FORMÁLNA KONCEPTOVÁ ANALÝZA V KYBERBEZPEČNOSTI

UNIVERZITA PAVLA JOZEFA ŠAFÁRIKA V KOŠICIACH

Financované Európskou úniou NextGenerationEU

PLÁN OBNOVY

MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

Projekt Kompetenčné centrum kybernetickej bezpečnosti na UPJŠ (KCKB UPJŠ) financovaný Európskou úniou z prostriedkov Plánu obnovy a odolnosti Slovenskej republiky, kód projektu: 17R05-04-V01-00007.

Formálna konceptová analýza (FCA)

- Rudolf Wille, 1981
- Matematický rámec na štruktúrovanie a vyhľadávanie informácií
- Založená na teórii zväzov a usporiadaných množín
- Modeluje vzťahy medzi objektami a atribútmi pomocou formálneho kontextu
- Identifikuje skryté formálne koncepty v dátach
- Organizuje ich do hierarchickej štruktúry (konceptový zväz)
- Umožňuje formálne odvodenie závislostí a pravidiel z dát (atribútové implikácie)
- Uplatnenie v rôznych aplikačných doménach, vrátane kyberbezpečnosti

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

Základné definície FCA

Formálny kontext je trojica $\langle B, A, I \rangle$ pozostávajúca z dvoch množín B, A a binárnej relácie I medzi B a A, teda $I \subseteq B \times A$. Prvky množiny B nazývame objekty a prvky množiny A nazývame atribúty daného kontextu.

	i	ii	iii
a	×		
b	×	×	
c		×	

Nech $\langle B, A, I \rangle$ je formálny kontext a $X \in P(B), Y \in P(A)$ sú podmnožiny množiny objektov, resp. atribútov. Definujme dvojicu zobrazení $\uparrow: P(B) \to P(A)$ a $\downarrow: P(A) \to P(B)$ nasledovne:

$$\uparrow (X) = \{y \in A : (\forall x \in X)\langle x, y \rangle \in I\},$$

$$\downarrow (Y) = \{x \in B : (\forall y \in Y)\langle x, y \rangle \in I\}.$$

Dané zobrazenia nazývame **derivačné operátory** formálneho kontextu.

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

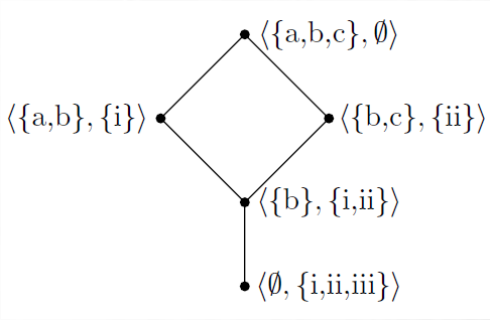
Základné definície FCA

Nech $\langle B, A, I \rangle$ je formálny kontext a nech $\langle \uparrow, \downarrow \rangle$ sú jeho derivačné operátory. Uvažujme podmnožiny množín objektov a atribútov $X \in P(B), Y \in P(A)$. Dvojica $\langle X , Y \rangle$ pre ktorú platí $\uparrow (X) = Y , \downarrow (Y) = X$ sa nazýva **formálny koncept** formálneho kontextu $\langle B, A, I \rangle$.

Podmnožinu objektov X nazývame extentom a podmnožinu atribútov Y intentom tohto formálneho konceptu.

Nech $\langle X_1, Y_1 \rangle, \langle X_2, Y_2 \rangle$ sú formálne koncepty formálneho kontextu $\langle B, A, I \rangle$. Uvažujme čiastočné usporiadanie $\preceq$ také, že $\langle X_1, Y_1 \rangle \preceq \langle X_2, Y_2 \rangle$ práve vtedy, keď $X_1 \subseteq X_2$. Čiastočne usporiadaná množina $\langle \mathfrak{B}(B, A, I), \preceq \rangle$ sa nazýva **konceptový zväz** formálneho kontextu $\langle B, A, I \rangle$.

Konceptový zväz formálneho kontextu $\langle B, A, I \rangle$ označujeme $\mathfrak{B}(B, A, I)$.



Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

Aplikácie FCA

Aplikácie FCA sa v kyberbezpečnosti vyskytujú v oblastiach:

- Digitálna forenzná analýza
- Detekcia podvodov
- Analýza malvéru
- Modelovanie hrozieb
- A ďalšie...

