SLOVENSKÁ TECHNICKÁ
UNIVERZITA V BRATISLAVE
FAKULTA ELEKTROTECHNIKY
A INFORMATIKY

Ing. Peter Švec

Autoreferát dizertačnej práce

Ontologická reprezentácia pre bezpečnosť informačných systémov

na získanie vedecko-akademickej hodnosti
philosophiae doctor, *PhD.*

v doktorandskom študijnom programe:	Aplikovaná informatika
v študijnom odbore:	Informatika
Forma štúdia:	denná

Bratislava, 2024

Slovenská Technická univerzita v Bratislave
Fakulta Elektrotechniky a Informatiky

Ing. Peter Švec

Autoreferát dizertačnej práce

Ontologická reprezentácia pre bezpečnosť informačných systémov

na získanie vedecko-akademickej hodnosti
philosophiae doctor, *PhD.*

v doktorandskom študijnom programe:	Aplikovaná informatika
v študijnom odbore:	Informatika
Forma štúdia:	denná

Bratislava, 2024

Predkladateľ: Ing. Peter Švec
FEI STU
Ilkovičova 3, 812 19 Bratislava 1

Školiteľ: prof. Ing. Pavol Zajac, PhD.
FEI STU
Ilkovičova 3, 812 19 Bratislava 1

Konzultant: Ing. Štefan Balogh, PhD.
FEI STU
Ilkovičova 3, 812 19 Bratislava 1

Oponenti: doc. RNDr. Damas Gruska, PhD.
Katedra aplikovanej informatiky
FMFI Univerzita Komenského
Mlynská dolina, 824 48 Bratislava

prof. Ing. Kristína Machová, PhD.
Katedra kybernetiky a umelej inteligencie
FEI Technická univerzita v Košiciach
Letná 9, 042 00 Košice

Autoreferát bol rozoslaný dňa .....

Obhajoba dizertačnej práce sa koná ..... o .....
na **FEI STU**, Ilkovičova 3, 812 19 Bratislava 1, v miestnosti .....

prof. Ing. Vladimír Kutiš, PhD.
dekan FEI STU v Bratislave

Obsah

Úvod

Rozvoj informačných technológií a internetu v posledných desaťročiach so sebou priniesol aj mnohé hrozby pre podniky, vládne inštitúcie alebo používateľov. Jednou z takýchto hrozieb je aj škodlivý kód (malvér), ktorý sa dokáže vďaka dostupnosti internetu šíriť oveľa rýchlejšie a jednoduchšie ako v minulosti. Podľa najnovších štatistík sa denne objavuje až 450 000 nových vzoriek, pričom od roku 1984 bolo zaznamenaných až 1,5 miliardy škodlivých súborov, čím vznikla potreba vyvájať rôzne detekčné mechanizmy [1]. V posledných rokoch sa ako najúspešnejšie riešenie javilo použitie algoritmov strojového učenia, ktoré dokážu spracovávať veľké množstvo dát a zároveň dosahovať vysokú úspešnosť pri detekcii. Napriek spomínaným výhodám, dané algoritmy majú aj určité nevýhody. Jednou z takýchto nevýhod je, že väčšina algoritmov nie je vysvetliteľná a fungujú ako čierna skrinka v zmysle, že nevieme pochopiť rozhodnutie detekčného systému, ktorý označil vzorku za škodlivú. Vysvetliteľnosť a schopnosť dôverovať rozhodnutiam algoritmom strojového učenia sa stala dôležitou výskumnou témou v tejto oblasti, najmä v posledných rokoch, kedy sa strojové učenie dostáva do kritických oblastí ako sú medicína alebo finančný sektor.

Táto práca sa zaoberá výskumom ontológií a sémantických technológií a ich využitím v oblasti počítačovej bezpečnosti pri vývoji vysvetliteľných systémov na detekciu škodlivého kódu. Jadro práce tvorí výskum algoritmov konceptového učenia (algoritmy, ktoré sa dokážu učiť nad ontológiami) a ich doposiaľ nepreskúmaná aplikácia pri detekcii škodlivého kódu.

