorom je každý vrchol spojený so všetkými ostatnými vrcholmi. **Regulárny graf** má všetky vrcholy rovnakého stupňa. **Bipartitný graf** možno rozdeliť na dve množiny vrcholov tak, aby každá hrana viedla iba medzi množinami, nie v rámci jednej z nich. **Strom** je súvislý graf, ktorý neobsahuje žiadnu kružnicu. Graf, ktorého každý komponent je strom, nazývame **les**.

Najkratšia cesta medzi dvoma vrcholmi je cesta s minimálnou dĺžkou. Ak je graf ohodnotený, dĺžka cesty je súčet ohodnotení jej hrán. **Vzdialenosť vrcholov** je dĺžka najkratšej cesty medzi dvoma vrcholmi. **Excentricita vrcholu** je najväčšia vzdialenosť daného vrcholu od ostatných vrcholov grafu. **Priemer grafu** je najväčšia excentricita vrcholov v grafe. **Polomer grafu** je najmenšia excentricita vrcholov v grafe. **Kostra grafu** je acyklický podgraf (neobsahujúci kružnicu, teda les), ktorý zachováva komponenty súvislosti. **Minimálna kostra** je kostra s najmenším možným súčtom váh hrán. **Tok** je množstvo, ktoré preteká orientovanou sieťou z počiatočného-ých vrchola-ov (**zdroja**) do koncového-ých vrchola-ov (**ústie**). Pri úlohách s tokom sa sleduje aj kapacita hrán, teda maximálne možné množstvo, ktoré môže hranou prejsť. **Maximálny tok** v sieti je tok, ktorý už nemožno ďalej zväčšiť. **Topologické usporiadanie** v orientovanom acyklickom grafe je také poradie vrcholov, že každá hrana ide z vrchola, ktorý je v poradí skôr, do vrchola, ktorý je v danom poradí neskôr.

Dva grafy sú **izomorfné** vtedy, keď majú rovnakú štruktúru, hoci môžu byť nakreslené inak. To znamená, že medzi ich vrcholmi existuje presné priradenie zachovávajúce susednosť.

4.2 Stavové diagramy

Stavový diagram (*angl.* State Machine Diagram) predstavuje v UML jeden zo základných behaviorálnych nástrojov na modelovanie dynamiky systému. Jeho hlavným cieľom je zachytiť, **ako konkrétny objekt počas svojho životného cyklu prechádza medzi konečným počtom stavov**, pričom tieto prechody sú vyvolané udalosťami. Ide o formu grafického zápisu, ktorá umožňuje presne popísať reakcie systému na vonkajšie aj vnútorné podnety a zároveň ukazuje, že správanie objektu je závislé od jeho predchádzajúceho stavu. Takýto prístup je vhodný najmä pre systémy riadené udalosťami, kde rovnaká udalosť môže mať odlišný efekt v závislosti od aktuálneho stavu objektu.

Stavový diagram vychádza z konceptu konečno-stavového automatu (vizualizácia automatu vychádza z grafovej reprezentácie), ktorý reprezentuje objekt ako entitu schopnú nachádzať sa v jednom z definovaných stavov. Každý **stav predstavuje určitú konfiguráciu podmienok**, ktoré sú v danom okamihu pravdivé — môže **ísť o hodnoty atribútov**, vzťahy s inými objektmi alebo práve prebiehajúcu aktivitu. **Prechod medzi stavmi** je reakciou na udalosť, ktorá môže byť externá (napr. vstup používateľa) alebo interná (napr. dokončenie aktivity). Prechody môžu byť doplnené o **podmienky a akcie**, ktoré sa vykonajú pri zmene stavu.

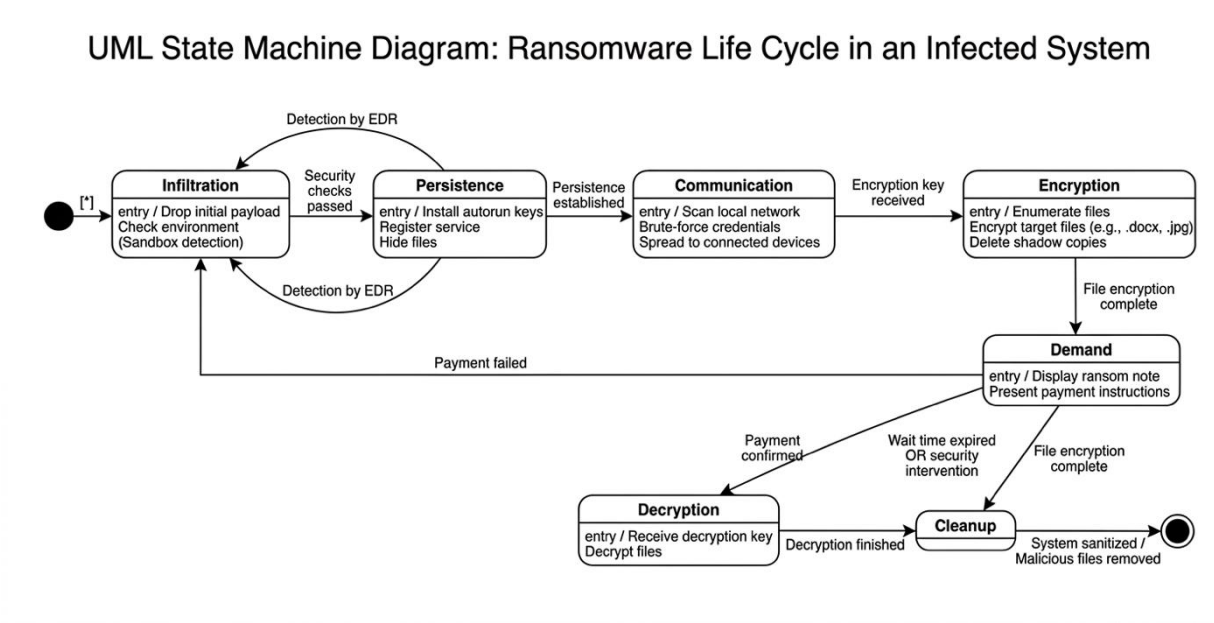
Stavový diagram je vhodný model pre analýzu správania **reaktívnych objektov**, teda takých, ktoré reagujú na podnety počas behu systému. V objektovo orientovanom návrhu sa používa na opis správania tried, subsystémov či celých systémov, pričom umožňuje presne definovať, ako objekt reaguje na rôzne udalosti v rôznych fázach svojho životného cyklu. Tým podporuje analýzu, návrh aj testovanie softvéru, pretože poskytuje jasnú vizualizáciu dynamiky systému.

Základné prvky stavového diagramu sú:

- **Počiatočný stav** – znázornený plnou čiernou bodkou, označuje začiatok životného cyklu objektu.
- **Stav** – zaoblený obdĺžnik reprezentujúci situáciu, v ktorej sa objekt nachádza (determinovanú hodnotami atribútov objektu); môže obsahovať akcie a aktivity.
- **Prechod** – orientovaná šípka medzi stavmi, označená udalosťou, voliteľne podmienkou a akciou.
- **Koncový stav** – dvojitý kruh označujúci ukončenie životného cyklu objektu.
- **Vetvenie a spojenie** – prvky umožňujúce modelovať začiatok a koniec paralelných vetiev správania.

- **Slučka** (*angl.* self-transition) – prechod, ktorý nespôsobuje zmenu stavu.

Zatiaľ čo každý životný cyklus má práve jeden počiatočný stav, životný cyklus môže mať viacero koncových stavov, ktoré zodpovedajú rôznym verziám vývoja systému.