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

Digitálna forenzná analýza

Financované Európskou úniou
NextGenerationEU

PLÁN OBNOVY

MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

The analysis of digital evidence by Formal concept analysis

P. Sokol, Ľ. Antoni, O. Krídlo, E. Marková, K. Kováčová, S. Krajči (2022)

Motivácia:

- Rastúci počet útokov – analytici sú zahltení alertami
- Potreba rýchlo pochopiť „čo sa deje“ a nájsť relevantné digitálne dôkazy, v praxi pritom stále prevažujú manuálne postupy
- Využívajú sa metadáta (napr. veľkosť súboru, cesta k súboru), analýza časovej osi (chronologické usporiadanie záznamov pred/počas/po incidente) a vyhľadávanie anomálií
- Cieľom je redukovať čas a úsilie potrebné na vykonanie digitálnej forenznej analýzy

Výskumné ciele:

- Hľadanie vzťahov medzi atribútmi digitálnych dôkazov na časovej osi
- Identifikácia anomálnych záznamov na časovej osi

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

The analysis of digital evidence by Formal concept analysis

P. Sokol, Ľ. Antoni, O. Krídlo, E. Marková, K. Kováčová, S. Krajči (2022)

Dataset:

- Windows + NTFS (New Technology File System)
- DFIR Madness portal – Case001: The Stolen Szechuan Sauce (DC01 Disk Image (E01))

Predspracovanie dát:

- Časová os: 1 263 787 záznamov
- Na analýzu boli použité len záznamy so zdrojom file (filesystem): 843 863 záznamov

Využitie FCA:

- Skúmanie zmysluplných zoskupení digitálnych objektov vzhľadom na spoločné atribúty + vizualizácia v konceptovom zväze
- Asociačné pravidlá na identifikáciu štandardného správania OS + výnimiek

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

The analysis of digital evidence by Formal concept analysis

P. Sokol, Ľ. Antoni, O. Krídlo, E. Marková, K. Kováčová, S. Krajči (2022)

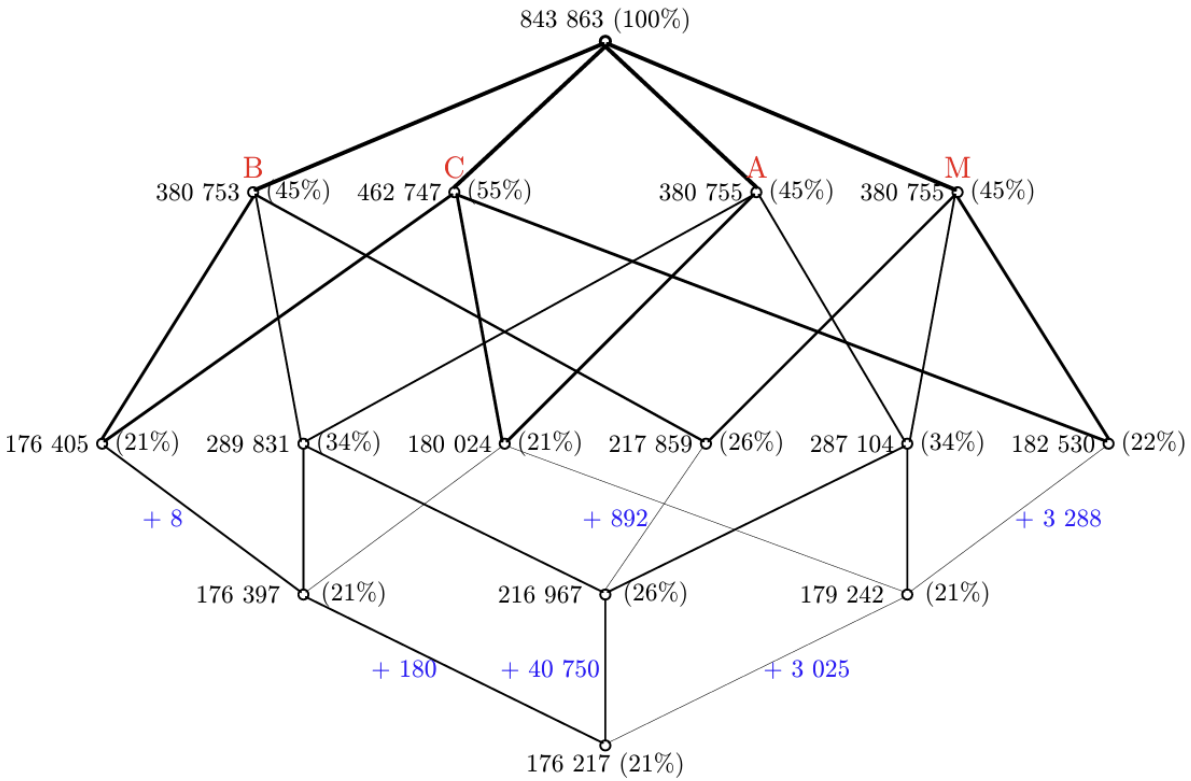


Figure 1: Concept lattice of MACB.