1 Ciele dizertačnej práce

Vedecké ciele práce môžeme zhrnúť do nasledujúcich bodov:

1. Návrh ontológie pre škodlivé PE súbory (pre operačný systém *Windows*) spolu s vytvorením a publikovaním štandardizovaných datasetov založených na spomínanej ontológii. Hlavným cieľom tejto tézy bolo vytvoriť dataset s interpretovateľnými vlastnosťami a zároveň poskytnúť možnosť jednoduchšej reprodukcie výsledkov s cieľom efektívnejšieho porovnávania rôznych algoritmov (nielen pre konceptové učenie). Výsledky boli publikované v [2, 3, 4].

2. Výskum existujúcich algoritmov konceptového učenia nad nami navrhnutou ontológiou. Cieľom bolo preskúmať jednotlivé algoritmy konceptového učenia dostupné v softvéri DL-Learner (OCEL, CELOE, PARCEL a SPACEL) a overiť ich úspešnosť pri detekcii škodlivého kódu pri rôznych experimentálnych nastaveniach. Výsledky boli publikované v [5, 6, 7, 8].

3. Analýza bezpečnostných aspektov konceptových výrazov (výsledkov z konceptového učenia). Cieľom bolo preskúmať odolnosť konceptového učenia voči rôznym útokom, ktorých cieľom je pomýliť klasifikátor (t.j. označiť škodlivý kód ako legitímny softvér).

2 Teória a metódy

V tejto kapitole si v stručnosti popíšeme metódu konceptového učenia a nástroj DL-Learner.

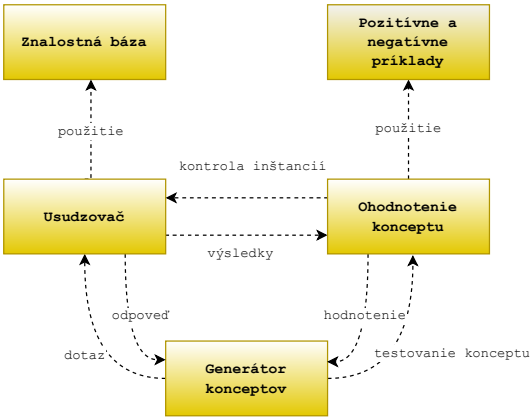
2.1 Konceptové učenie

Hlavnou výskumnou časťou práce je práve strojové učenie nad ontológiami, tzv. *konceptové učenie*, predstavené v práci [9]. Cieľom konceptového učenia je nájsť logický popis konceptu vzhľadom na existujúce (alebo neexistujúce) inštancie konceptu v znalostnej báze [10].

Samotné konceptové učenie má svoj pôvod v *induktívnom logickom programovaní* (ILP) [11, 12]. ILP metódy používajú logický program ako bázu znalostí a snažia sa nájsť logický program popisujúci pozitívne príklady, ktorý zároveň nepopisuje žiadne negatívne príklady. Hlavným rozdielom je, že ILP metódy využívajú logické programy ako znalostné bázy, zatiaľ čo konceptové učenie sa spolieha na deskripčné logiky a OWL.

Konceptové učenie má v praxi viacero využití. Prvotným zámerom konceptového učenia bolo strojové rozširovanie ontológie. Ak teda máme nejakú základnú verziu ontológie, aplikovaním konceptové učenia môžeme získať nové konceptové výrazy, ktoré vieme do danej ontológie doplniť a získať tak expresívnejšiu a komplexnejšiu bázu znalostí. Ďalším využitím konceptového učenia (čo je aj hlavným cieľom práce), je aplikácia naučených výrazov na riešenie klasifikačných problémov. Z vyššie uvedeného textu vyplýva, že konceptové učenie je vhodné najmä na binárnu klasifikáciu (t.j. rozlišovanie medzi dvoma triedami). Existujú však aj iné variácie konceptového učenia.