Obrázok 1 - **Stavový diagram pre systém infikovaný ransomwarom (vygenerovaný Gemini 3 Flash dňa 30. 04. 2026)**

Stavový diagram sa používa v situáciách, kde je dôležité presne definovať reakcie systému na udalosti a kde správanie závisí od predchádzajúcich stavov. Z pohľadu digitálnej forenznej analýzy sú vhodnými objektmi, ktoré je možné popisovať stavovým diagramom:

- používateľská relácia,
- proces,
- súbor,
- zápis v registri.

Na základe jednotlivých životných cyklov (inštancií) je môže vytvárať stavové diagramy, ktoré predstavujú triedy – všeobecné vzory správania (pozri napr. [2]). Vďaka jasnej štruktúre stavov a prechodov umožňuje pochopiť, ako systém reaguje na rôzne udalosti, a poskytuje dôležitý nástroj na analýzu systému. Keďže viacero kybernetických útokov v sebe zahŕňa typické šablóny správania, môže tento prístup napomôcť pri predikcii alebo postmortem analýze kybernetických útokov. Jeho sila spočíva v schopnosti zachytiť dynamiku systému v čase a zobraziť ju spôsobom, ktorý je zrozumiteľný tak pre strojovú analýzu ako aj forenzného analytika. Príklad takého diagramu znázorňuje Obrázok 1.

4.3 Petriho siete

Petriho siete sú formálnym nástrojom na **modelovanie diskrétnych paralelných systémov**. Koncepčne vychádzajú z grafovej štruktúry, a preto okrem algebrickej reprezentácie poskytujú aj možnosť vizualizácie procesov, čo je zvlášť užitočné pri simulácií procesov. Petriho siete umožňujú analyzovať zdieľanie prostriedkov a synchrónne aj asynchrónne správanie systémov, a sú istým typom zovšeobecnenia konečnostavových automatov.

Z formálneho hľadiska je Petriho sieť **orientovaný bipartitný graf**, čo znamená, že obsahuje dva typy uzlov, ktoré sa striedajú na jednotlivých cestách grafu:

- **miesta (***angl.* **places) -** reprezentované kruhmi, predstavujú napr. fyzické miesta, podmienky alebo stavy.
- **prechody (***angl.* **transitions) -** reprezentované kruhmi, predstavujú udalosti či akcie.

Orientované hrany v sieti môžu spájať iba miesto s prechodom alebo prechod s miestom, nikdy nie dva uzly rovnakého druhu. Hrany reprezentujú smer evolúcie udalostí v systéme.

Dynamický aspekt siete zabezpečujú **značky/žetóny (***angl.* **tokens)**, ktoré sa umiestnené v jednotlivých miestach. Priradenie týchto značiek jednotlivým miestam sa dá zapísať vo forme vektora a nazýva sa **označkovanie (***angl*. **marking)**. Označkovanie definuje aktuálny stav systému.

Pravidlá prechodu sú jednoduché, ale umožňujú komplexné správanie. Miesto nazývame vstupných miestom prechodu, ak z neho vedie orientovaná hrana do uvedeného prechodu. Analogicky definujeme výstupné miesto. V základnej verzii majú hrany kapacitu jedna, v rozšírených verziách sú kapacity vyjadrené kladným celým číslom. Toto číslo reprezentuje počet žetónov, ktoré sú potrebné na aktiváciu prechodu z daného miesta.

- Prechod je **pripravený (***angl.* **enabled)**, ak každé jeho vstupné miesto obsahuje dostatočný počet značiek.
- Pri **odpálení (***angl.* **firing)** prechodu sa značky zo vstupných miest odoberú a do výstupných miest sa pridajú nové značky v počte zodpovedajúcom kapacite prechodov.

Takto formalizovaný model umožňuje pomocou grafových algoritmov analyzovať niekoľko charakteristík procesov:

- **Dosiahnuteľnosť (***angl.* **reachability):** Základný problém analýzy, ktorý skúma, či sa systém z počiatočného stavu môže niekedy dostať do cieľového (napr. koncového alebo chybového) stavu.
- **Priechodnosť (***angl.* **liveness):** Zaručuje, že systém nezamrzne. Ak je sieť živá, v každom dosiahnuteľnom stave je možné po určitej sekvencii krokov odpáliť ľubovoľný prechod. Tým sa predchádza stavu uviaznutia (*angl.* **deadlock**).
- **Ohraničenosť (***angl.* **boundedness):** Určuje, či počet značiek v ľubovoľnom mieste nepresiahne stanovenú hranicu. Ak je hranica , hovoríme o **bezpečnej sieti**, čo je ideálne pre modelovanie kritických procesov.

V praxi sa základný model často dopĺňa o ďalšie dimenzie, aby lepšie reflektoval komplexnosť reálnych situácií:

- **Farebné Petriho siete (CPN):** Značky nie sú jednoduché objekty, ale nesú dátovú hodnotu (farbu), čo umožňuje modelovať zložité dátové toky a procesy.
- **Časové Petriho siete (TPN):** Do modelu vstupuje časová zložka (oneskorenie prechodov), nevyhnutná pre systémy fungujúce v reálnom čase.
- **Stochastické Petriho siete:** Využívajú sa na modelovanie výkonnosti a spoľahlivosti, kde odpálenie prechodu závisí od pravdepodobnostného rozdelenia vychádzajúceho z reálnych štatistických pozorovaní.

Vďaka svojej exaktnosti sa Petriho siete využívajú pri návrhu distribuovaných systémov, komunikačných protokolov, automatizácie výroby či dokonca pri analýze biologických procesov. Keďže typickou črtou kybernetických útokov je časová závislá postupnosť viacerých udalostí, ktoré navyše manipulujú s viacerými typmi objektov, sú Petriho siete dobrým kandidátom na modelovanie a formálnu analýzu zdokumentovaných a tiež predpokladaných útokov. Pre hlbšie porozumenie Petriho sietí odporúčame napr. publikáciu [3].

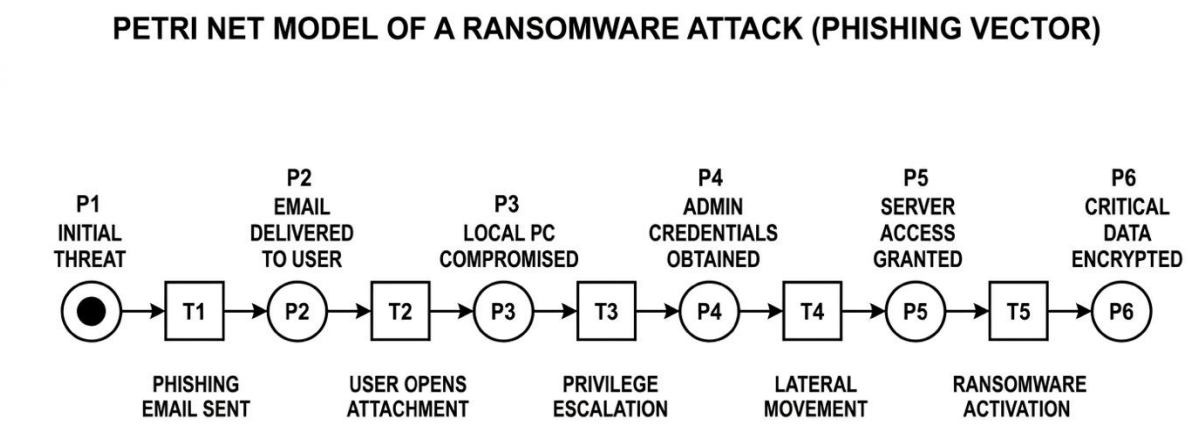
Na ilustráciu si uveďme príklad modelu jednoduchého phishingového útoku. Tento scenár je znázornený pomocou Petriho siete na Obrázku 2. Jednotlivé prvky Petriho siete sú:

Miesta (stavy):

- **P1: Start/Hrozba** (Útočník je pripravený).
- **P2: E-mail v schránke** (Phishingový e-mail doručený obeti).
- **P3: Lokálny prístup** (Malvér beží na PC zamestnanca HR).
- **P4: Admin práva** (Útočník ovládol privilegovaný účet).
- **P5: Prístup k serveru** (Cesta k databáze je voľná).
- **P6: Šifrovanie dokončené** (Cieľ útoku dosiahnutý).