Konceptový zväz pre časové pečiatky (MACB atribúty)

- 15 konceptov
- Pri každom koncepte: počet objektov v extente a percento zo všetkých objektov
- Analýza časových pečiatok pomáha pochopiť aké operácie boli vykonané nad súborom

Konceptový zväz umožňuje vidieť **typické** aj **zriedkavé kombinácie atribútov** (potenciálne anomálie)

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

Zaujímavé zistenia z konceptového zväzu (príklad):

- 21% objektov patrí do konceptu, kde sú prítomné všetky 4 MACB atribúty (= vytvorenie súboru)
- Koncept s intentom BCM sa vo zväze nenachádza
- C je najčastejšia časová pečiatka (55%)

Výskum sa ďalej zameriava na vzťah atribútov MACB k atribútom opisujúcim cestu, typ súboru a veľkosť súboru vytvorením nového konceptového zväzu

- Zistený príklad **anomálie**: kombinácia atribútov ukázala, že v adresári bol vytvorený iba jeden spustiteľný súbor

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

The analysis of digital evidence by Formal concept analysis

P. Sokol, Ľ. Antoni, O. Krídlo, E. Marková, K. Kováčová, S. Krajči (2022)

Table 4
Selected association rules for all attributes with confidence less than 100%

Association rule	Support	Confidence
{C, B, size_none} → {A}	175 441	99,99%
{dir_win} → {file_stat}	313 053	99,99%
{A, file_stat, directory, dir_win, size_none} → {M}	19 558	99,95%
{A, dir_other, size_none} → {NTFS_file_stat, mft}	268 996	99,88%
{C, B, dir_user} → {M, A}	24	96,00%
{file_stat, filef, file_executable} → {dir_win}	125	96,00%
{C, file_stat, filef, file_executable} → {dir_win}	41	95,00%
{A, file_stat, filef, file_executable} → {dir_win}	41	95,00%
{size_Q4, file_stat, filef, file_executable} → {dir_win}	90	94,00%

Asociačné pravidlá:

- Confidence 100% – identifikácia štandardov – záznamy, ktorým nie je potrebné venovať pozornosť
- Confidence blízka 100% – skúmanie výnimiek

2 skupiny výnimiek:

- Výnimky spôsobené známymi špecifikami OS Windows
- Výnimky reprezentujúce potenciálne anomálie, ktoré treba riešiť

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

Formal concept analysis approach to understand digital evidence relationships

P. Sokol, Ľ. Antoni, O. Krídlo, E. Marková, K. Kováčová, S. Krajči (2023)

Rozšírenie predošlého výskumu (2022):

- O ďalší typ zdroja digitálnych dôkazov (event logs)
- Do fuzzy FCA

Dataset:

- DFIR Madness portal – Case001: The Stolen Szechuan Sauce (DC01 Disk Image (E01))
- Filesystem dataset (843 863 záznamov) + Event log dataset (86 180 záznamov) podľa zdroja
- (Celkom použité 4 datasety s rôznymi množinami atribútov)

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

Formal concept analysis approach to understand digital evidence relationships

P. Sokol, Ľ. Antoni, O. Krídlo, E. Marková, K. Kováčová, S. Krajči (2023)

Využitie FCA:

- Identifikácia podozrivých záznamov špecifických pre event logs, súborový systém NTFS, operačný systém Windows alebo určitý typ anomálie
- Skúmanie závislostí medzi atribútmi digitálnych dôkazov, ktoré môžu pomôcť pri identifikácii vzťahov medzi rôznymi časťami dôkazov a potenciálne odhaliť nové poznatky
- Identifikácia relevantných záznamov na ďalšiu kontrolu forenznými analytikmi
- Poskytovanie upozornení pre forenzných analytikov
- Identifikácia vhodných časových období na detailné preskúmanie forenznými analytikmi

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

Formal concept analysis approach to understand digital evidence relationships

P. Sokol, Ľ. Antoni, O. Krídlo, E. Marková, K. Kováčová, S. Krajči (2023)

Konceptové zväzy

Autori formulujú **2 všeobecné odporúčania pre bezpečnostných analytikov** týkajúce sa vlastných atribútov:

- Skontrolovať vzťahy medzi atribútmi na základe štruktúry vlastných atribútov (všeobecné atribúty bližšie k vrcholu zväzu, špecifické hlbšie vo zväze)
- Skontrolovať koncepty s najmenším počtom vlastných objektov (podozrivé/zriedkavé záznamy bez manuálneho prehľadávania celého datasetu)

Asociačné pravidlá

Autori navrhujú kategorizáciu asociačných pravidiel do 3 skupín na základe hodnoty confidence:

- Confidence 100% = očakávané správanie
- Confidence [90%, 100%) = podozrivé správanie (najzaujímavejšie pre forenznú analýzu)
- Confidence [50 %, 90 %) = nezaujímavé správanie

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

Formal concept analysis approach to understand digital evidence relationships

P. Sokol, Ľ. Antoni, O. Krídlo, E. Marková, K. Kováčová, S. Krajči (2023)

Fuzzy atribútové implikácie

- Analýza event logs aj s časovým aspektom – 3 fuzzy atribúty (epochtime, hour, minute)
- Výsledné fuzzy atribútové implikácie možno interpretovať ako špecifické kombinácie atribútov v konkrétnom časovom období
- Tieto informácie môžu pomôcť bezpečnostným analytikom nájsť vhodné časové intervaly na forenznú analýzu

Fuzzy FCA

- Fuzzy rozšírenie FCA je možné použiť aj na generovanie fuzzy formálnych konceptov
- Ak je počet fuzzy konceptov vo zväze príliš veľký na manuálnu kontrolu analytikom, autori navrhnujú redukciu presnejšou špecifikáciou podmienok na atribúty (resp. objekty) alebo vytvorením podzväzu konceptov so support > určitý prah

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

Towards a granular computing approach based on Formal Concept Analysis for discovering periodicities in data

V. Loia, F. Orciuoli, W. Pedrycz (2018)

Motivácia:

- Analýza výskytov a spoločných výskytov udalostí môže viesť k významným poznatkom v oblasti verejnej bezpečnosti

V kontexte **digitálnej forenznej analýzy** je užitočné:

- Zostrojiť časovú os podozrivej osoby (napr. zo sociálnych sietí)
- Identifikovať skryté vzťahy medzi osobami (napr. osoby v rovnakom čase na rovnakom mieste)
- Odhaliť behaviorálne vzory prostredníctvom periodických aktivít

Výskumné ciele:

- Autori sa pozerajú na **temporálne dáta**, ich cieľom je hľadanie vzorov, spoločných výskytov udalostí a periodicít, pričom je potrebné analyzovať dáta z rôznych časových pohľadov (multiple time-related views) a vedieť tieto pohľady hodnotiť, aby sa vybrali tie relevantné

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

Towards a granular computing approach based on Formal Concept Analysis for discovering periodicities in data

V. Loia, F. Orciuoli, W. Pedrycz (2018)

Využitie FCA:

- Vytvárajú sa formálne kontexty, kde objekty sú udalosti a atribúty sú časového charakteru
- Využíva sa Granular Computing – nepozerá sa na jednotlivé udalosti, ale na zmysluplné skupiny udalostí – granuly (Timed Information Granules) získané ako formálne koncepty
- Viaceré pohľady na dáta sa získavajú viacerými granuláciami (time-guided granulation)
- Na výber zaujímavých a relevantných granúl sa zavádzajú miery hodnotenia ich kvality

Dataset:

- Lesné požiare (Montesinho Natural Park, Portugalsko) – avšak metodológiu je možné aplikovať aj na temporálne dáta z digitálnej forenznej analýzy
- Výsledky môžu byť použité na predikciu lesných požiarov v danej lokalite

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

Towards a granular computing approach based on Formal Concept Analysis for discovering periodicities in data

V. Loia, F. Orciuoli, W. Pedrycz (2018)

Table 1
Sample formal context.