Formálne je možné konceptového učenie definovať následovne (variant použitý v práci) [13]. Majme znalostnú bázu $\mathcal{K} = (\mathcal{T}, \mathcal{A})$. Majme dve navzájom disjunktné množiny E^+ (pozitívne príklady) a E^- (negatívne príklady). Platí, že $E \subseteq N_I$ (N_I je mno-



Obr. 1: Schematické znázornenie konceptového učenia.

žina všetkých individuálov v znalostnej báze), kde $E = E^+ \cup E^-$. Cieľom učenia je následne nájsť taký konceptový výraz C, pre ktorý platí $\forall e \in E^+ : \mathcal{K} \models C(e)$ a zároveň $\forall e \in E^- : \mathcal{K} \not\models C(e)$; t.j. výraz, ktorý pokrýva všetky pozitívne príklady a žiadne z negatívnych.

Všeobecne môžeme konceptové učenie definovať taktiež ako prehľadávací proces v množine všetkých konceptov (ktorých je samozrejme nekonečne veľa). Schematické znázornenie konceptového učenia môžeme vidieť na obrázku 1. Jedným zo základných elementov je generátor konceptov, ktorý postupne generuje nové hypotézy (vo forme komplexných konceptových výrazov). Generátor konceptov pri vytváraní nových hypotéz používa automatický

usudzovač, ktorý pracuje nad znalostnou bázou (softvér, ktorý dokáže automaticky odvodzovať nové fakty v ontológii). Kvalita vygenerovaných konceptov je následne ohodnocovaná podľa vybranej metriky. Jednou z metrík môže byť napríklad *pokrytie* pozitívnych a negatívnych príkladov. Na kontrolu *pokrytia* býva zvyčajne opäť využívaný automatický usudzovač. Medzi ďalšie metriky môže patriť napr. informačný zisk, dĺžka konceptu a pod.

2.2 Softvér DL-Learner

V tejto kapitole si popíšeme softvér DL-Learner, ktorý patrí v súčasnosti medzi najpoužívanejší *framework* pre štrukturované strojové učenie v deskripčných logikách [14]. Všetky experimenty uvádzané v tejto práci boli vykonávané prostredníctvom nástroja DL-Learner. Samotný nástroj je implementovaný v jazyku Java, prostredníctvom aplikačného rámca Java Spring. DL-Learner taktiež patrí medzi nástroje s otvoreným zdrojovým kódom[1]. Celkovú architektúru nástroja môžeme vidieť na obrázku 2. Ako môžeme vidieť, architektúru tvorí päť hlavných komponentov:

- *Zdroj znalostí*: tento komponent nástroja definuje, kde sa nachádzajú samotné znalosti a akým spôsobom ich vieme získavať. DL-Learner podporuje všetky známe serializačné formáty pre RD a OWL. Okrem lokálnych znalostných báz, podporuje aj vzdialené získavanie znalostí prostredníctvom SPARQL. V rámci učenia je taktiež možné použiť viacero zdrojov znalostí, pričom je možné kombinovať lokálne a vzdialené zdroje.

[1] https://github.com/SmartDataAnalytics/DL-Learner

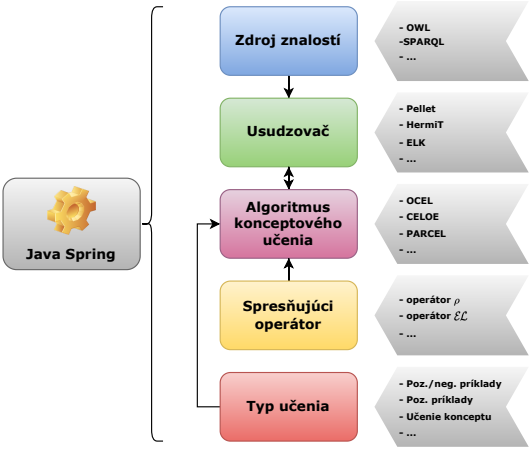
- *Usudzovač*: DL-Learner umožňuje pripojiť sa k usudzovaču prostredníctvom štandardného OWL API, protokolu *OWL-Link* alebo taktiež umožňuje aj priamy prístup v prípade, ak sú potrebné určité pokročilé vlastnosti, ktoré neponúka API rozhranie. Okrem známych usudzovačov, ako sú *Pellet*, *HermiT* alebo *FaCT++*, ponúka DL-Learner aj vlastný tzv. *closed world reasoner*, ktorý pracuje na báze uzavretého sveta (CWA). Tento usudzovač poskytuje aj určité optimalizácie pre overovanie inštancií (jedna z výkonovo najnáročnejších častí konceptového učenia), predpočítaním inferencií a ich uložením v pamäti.