Prechody (akcie):

- **T1: Odoslanie phishingu** (Akcia vyžaduje zdroj/odhodlanie v P1).
- **T2: Spustenie makra** (Interakcia používateľa, ktorá presúva stav do P3).
- **T3: Eskalácia práv** (Využitie zraniteľnosti systému).
- **T4: Laterálny pohyb** (Prienik zo stanice na server).
- **T5: Aktivácia ransomware** (Spustenie deštruktívnej fázy).



Obrázok 2 **- Petriho sieť pre phishinový útok**
(vygenerovaný Gemini 3 Flash dňa 28. 04. 2026)

Pri zložitejších útokoch je samozrejme možné modelovať aj paralelné procesy a zložitejšie podmienky ovplyvňujúce jednotlivé udalosti.

4.4 Procesné diagramy

Procesné diagramy (*angl.* Bussiness Process Diagrams) slúžia na modelovanie a analýzu komplexných procesov rôzneho typu. Predstavujú vizuálny most medzi strategickým pohľadom a implementáciou v konkrétnej platforme. Najrozšírenejším a globálne akceptovaným štandardom pre tieto modely je **BPMN** (*angl.* Business Process Model and Notation), Konceptuálne procesné modelovanie vychádza z **Petriho sieti** a v UML je jeho analógiou modelovanie aktivít. BPMN však ponúka oveľa viac výrazových prostriedkov, a preto sa spravidla využíva v kombinácii s diagramami UML.

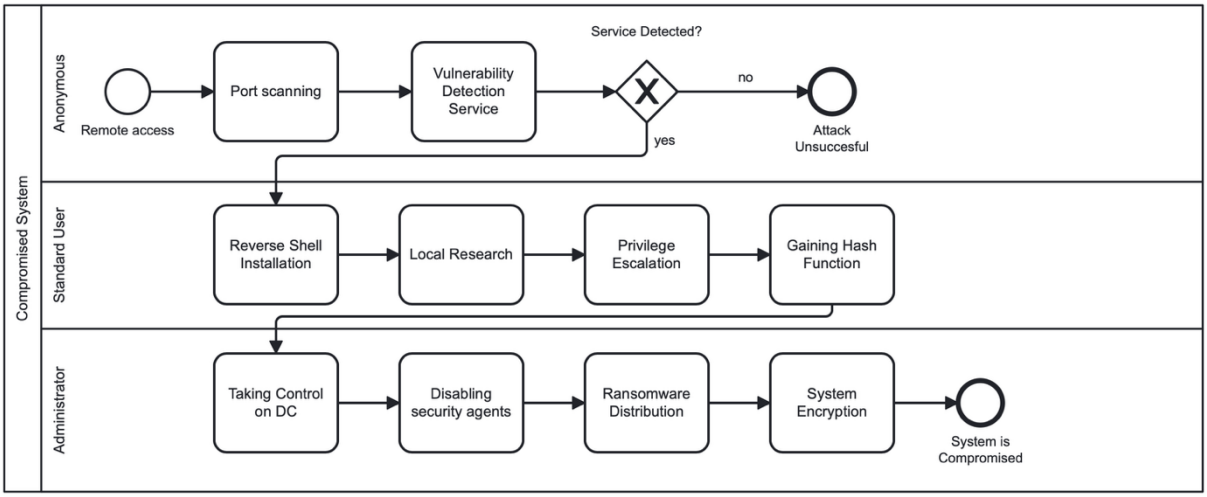
BPMN je modelovací rámec, ktorý poskytuje jednotnú syntax pre biznis analytikov, technických vývojárov a manažérov. Hlavným cieľom je formálne popísať a vizualizovať procesy v systéme. Na rozdiel od vývojových diagramov (*angl.* flowcharts) alebo UML diagramov aktivít, procesné

diagramy obsahujú špecifické sémantické pravidlá, ktoré umožňujú transformáciu vizuálneho modelu do spustiteľného kódu (napr. XML, BPEL).

Notácia BPMN využíva štyri hlavné kategórie grafických prvkov, ktoré definujú štruktúru procesu (bližšie informácie sú uvedené v [3]):

- **Objekty toku (***angl.* flow objects**):**
  - **udalosti (***angl.* events**):** znázornené kruhmi, reprezentujú niečo, čo sa „stane“ (napr. prijatie e-mailu, uplynutie času); rozlišujeme začiatočné, priebežné a ukončovacie udalosti.
  - **aktivity (** *angl.* activities**):** obdĺžniky so zaoblenými rohmi znázorňujú činnosti vykonávané v rámci procesu (úlohy alebo podprocesy).
  - **brány (** *angl.* gateways**):** kosoštvorce riadiace vetvenie a spájanie tokov na základe rozhodovacích podmienok (logické spojky AND, OR, XOR, komplexné, ...).
- **pripájacie objekty (***angl.* connecting objects**):** definujú smer toku sekvencií, správ (komunikácia medzi účastníkmi) a asociácií k dátam.
- **plavecké dráhy (***angl.* swimlanes**):**
  - **bazén (***angl.* poll**):** reprezentuje účastníka procesu,
  - **dráha (***angl.* lane**):** samostatná oblasť v rámci poolu, ktoré priraďuje zodpovednosť konkrétnym rolám alebo oddeleniam.
- **artefakty:** slúžia na doplnenie informácií, ako sú dátové objekty alebo textové anotácie, bez priameho vplyvu na logiku toku.

Procesné diagramy umožňujú modelovať jeden systém alebo aj interakciu viacerých systémov. S tým súvisia pojmy **orchestrácia** a **choreografia. Orchestrácia** predstavuje pohľad z perspektívy jedného centrálneho subjektu (procesu), ktorý riadi a koordinuje ostatné aktivity podobne ako dirigent v orchestri. Tento proces má pod kontrolou poradie úloh a presne vie, kedy sa majú vykonať. **Choreografia** sa naopak zameriava na interakciu medzi dvoma alebo viacerými účastníkmi bez prítomnosti centrálneho riadenia. Modeluje výmenu správ a pravidlá spolupráce medzi partnermi, podobne ako tanečníci, ktorí poznajú svoje kroky a reagujú na pohyby ostatných bez toho, aby ich niekto priamo riadil.



Obrázok 3 **- Procesný model ransomvérového útoku**

Obrázok 3 predstavuje príklad procesného modelu ransomvérového útoku, ktorý zachytáva postupnosť jednotlivých krokov od iniciálnej kompromitácie systému až po finálnu fázu šifrovania dát. Model ilustruje, ako útočník postupne prechádza rôznymi fázami útoku, pričom jednotlivé aktivity sú reprezentované ako sekvencia na seba nadväzujúcich procesov. Prvotná fáza začína skenovaním zariadenia a dostupných služieb, koniec sekvencie je kompromitovaný systém so zašifrovanými súbormi. Zároveň riadky v procesnom modeli ukazujú ako útočník postupne menil svoju rolu – od anonymného používateľa zo siete, cez štandardného používateľa, až po administrátora.