- *Typ učenia*: tento komponent slúži na definíciu problému, ktorý sa snažíme pomocou konceptového učenia riešiť. Učiace algoritmy používajú tento komponent následne na testovanie novo vygenerovaných hypotéz. Ako sme spomínali v kapitole 2.1, existujú tri základné typy (ktoré zároveň aj podporuje DL-Learner): učenie s pozitívnymi a negatívnymi prvkami, učenie iba s pozitívnymi prvkami a učenie konceptu.

- *Spresňujúci operátor*: tento komponent definuje spresňujúci operátor, ktorý sa používa na generovanie konceptov. DL-Learner ponúka dva základné operátory: ρ a špeciálny operátor pre logiku $\mathcal{EL}$. DL-Learner umožňuje taktiež konfigurovať spresňujúci operátor tak, aby zodpovedal konkrétnemu fragmentu OWL (napr. vypnutie/zapnutie negácie, nominálov, operátora kardinality, atď.).

- *Algoritmy konceptového učenia*: sem patria algoritmy, ktoré

implementujú konkrétnu učiacu stratégiu. Patria sem algoritmy ako OCEL, CELOE, PARCEL, ELTL atď.



Obr. 2: Architektúra softvéru DL-Learner.

3 Dosiahnuté výsledky dizertačnej práce

3.1 Ontológia

V rámci práce sme navrhli a implementovali novú ontológiu, ktorá reprezentuje škodlivé súbory pre operačný systém Windows. Okrem samotného návrhu sme publikovali aj množinu datasetov, ktorá môže slúžiť na ďalšie napredovanie výskumu

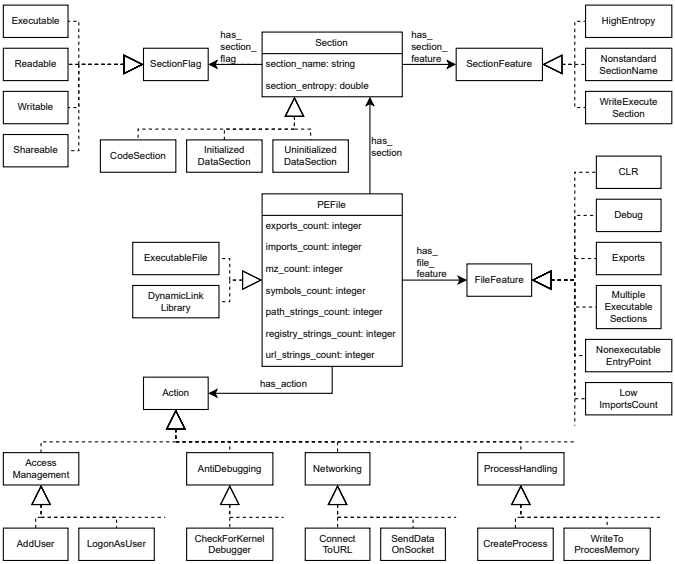
v tejto oblasti. Okrem samotného faktu, že podobná ontológia doposiaľ neexistovala, v porovnaní s inými ontológiami môžeme vyzdvihnúť hlavne jej zameranie na aktuálny vedecký problém spolu s robustnosťou datasetu, ktorá môže byť užitočná pri výskume nových algoritmov. Samotná ontológia nie je limitovaná iba pre konceptové učenie, ale môže byť využitá aj pri ďalších podobných prístupoch. Samotnú ontológiu (resp. znázornenie niektorých základných tried) môžeme vidieť na obrázku 3. Celková ontológia pozostáva zo 195 tried, 6 rolí a 9 dátových vlastností. Čo sa týka ponúkanej expresivity a výpočtovej náročnosti, ontológiu môžeme zaradiť do profilu OWL 2 QL. Tento profil ponúka efektívne overovanie inštancií, čím je vhodným kandidátom pre prípady, keď ontológia obsahuje veľké množstvo individuálov (ako v našom prípade).

Algoritmus	Správnosť	Presnosť	Senzitivita	FP miera	F1 miera
OCEL	0.90 ± 0.00	0.99 ± 0.00	0.85 ± 0.00	0.00 ± 0.00	0.91 ± 0.00
CELOE	0.85 ± 0.00	0.88 ± 0.02	0.88 ± 0.01	0.18 ± 0.03	0.88 ± 0.00
PARCEL	0.92 ± 0.00	0.93 ± 0.00	0.93 ± 0.00	0.11 ± 0.00	0.93 ± 0.00
SPACEL	0.83 ± 0.00	0.97 ± 0.00	0.74 ± 0.02	0.03 ± 0.00	0.84 ± 0.01

Tabuľka 1: Výsledky pre kombinovaný klasifikátor.

3.2 Experimenty

V tejto práci sme taktiež skúmali existujúce algoritmy konceptového učenia (dostupné v rámci softvéru DL-Learner). Konkrétne sme sa venovali ich aplikácii na tvorbu vysvetliteľných modelov pre detekciu škodlivého kódu. Takáto aplikácia konceptového učenia doposiaľ nebola preskúmaná. Ako najlepšie algoritmy z



Obr. 3: Znázornenie základných tried v ontológii.

hľadiska úspešnosti sme označili OCEL a PARCEL, kde najlepšie modely dosiahli F1 mieru 0.91, resp. 0.93 (pri kombinácii rodín). Výsledky je možno vidieť v tabuľke 1. Celkovo sme tak zhodnotili, že konceptové učenie preukazuje vysoký potenciál pri riešení problému detekcie malvéru, keďže metriky pri niektorých prípadoch príliš nezaostávali za štandardným strojovým učením, pričom výsledné modely poskytovali plnú vysvetliteľnosť. Ako zrejmá nevýhoda algoritmov konceptového učenia, však stále zostáva výpočtová náročnosť. Okrem samotnej úspešnosti algoritmov a ich vysvetliteľnosti sme skúmali aj ich bezpečnosť. V rámci týchto experimentov sme zistili, že bezpečnosť konceptových výrazov je porovnateľná s modelmi strojového učenia.

4 Zoznam publikácií dizertanta

- BALOGH, Štefan - **ŠVEC, Peter** - ŠIMKO, Alexander. Expected results of conceptual learning. In Application of Knowledge Methods in Information Security : Smolenice, Slovakia. June 27-29, 2022. 1. vyd. Bratislava : SRDA, 2022, [2] s. ISBN 978-80-974468-0-2.

- BALOGH, Štefan - **ŠVEC, Peter**. Integration of outputs from tools for static and dynamic code analysis into an ontological model. In Application of Knowledge Methods in Information Security : Smolenice, Slovakia. June 27-29, 2022. 1. vyd. Bratislava : SRDA, 2022, [3] s. ISBN 978-80-974468-0-2.

- BISTÁK, Tomáš - **ŠVEC, Peter** - KĽUKA, Ján - ŠIMKO, Alexander - BALOGH, Štefan - HOMOLA, Martin. Im-

proving DL-Learner on a Malware Detection Use Case. In DL 2023 : 36th International Workshop on Description Logics. Rhodes, Greece. September 2-4, 2023. Aachen : CEUR-WS, 2023, [13] s. ISSN 1613-0073 (2022: 0.202 - SJR). V databáze: SCOPUS: 2-s2.0-85176450251.