5 Aplikácie grafových modelov v rámci projektu ADFIR

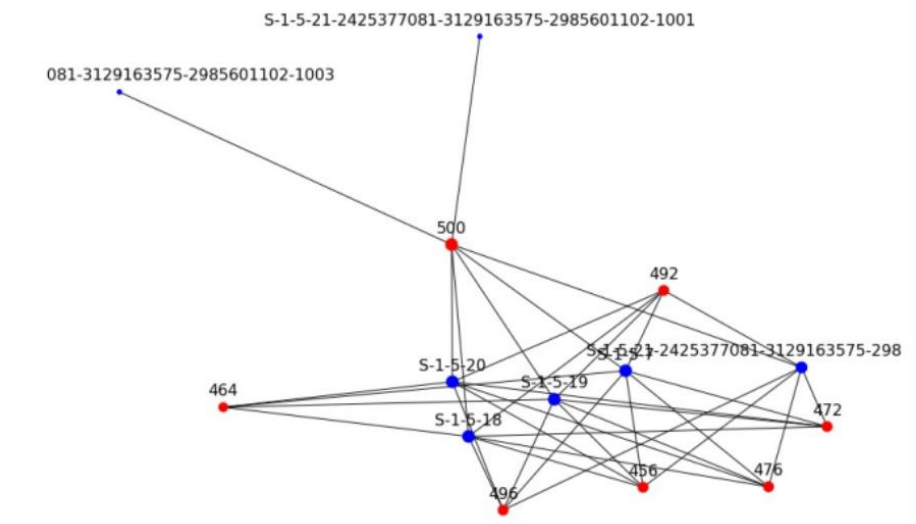
V závislosti od charakteru dostupnej dátovej sady, typu a štruktúry objektov v dátovej sade a sledovaného javu, môžeme zvoliť rôznu grafovú reprezentáciu. Vhodná reprezentácia nám umožní vizualizovať vybrané črty analyzovanej sady a vytvoriť predpoklady na použitie vhodných algoritmov.

V nasledovných častiach si ukážeme využitie klasického grafu, stavového diagramu a procesného modelu na analýzu nám dostupných dátových zdrojov. Pri ukážkach jednotlivých grafov vychádzame z práce [4].

5.1 Tvorba grafov z digitálnych stôp

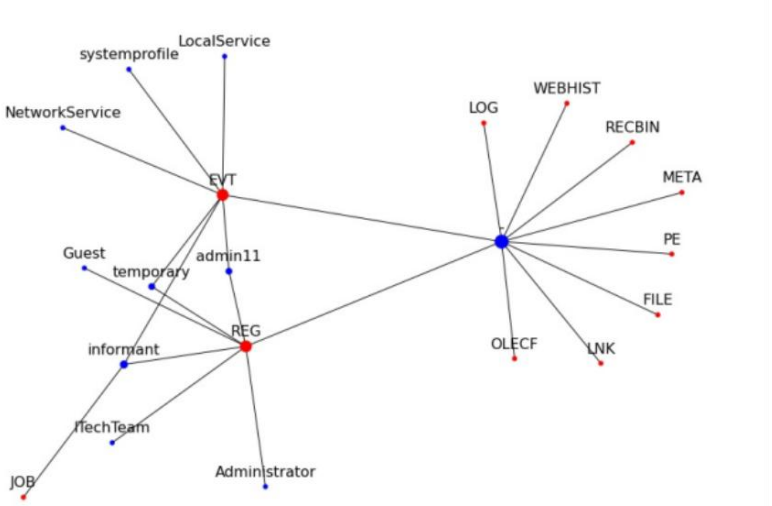
V prípade **post-mortem analýzy** digitálnych stôp sú zvyčajne dostupné záznamy tie artefakty, ktoré sú uložené na sekundárnych úložiskách (mimo primárnej operačnej pamäte). Tieto artefakty môžeme analyzovať pomocou rôznych analytických nástrojov (v rámci tohto materiálu sa zameriavame na nástroje Plaso [5] a nástroje od Erica Zimmermana [6]. Pomocou týchto nástrojov možno získať extrahované dáta (zvyčajne vo formáte CSV), z ktorých je možné vytvoriť rôzne grafy (podľa jednotlivých forenzných artefaktov).

Z digitálnych stôp uvedených vo Windows Event Logoch sme vytvorili graf z dvoch atribútov *user_id* a *execution_process_id*. Takto vznikne bipartitný graf, kde vrcholy reprezentujú používateľov (na obrázku označené modrou farbou) a procesy (označené červenou farbou). Hrana v grafe znamená, že v záznamoch sa vyskytol riadok obsahujúci daného používateľa a príslušný proces. Na Obrázku 4 je znázornený graf z dátovej sady ”NIST Data Leakage Case” [10].



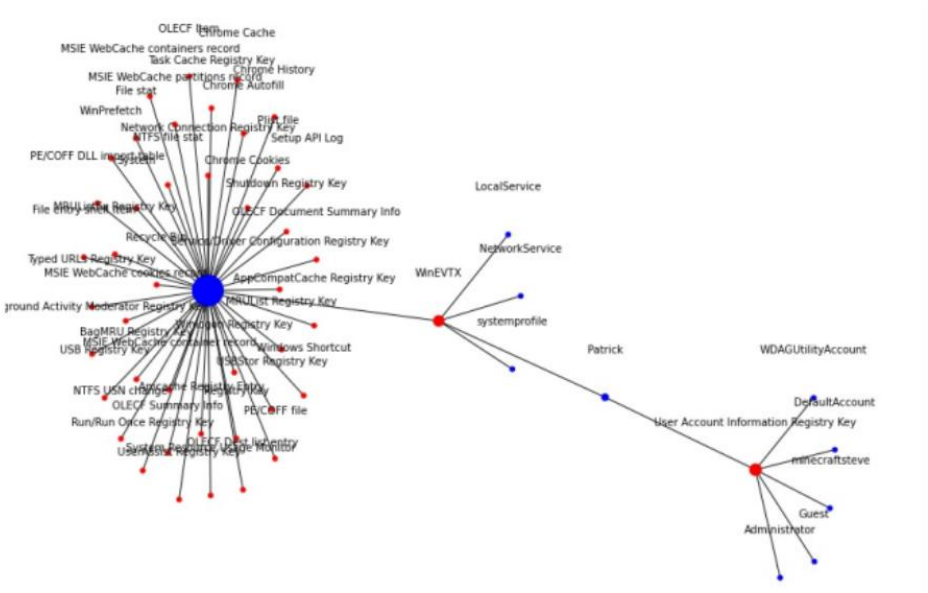
Obrázok 4 - Používatelia a procesy v grafe prípadu NIST Data Leakage Case

Na Obrázku 5 je graf vytvorený z atribútov *user* a *source*, takisto z ”NIST Data Leakage Case” [10]. Menom používateľa sú označené vrcholy červenej farby, zdrojom vrcholy modrej farby. Ako zdroj je uvedené umiestnenie artefaktov (napr. registre, logy, ...). Hrana reprezentuje existenciu záznamu príslušného používateľa v zdrojových dátach. Pre forenzného analytika to môže prehľadnejšie naznačiť, na ktorých používateľov sa má zamerať pri forenznej analýze.