- PLOSZEK, Roderik - **ŠVEC, Peter** - DEBNÁR, Patrik. Analysis of encryption schemes in modern ransomware. In RAD Hrvatske akademije znanosti i umjetnosti : Matematičke znanosti, Vol 25, No. 546. Zagreb : Hrvatska akademija znanosti i umjetnosti, 2021, S. 1-13. ISSN 1845-4100. V databáze: DOI: 10.21857/mnlqgc58gy ; WOS: 000690973400001 ; SCOPUS: 2-s2.0-85118183775.

  Táto práca bola citovaná v nasledujúcich publikáciach (podľa Google Scholar):

  - BALOGH, Stefan, et al. IoT security challenges: cloud and blockchain, postquantum cryptography, and evolutionary techniques. Electronics, 2021, 10.21: 2647.
  - DAVIES, Simon R.; MACFARLANE, Richard; BUCHANAN, William J. Comparison of entropy calculation methods for ransomware encrypted file identification. Entropy, 2022, 24.10: 1503.
  - WADHO, Shuaib Ahmed, et al. Emerging Ransomware Attacks: Improvement and Remedies-A Systematic Literature Review. In: 2023 4th International Conference on Artificial Intelligence and Data Sciences (AiDAS). IEEE, 2023. p. 148-153.

  - CAÑADAS, Agustín Moreno; MENDEZ, Odette M.; VEGA, Juan David Camacho. Algebraic Structures Induced by the Insertion and Detection of Malware. Computation, 2023, 11.7: 140.
  - DAVIES, Simon R.; MACFARLANE, Richard. Comparison Of Common Mathematical Techniques Used In The Calculation Of File Entropy. In: 2022 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME). IEEE, 2022. p. 1-6.
  - EGGE, Torstein Martinsheimen; JONUZI, Omer; STEMLAND, Åsmund Hunderi. Cloud Backup Architectures Resistant to Ransomware Attacks. 2022. Bachelor's Thesis. NTNU.
  - RAJ, Arpit, et al. Mitigating Ransomware Attacks: A Multi-Tiered Approach to Defence. Available at SSRN 4531839.

- **ŠVEC, Peter**. Generating adversarial malware examples using various build options. In ELITECH´20 [elektronický zdroj] : 22nd Conference of doctoral students. Bratislava, Slovakia. May 27, 2020. 1. ed. Bratislava : Vydavateľstvo Spektrum STU, 2020, [5] s. ISBN 978-80-227-5001-1.

- **ŠVEC, Peter** - BALOGH, Štefan - HOMOLA, Martin. Experimental evaluation of description logic concept learning algorithms for static malware detection. In ICISSP 2021 : 7th International Conference on Information Systems Security and Privacy. Vienna, Austria. February 11-13,

2021. Setúbal : SciTePress - Science and Technology Publications, 2021, S. 792-799. ISSN 2184-4356. ISBN 978-989-758-491-6. V databáze: WOS: 000664076200084 ; SCOPUS: 2-s2.0-85103013157.

Táto práca bola citovaná v nasledujúcich publikáciach (podľa Google Scholar):

- MARTIN, Trevor. On the need for collaborative intelligence in cybersecurity. Electronics, 2022, 11.13: 2067.
- CARDILLO, Franco Alberto; DEBOLE, Franca; STRACCIA, Umberto. PN-OWL: A two stage algorithm to learn fuzzy concept inclusions from OWL ontologies. arXiv preprint arXiv:2303.07192, 2023.
- BALOGH, Štefan. Knowledge and datasets as a resource for improving artificial intelligence. In: Data Science and Intelligent Systems: Proceedings of 5th Computational Methods in Systems and Software 2021, Vol. 2. Springer International Publishing, 2021. p. 828-837.
- MOJŽIŠ, Ján; KENYERES, Martin. Interpretable Rules with a Simplified Data Representation-a Case Study with the EMBER Dataset. In: Proceedings of the Computational Methods in Systems and Software. Cham: Springer International Publishing, 2023. p. 1-10.

- **ŠVEC, Peter** - BALOGH, Štefan - HOMOLA, Martin - KĽUKA, Ján. Ontological representation of the EMBER

dataset. In Application of Knowledge Methods in Information Security : Smolenice, Slovakia. June 27-29, 2022. 1. vyd. Bratislava : SRDA, 2022, [2] s. ISBN 978-80-974468-0-2.

- **ŠVEC, Peter** - JÓKAY, Matúš. Video steganography based on synchronization timestamps. In Mikulášská kryptobesídka 2019 : sborník příspěvků. Praha, Česká republika. 5.-6.12.2019. 1. vyd. Bílovice nad Svitavou : Trusted Network Solutions, 2019, S. 33-34.

- **ŠVEC, Peter** - PLOSZEK, Roderik. A review of encryption schemes used in modern ransomware. In CECC 2020 : Book of abstracts : 20th Central European conference on cryptology. Zagreb, Croatia. June 24-26, 2020. Zagreb : University of Zagreb, 2020, S. 50-51.

- **ŠVEC, Peter** - BALOGH, Štefan. Description logics concept learning in malware detection. In Application of Knowledge Methods in Information Security : Bratislava, Slovakia. September 18, 2021. 1. vyd. Bratislava : SRDA, 2021, [1] s. ISBN 978-80-970145-2.

- **ŠVEC, Peter** - BISTÁK, Tomáš - HOMOLA, Martin - BALOGH, Štefan - KĽUKA, Ján - ŠIMKO, Alexander. Towards Explainable Malware Detection with Structured Machine Learning. In XLoKR 2023 : 4th Workshop on Explainable Logic-Based Knowledge Representation. Rhodes, Greece. September 2-8, 2023. Dresden : CPEC – TRR 248, 2023, [7] s.

- **ŠVEC, Peter** - BALOGH, Štefan - HOMOLA, Martin - KĽUKA, Ján. Knowledge-based dataset fot training PE malware detection models : Technical report. In arXiv.org. no: 2301.00153: 31.12. (2022), [13] s. ISSN 2331-8422.

  Táto práca bola citovaná v nasledujúcich publikáciach (podľa Google Scholar):

  - CARDILLO, Franco Alberto; DEBOLE, Franca; STRACCIA, Umberto. PN-OWL: A two stage algorithm to learn fuzzy concept inclusions from OWL ontologies. arXiv preprint arXiv:2303.07192, 2023.
  - GUPTA, Amjani; SINGH, Dr Karan. Malware Analysis on AI Technique. arXiv preprint arXiv:2311.14501, 2023.
  - BALOGH, Štefan; GALKO, Tibor. Integration of Results from Static and Dynamic Code Analysis into an Ontological Model. In: 2023 IEEE 12th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS). IEEE, 2023. p. 680-685.

- **ŠVEC, Peter** - BALOGH, Štefan - HOMOLA, Martin - KĽUKA, Ján - BISTÁK, Tomáš. Semantic Data Representation for Explainable Windows Malware Detection Models. In arXiv.org. no: 403.11669: 18.03. (2024), [52] s. ISSN 2331-8422. (Poslané do časopisu).

  Táto práca bola citovaná v nasledujúcich publikáciach (podľa Google Scholar):

  - ANTHONY, Peter, et al. Explainable Malware Detection with Tailored Logic Explained Networks. arXiv preprint arXiv:2405.03009, 2024.

Riešiteľ projektov

- Ontologická reprezentácia pre bezpečnosť informačných systémov, APVV19-0220, 2020-2024

Recenzie

- DROZDA, Martin. Krátky úvod do jazyka C++. Bratislava: Vydavateľstvo Spektrum STU, 2020. ISBN 978-80-227-4994-7.