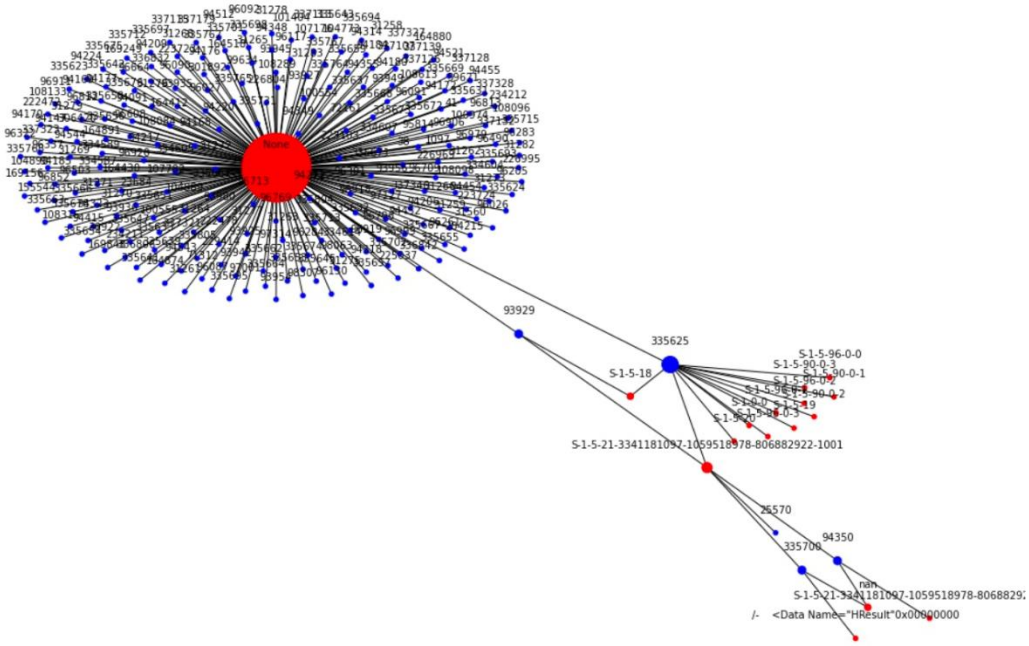
Obrázok 5 - Používatelia a zdroje v grafe prípadu NIST Data Leakage Case

Na dátach zo súťaže Capture The Flag MAGNET [8, 9] je možné v grafe vygenerovaného z atribútov *user* a *source* sledovať excentricitu vrcholov. Centrum grafu tvorí vrchol reprezentujúci logy (WinEVTX) a vrchol reprezentujúci jedného používateľa (na Obrázku 6).



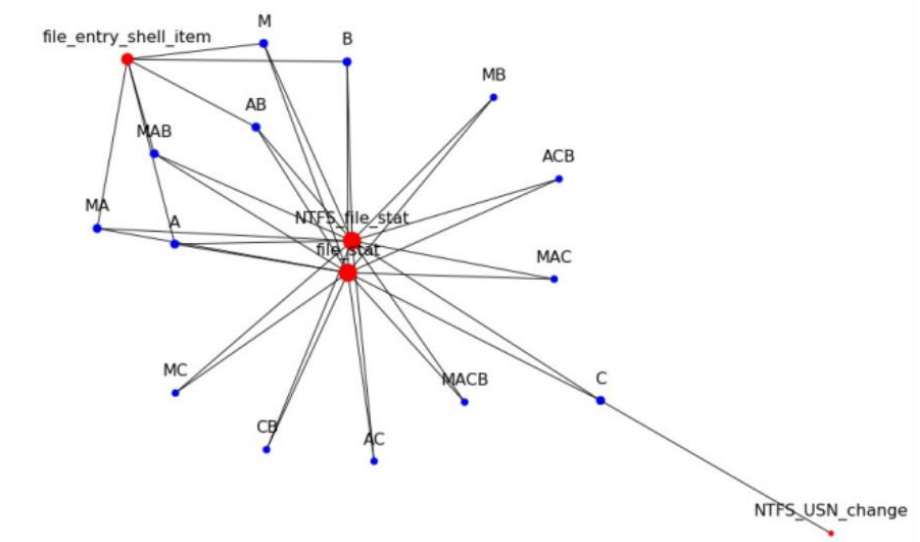
Obrázok 6 - Používatelia a zdroje v grafe prípadu Magnet CTF 2022

Pri vytvorení grafu medzi používateľmi a súbormi (či už názvy alebo číselné inody) je možné v grafe hľadať vrcholy s vysokým stupňom vrcholu – to môže znamenať napr. ransomware šifrujúci súbory, ale takisto aj zálohovanie. Na Obrázku 7 je ukážka takéhoto grafu z prípadu Magnet CTF 2022 [9], údaje sú z Windows Event Log. Červenou farbou sú označení používatelia, modrou číselné inody súborov.



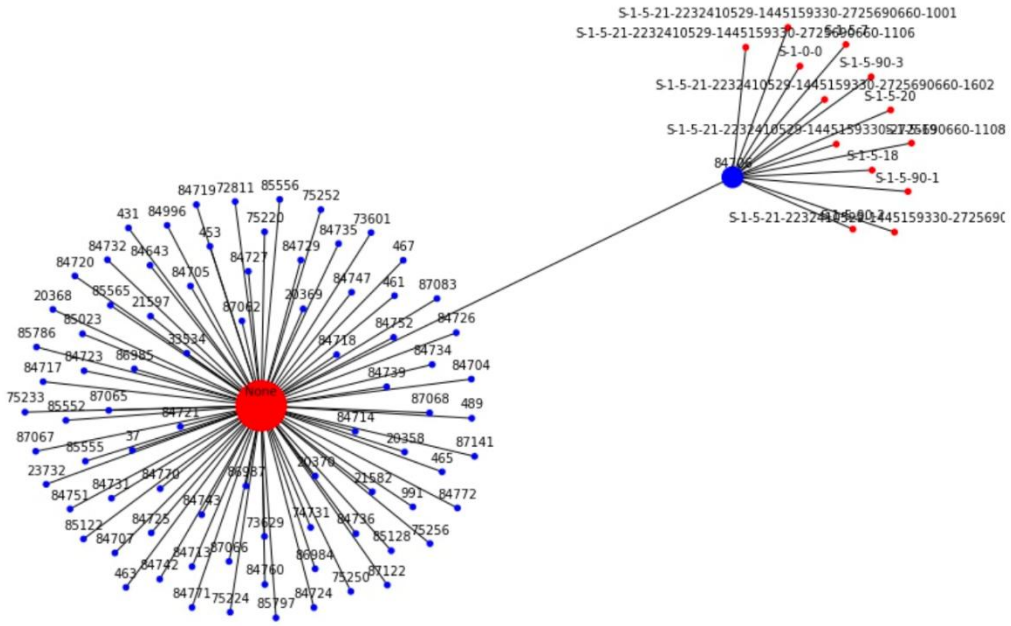
Obrázok 7 - Používatelia a inody súborov v grafe prípadu Magnet CTF 2022

Z digitálnych stôp týkajúcich sa súborovému systému je možné sledovať jednotlivé súbory a časové pečiatky (*Modification*, *Access*, *Change* a *Birth*). Na Obrázku 8 je ukážka takéhoto grafu zo súťaže ”Stolen Szechuan Sauce” [7]. Červenou označené vrcholy reprezentujú súbory a modrou farbou sú označené vrcholy reprezentujúce kombináciu časových pečiatok.



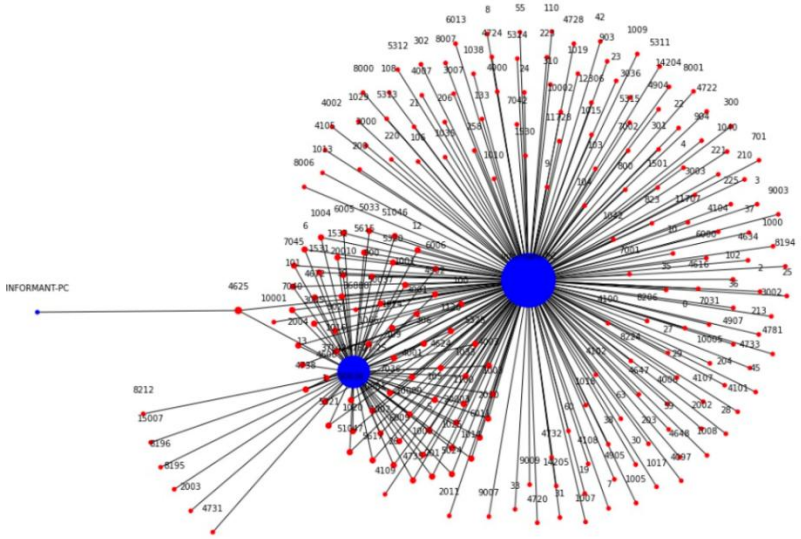
Obrázok 8 - Časové pečiatky a typ súborov v grafe prípadu Stolen Szechuan Sauce

Obrázok 9 zobrazuje graf z prípadu Stolen Szechuan Sauce [7] vytvorený z atribútov *user* a *inode*. Počet unikátnych hodnôt atribútu *inode* bol 81 a počet unikátnych hodnôt *user* bol 14, ale z grafu je viditeľné, že so súbormi nepracovali všetci používatelia. Údaje sú z Windows Event Log. Červenou farbou sú označení používatelia, modrou číselné inody súborov.



Obrázok 9 - Používatelia s inody v prípade Szechuan Sauce EVT

Na Obrázku 10 je graf vytvorený z atribútov *computer_id* a *event_id* z "NIST Data Leakage Case" [10]. Tu hľadanie kružníc ukázalo, že všetky kružnice majú spoločné tri vrcholy a líšia sa len v jednom vrchole reprezentujúcom *event_id* (červený vrchol).

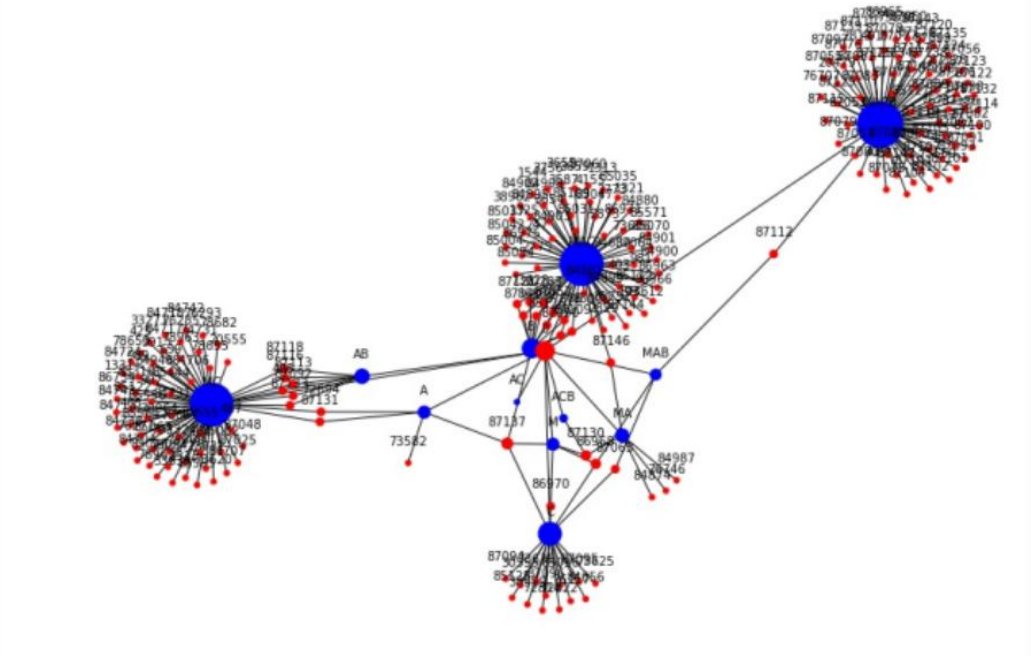


Obrázok 10 - Počítače a záznamy v grafe prípadu NIST Data Leakage Case

Obrázok 11 zobrazuje graf z prípadu "Stolen Szechuan Sauce" [7] vytvorený z atribútov *type* a *MACB* obsahuje samostatné komponenty súvislosti. Z nich vidieť, ako jednotlivé typy (červené vrcholy) menia rôzne časové pečiatky súborov (modré vrcholy).

Obrázok 11 – Typ a množina časových pečiatok v grafe prípadu Stolen Szechuan Sauce

Poslednou ukážkou je graf takisto zo súťaže ”Stolen Szechuan Sauce” [7]. Reprezentuje údaje zo súborového systému - číselné inody a kombináciu časových pečiatok. Z komplexnejšieho Obrázka 12 je viditeľné, na čo konkrétne sa zamerať, takže sme skúšali rôzne vlastnosti grafu. Hľadali sme kružnice v grafe. Našli sme ich 43 a sedem kružníc obsahovalo inody, ktoré boli analýzou manuálne identifikované ako relevantné pre prípad.



Obrázok 12 - Inody a časové pečiatky v grafe prípadu Stolen Szechuan Sauce

Zobrazené boli iba neohodnotené grafy, samozrejme je možné doplniť hrany aj o váhu (napr. početnosť výskytu danej udalosti).

Pri online analýze je navyše možné analyzovať aj zmeny stavu – ako procesy vznikajú/zanikajú. Je teda možné vytvoriť orientovaný graf obsahujúci viacero typov vrcholov (používateľ, proces, súbor, register, sieťové spojenie, ...). Orientácia hrany sa štandardne nastaví podľa časových pečiatok jednotlivých zaznamenaných udalostí.

5.2 Tvorba stavových diagramov z digitálnych stôp

V Tabuľke 1 sú zachytené vybrané záznamy, ktoré sémantická analýza označila ako podozrivú aktivitu. Tabuľka predstavuje výsek z časovej osi udalostí v operačnom systéme, pričom jednotlivé záznamy sú zoradené chronologicky a zachytávajú sled operácií, ktoré spolu môžu tvoriť ucelený kybernetický bezpečnostný incident.

row_id	timestamp	attack_ types	mitre_ techniques	killchain_ stage	source_type
SSS_DC:10	2020-09-19 T03:56:03	Remote Desktop Protocol Connection	T1021.001	Delivery	Content Modification Time
SSS_DC:14	2020-09-19 T03:56:03	Privilege Escalation			
SSS_DC:15	2020-09-19 T03:56:03	Privilege Escalation	T1078	Exploitation	Content Modification Time
SSS_DC:143	2020-09-19 T03:56:04	Service Manipulation	T1543.003	Installation	Content Modification Time
SSS_DC:731	2020-09-19 T03:56:11	Privilege Escalation	T1068	Exploitation	Content Modification Time
SSS_DC:787	2020-09-19 T03:56:15	Malware Installation	T1105	Installation	Metadata Modification Time
SSS_DC:788	2020-09-19 T03:56:15	Malware Installation	T1105	Installation	Metadata Modification Time
SSS_DC:789	2020-09-19 T03:56:15	Malware Installation	T1105	Installation	Metadata Modification Time
SSS_DC:790	2020-09-19 T03:56:15	Malware Installation	T1105	Installation	Metadata Modification Time
SSS_DC:792	2020-09-19 T03:56:15	Malware Installation	T1105	Installation	Metadata Modification Time
SSS_DC:799	2020-09-19 T03:56:15	Malware Installation	T1105	Installation	Metadata Modification Time
SSS_DC:834	2020-09-19 T03:56:53	Privilege Escalation	T1078	Exploitation	Content Modification Time
SSS_DC:837	2020-09-19 T03:56:55	Service Installation	T1543.003	Installation	Content Modification Time
SSS_DC:836	2020-09-19 T03:56:55	Privilege Escalation	T1078	Exploitation	Content Modification Time
SSS_DC:840	2020-09-19 T03:56:55	Privilege Escalation	T1078	Exploitation	Content Modification Time
SSS_DC:844	2020-09-19 T03:56:55	Privilege Escalation	T1078	Exploitation	Content Modification Time