Pedagogická činnosť

- Cvičenia z predmetu Úvod do počítačovej bezpečnosti (2019-2024)
- Cvičenia z predmetu Bezpečnosť informačných systémov z pohľadu praxe (2019-2024)
- Cvičenia z predmetu Operačné systémy (2019)
- Cvičenia z predmetu Programovanie 2 (2020)
- Cvičenia z predmetu Počítačová kriminalita (2021)
- Pozvané prednášky v rámci predmetu Úvod do počítačovej bezpečnosti a projektu MiniErasmus pre stredoškolákov
- Vedúci 16 obhájených bakalárskych prác a 2 tímových projektov

Zoznam použitej literatúry

1. *Malware Statistics Trends Report*. [B.r.]. Dostupné tiež z: `https://www.av-test.org/en/statistics/malware/`. [Online; cit. 2023-10-21].
2. ŠVEC, Peter, BALOGH, Štefan, HOMOLA, Martin a KL'UKA, Ján. Ontological Representation of the EMBER Dataset. *Application of Knowledge Methods in Information Security*. 2022, s. 3.
3. ŠVEC, Peter, BALOGH, Štefan, HOMOLA, Martin a KL'UKA, Ján. Knowledge-Based Dataset for Training PE Malware Detection Models. *arXiv preprint arXiv:2301.00153*. 2022.
4. ŠVEC, Peter, BALOGH, Štefan, HOMOLA, Martin, KL'UKA, Ján a BISTÁK, Tomáš. Semantic Data Representation for Explainable Windows Malware Detection Models. *arXiv preprint arXiv:2403.11669*. 2024.
5. ŠVEC, Peter a BALOGH. Description Logics Concept Learning in Malware Detection. *Application of Knowledge Methods in Information Security*. 2021, s. 1.
6. ŠVEC, Peter, BALOGH, Štefan a HOMOLA, Martin. Experimental Evaluation of Description Logic Concept Learning Algorithms for Static Malware Detection. In: *ICISSP*. 2021, s. 792–799.
7. ŠVEC, Peter, BISTÁK, Tomáš, HOMOLA, Martin, BALOGH, Štefan, KĽUKA, Ján a ŠIMKO, Alexander. Towards Explainable Malware Detection with Structured Ma-

chine Learning. In: *Workshop on Explainable Logic-Based Knowledge Representation XLoKR 2023*. 2023.

8. BISTÁK, Tomáš, ŠVEC, Peter, KĽUKA, Ján, ŠIMKO, Alexander, BALOGH, Štefan a HOMOLA, Martin. Improving DL-Learner on a Malware Detection Use Case. In: *36th International Workshop on Description Logics, DL 2023*. CEUR-WS, 2023.
9. LEHMANN, Jens a HITZLER, Pascal. Concept learning in description logics using refinement operators. *Machine Learning*. 2010, roč. 78, č. 1, s. 203–250.
10. FUNK, Maurice, JUNG, Jean Christoph, LUTZ, Carsten, PULCINI, Hadrien a WOLTER, Frank. Learning Description Logic Concepts: When can Positive and Negative Examples be Separated? In: KRAUS, Sarit (ed.). *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI 2019, Macao, China, August 10-16, 2019*. 2019, s. 1682–1688. Dostupné z DOI: 10.24963/ijcai.2019/233.
11. D'AMATO, Claudia, FANIZZI, Nicola a ESPOSITO, Floriana. Inductive learning for the semantic web: what does it buy? *Semantic Web*. 2010, roč. 1, č. 1-2, s. 53–59.
12. LAVRAC, Nada a DZEROSKI, Saso. Inductive Logic Programming. In: *WLP*. Springer, 1994, s. 146–160.
13. LEHMANN, Jens. *Learning OWL class expressions*. Zv. 22. IOS Press, 2010.

14. BÜHMANN, Lorenz, LEHMANN, Jens a WESTPHAL, Patrick. DL-Learner—A framework for inductive learning on the Semantic Web. *Journal of Web Semantics*. 2016, roč. 39, s. 15–24.