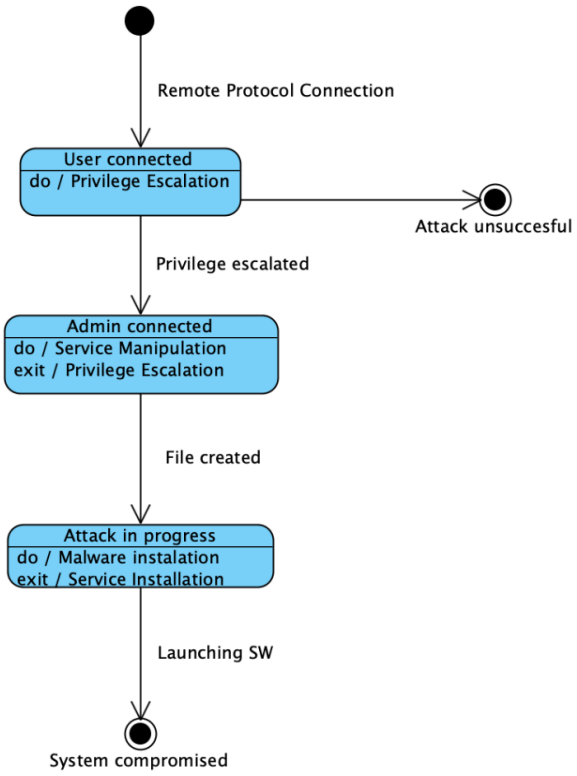
Tabuľka 1 – Výsek z časovej osi podozrivých aktivít v systéme I

Z kompletného logu pre časovú os udalostí v predchádzajúcej tabuľke je možné vyčítať, že podozrivé aktivity mali nasledovný charakter:

- nadviazanie anonymného spojenia,

- pokus o získanie administrátorských práv,
- modifikácia prístupových práv,
- inicializácia inštalácie súboru,
- zmena prístupových práv,
- inštalácia podozrivého súboru 87131-3 vo viacerých krokoch,
- opakovaný pokus o zmenu prístupových práv k súboru.

Jednu možnú reprezentáciu časovej osi uvedeného podozrivého správania pomocou stavového diagramu znázorňuje Obrázok 13. Stavový diagram modeluje jednotlivé fázy aktivity ako diskrétne stavy systému a prechody medzi nimi ako reakcie na konkrétne udalosti zachytené v logoch.



Obrázok 12 - Stavový diagram pre časovú os podozrivých aktivít v systéme I

Na Obrázku 13 je znázornený postupný prechod systému od počiatočného nadviazania vzdialeného spojenia (Remote Protocol Connection) do stavu „User connected“, v ktorom dochádza k pokusu o eskaláciu oprávnení. V prípade neúspechu je proces ukončený stavom „Attack unsuccessful“.

V prípade úspešnej eskalácie práv systém prechádza do stavu „Admin connected“, ktorý reprezentuje získanie administrátorských oprávnení a umožňuje ďalšie operácie, ako napríklad manipuláciu so službami. Následne dochádza k vytvoreniu súboru a prechodu do stavu „Attack in progress“, v ktorom prebieha inštalácia škodlivého softvéru.

Finálnym krokom je spustenie škodlivého programu („Launching SW“), ktoré vedie do koncového stavu „System compromised“, reprezentujúceho úspešné kompromitovanie systému. Diagram zároveň prostredníctvom alternatívnej vetvy zachytáva možnosť zlyhania útoku, čím poskytuje komplexnejší pohľad na možné scenáre vývoja incidentu.

Takto vytvorený stavový diagram umožňuje abstrahovať od jednotlivých logovacích záznamov a zamerať sa na kľúčové fázy útoku, čím uľahčuje jeho analýzu a pochopenie.

5.3 Tvorba procesných diagramov z digitálnych stôp

V tejto časti si ukážeme využitie procesného modelovania pomocou BPMN, ktoré nám umožňuje zvýrazniť iné typy forenzných stôp. Ako vstup slúži Tabuľka 2 – výsek z časovej osi podozrivých aktivít v systéme II, ktorá zachytáva sled udalostí relevantných pre ďalšiu analýzu.

row_id	timestamp	attack_ types	mitre_ tactics	mitre_ techniques	killchain_ stage
SSS_DC:7	2020-09-19 T03:22:07	Remote Access	Initial Access	T1021.001	Delivery
SSS_DC:77	2020-09-19 T03:22:08	Credential Access			
SSS_DC:155	2020-09-19 T03:22:09	Credential Use	Persistence	T1078.002	Exploitation
SSS_DC:161	2020-09-19 T03:22:09	Credential Use	Persistence	T1078.002	Exploitation
SSS_DC:538	2020-09-19 T03:22:37	Remote Access	Command and Control	T1021.001	Command and Control
SSS_DC:1123	2020-09-19 T03:23:01	Persistence	Persistence	T1547.001	Installation
SSS_DC:1127	2020-09-19 T03:23:01	Persistence	Persistence	T1547.001	Installation
SSS_DC:1128	2020-09-19 T03:23:01	Persistence	Persistence	T1547.001	Installation
SSS_DC:1133	2020-09-19 T03:23:01	Persistence	Persistence	T1547.001	Installation
SSS_DC:1143	2020-09-19 T03:23:01	Persistence	Persistence	T1547.001	Installation

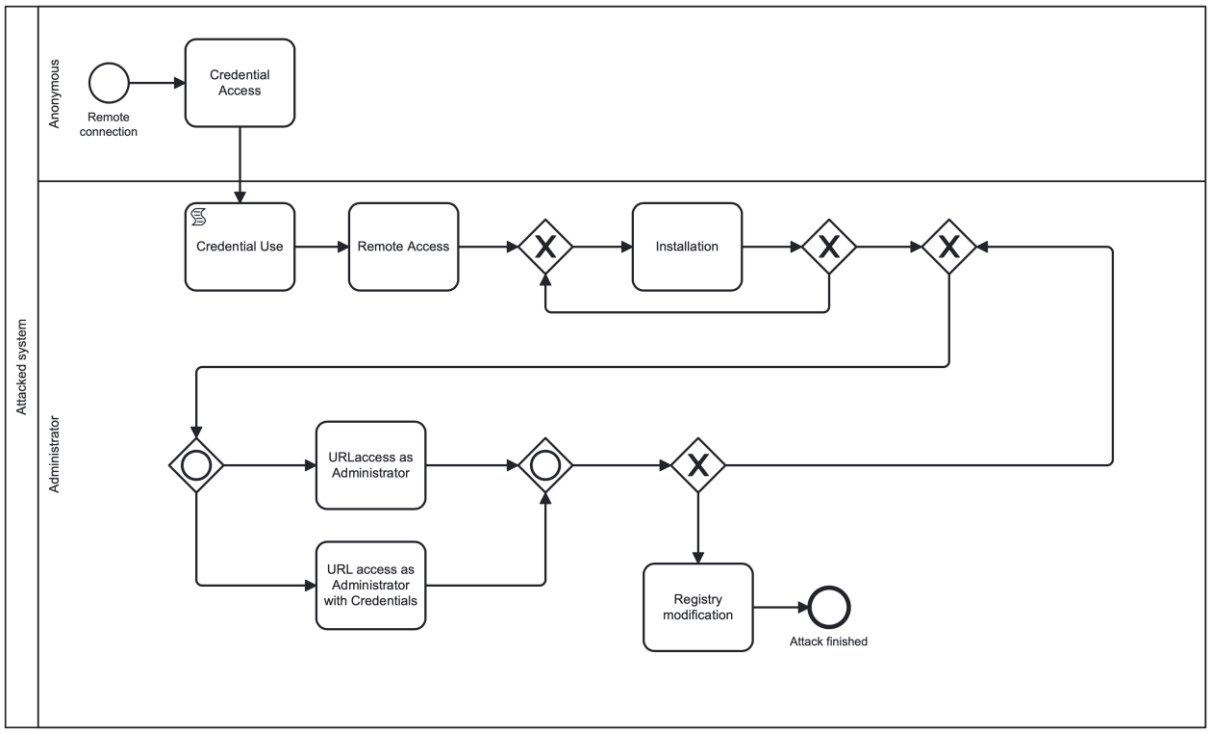
SSS_DC:1566	2020-09-19 T03:23:41	Web Traffic to Malicious IP	Command and Control	T1071.001	Command and Control
SSS_DC:1567	2020-09-19 T03:23:41	Web Traffic to Malicious IP	Command and Control	T1071.001	Command and Control
SSS_DC:1569	2020-09-19 T03:23:41	Web Traffic to Malicious IP	Command and Control	T1071.001	Command and Control
SSS_DC:1575	2020-09-19 T03:23:41	Web Traffic to Malicious IP	Command and Control	T1071.001	Command and Control
SSS_DC:1577	2020-09-19 T03:23:41	Web Traffic to Malicious IP	Command and Control	T1071.001	Command and Control
SSS_DC:1579	2020-09-19 T03:23:41	Web Traffic to Malicious IP	Command and Control	T1071.001	Command and Control
SSS_DC:1580	2020-09-19 T03:23:41	Web Traffic to Malicious IP	Command and Control	T1071.001	Command and Control
SSS_DC:1581	2020-09-19 T03:23:41	Web Traffic to Malicious IP	Command and Control	T1071.001	Command and Control
SSS_DC:1582	2020-09-19 T03:23:41	Web Traffic to Malicious IP	Command and Control	T1071.001	Command and Control
SSS_DC:1591	2020-09-19 T03:23:41	Web Traffic to Malicious IP	Command and Control	T1071.001	Command and Control
SSS_DC:1647	2020-09-19 T03:24:00	C2 Communication	Command and Control	T1102	Command and Control

Tabuľka 2 – Výsek z časovej osi podozrivých aktivít v systéme II

Z kompletného logu zodpovedajúceho vyššie uvedenej tabuľke 2 je možné vyčítať nasledovný sled podozrivých aktivít v systéme:

- vzdialený prístup do systému
- získanie prístupu ako administrátor podhodením prihlasovacích údajov
- opakovaný prieskum systému
- inštalácia podozrivého súboru
- zápis do registrov
- pokus o navštívenie URL 194.61.24.102 v roli administrátora, stiahnutie súboru *favicon.ico*
- zápis do registrov
- opakovaný pokus o prístup na URL v roli administrátora
- opakovaný prístup na URL v roli administrátora s podhodenými údajmi
- opakovaný prístup na URL v roli administrátora
- modifikácia registrov.

Pre každú takúto postupnosť udalostí existuje viacero interpretácií a možností ako ich reprezentovať. Jeden z možných modelov znázorňuje Obrázok 14.



Obrázok 14 - Procesný model pre časovú os podozrivých aktivít v systéme II

Obrázok 14 predstavuje procesný model vytvorený pomocou BPMN, ktorý zachytáva priebeh podozrivých aktivít v systéme z pohľadu jednotlivých aktérov a ich interakcií. Diagram je rozdelený do viacerých plaveckých dráh (lanes), ktoré reprezentujú rôzne entity zapojené do incidentu, napríklad útočníka (anonymous), napadnutý systém a administrátora.

Model začína nadviazaním anonymného vzdialeného spojenia, po ktorom nasleduje získanie a použitie prihlasovacích údajov (credential access a credential use). Tieto kroky vedú k získaniu vzdialeného prístupu do systému, čo predstavuje kľúčový bod kompromitácie. Následne dochádza k inštalácii škodlivého komponentu, pričom diagram zachytáva aj možné alternatívne alebo opakované priebehy prostredníctvom rozhodovacích uzlov (gateway).

V ďalšej fáze model zobrazuje aktivity súvisiace s administrátorským prístupom, konkrétne prístup k URL zdrojom buď priamo, alebo prostredníctvom získaných prihlasovacích údajov. Tieto vetvy sa následne spájajú a vedú k modifikácii systémového registra, ktorá môže predstavovať mechanizmus perzistencie alebo ďalšej manipulácie so systémom.

Celý proces je ukončený stavom „attack finished“, ktorý reprezentuje úspešné dokončenie útoku. Diagram zároveň prostredníctvom rozhodovacích bodov a spätných väzieb naznačuje,

že útok nemusí prebiehať striktne lineárne, ale môže obsahovať opakované pokusy alebo alternatívne scenáre.

Takto vytvorený procesný model umožňuje lepšie pochopiť logiku útoku, identifikovať kritické fázy a zároveň poskytuje základ pre návrh detekčných alebo obranných mechanizmov.

6 Záver

Navrhovaný model extrakcie digitálnych stôp do grafovej reprezentácie predstavuje efektívny rámec pre formálne zachytenie vzťahov medzi forenznými artefaktmi. Vytvorené grafové procesné a stavové diagramy poskytujú základ pre následnú analýzu priebehu kybernetického bezpečnostného incidentu, identifikáciu kľúčových udalostí a odhalenie anomálnych štruktúr v dátach.

Pomocou grafových algoritmov (napr. hľadanie podgrafov) je možné identifikovať vzory reprezentujúce typické známe schémy útokov a buď predikovať ich vývoj v reálnom čase alebo v rámci postmortem analýzy získať nové informácie o správaní sa útočníka.

Ďalšie pokračovanie analýzy by malo smerovať k automatizovanej detekcii vzťahov, k porovnaniu viacerých kybernetických bezpečnostných incidentov a k overeniu, do akej miery grafová reprezentácia podporuje presnejšiu a rýchlejšiu rekonštrukciu bezpečnostných udalostí.

7 Bibliografia

[1] Diestel, R. (2025). *Graph theory* (6th ed.). Springer. https://doi.org/10.1007/978-3-662-70107-2

[2] Olivé, A. (2007). *Conceptual modeling of information systems* (1st ed.). Springer. https://doi.org/10.1007/978-3-540-39390-0

[3] Weske, M. (2024). *Business process management: Concepts, languages, architectures* (4th ed.). Springer. https://doi.org/10.1007/978-3-662-69518-0

[4] Krišáková, S. P., Sokol, P., & Krivoš-Belluš, R. (2024). Forensic Artifacts' Analysis using Graph Theory. CEUR Workshop Proceedings. https://ceur-ws.org/Vol-3792/paper26.pdf

[5] Plaso (log2timeline), 2022 [online]. Dostupné na: https://github.com/log2timeline/plaso

[6] Eric Zimmerman's Tools*, 2026* [online]. Dostupné na: https://ericzimmerman.github.io/

[7] Case 001 – the stolen szechuan sauce, 2020 [online]. Dostupné na: https://dfirmadness.com/the-stolen-szechuan-sauce/

[8] Magnet CTF 2019 Windows Desktop, 2019 [online]. Dostupné na: https://digitalcorpora.s3.amazonaws.com/corpora/scenarios/magnet/2019 CTF - Windows-Desktop.zip

[9] Magnet CTF 2022 Windows, 2022 [online]. Dostupné na: https://digitalcorpora.s3.amazonaws.com/corpora/scenarios/magnet/2022 CTF - Windows.zip

[10] Data LeakageCase, 2018 [online]. Dostupné na: https://cfreds-archive.nist.gov/data_leakage_case/data-leakage-case.html