	Rome	Milan	Monday	Tusday	Wednesday	Thursday	Friday
e1_1	x		x				
e1_2	x		x				
e1_3	x		x				
e1_4	x		x				
e2_1		x		x			
e2_2	x		x				
e2_3		x					x
e2_4		x		x			

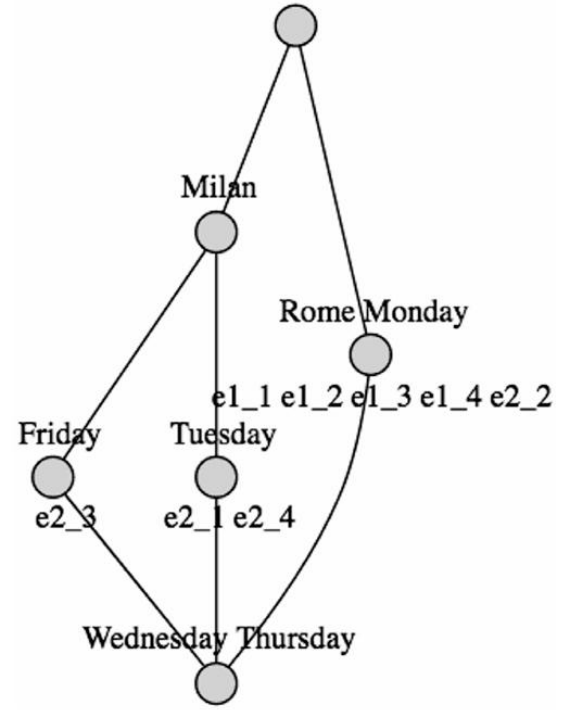


Fig. 2. Sample formal context.

Formálny kontext a konceptový zväz (príklad)

- Formálne koncepty = časované informačné granuly, ktoré reprezentujú periodické výskyty alebo spoločné výskyty udalostí

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

Towards a granular computing approach based on Formal Concept Analysis for discovering periodicities in data

V. Loia, F. Orciuoli, W. Pedrycz (2018)

Kvantitatívne miery vyhodnotenia kvality granúl a časových pohľadov:

- Information Granularity (IG) – meria úroveň granularity výsledných granúl, teda do akej miery je výsledok príliš hrubý (všeobecný), alebo príliš jemný (fragmentovaný)
- Information Entropy (IE) – meria entropiu/informačnú neurčitosť v rozložení granúl, teda či granularizácia vedie k „informatívnejšej“ štruktúre alebo naopak k neprehľadnej/chaotickej štruktúre
- Separation (SEP) – meria separáciu granúl, teda či granuly predstavujú dostatočne odlíšiteľné skupiny
- Coverage (COV) – meria pokrytie, teda do akej miery granuly pokrývajú dáta alebo významné časti dát
- Specificity (SP) – meria špecifickosť granúl, teda či granula zachytáva konkrétny a jasne definovaný vzor (nie príliš všeobecný)

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

Detekcia podvodov

Financované Európskou úniou
NextGenerationEU

PLÁN OBNOVY

MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

Formal concept analysis with negative attributes for forgery detection

M. Ojeda-Aciego, J. M. Rodríguez-Jiménez (2013)

Motivácia:

- Problém falšovania dokumentov
- Pri policajných kontrolách sa kontrolujú doklady z rôznych krajín, pričom neexistuje jednotná medzinárodná databáza, ktorá by umožňovala overiť pravosť dokumentov
- Kľúčová je rýchla identifikácia falzifikátu na základe dostupných znakov a bezpečnostných prvkov dokumentu

Výskumné ciele:

- Rozpoznávanie „podpisov“ falšovateľov (counterfeiter signature) v dokumentoch a hľadanie vzťahov medzi „podpismi“ s cieľom získať informácie o aktivitách falšovateľov
- Výskum sa zameriava na absenciu vlastností/atribútov – pri falšovaní nie je relevantné iba to, čo falšovateľ napodobnil, ale aj to, čo nedokázal napodobniť (= „podpis“ faľšovateľa)

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

Formal concept analysis with negative attributes for forgery detection

M. Ojeda-Aciego, J. M. Rodríguez-Jiménez (2013)

Využitie FCA:

- Binárna reprezentácia atribútov + zameranie na **negatívne atribúty** (atypické pre FCA)
- Získanie štruktúrovaných poznatkov o podobnostiach a vzťahoch medzi falzifikátmi (objektami) na základe ich atribútov
- Organizácia poznatkov v **zmiešanom konceptovom zväze** s pozitívnymi aj negatívnymi atribútmi

Dataset:

- Falšované talianske vodičské preukazy (36 dokumentov)
- Falšované rumunské občianske preukazy

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

Formal concept analysis with negative attributes for forgery detection

M. Ojeda-Aciego, J. M. Rodríguez-Jiménez (2013)

TABLE 1 Example of formal context

	m_1	m_2	m_3	m_4
g_1	1	0	0	1
g_2	1	1	0	0
g_3	1	0	1	1
g_4	0	1	0	1

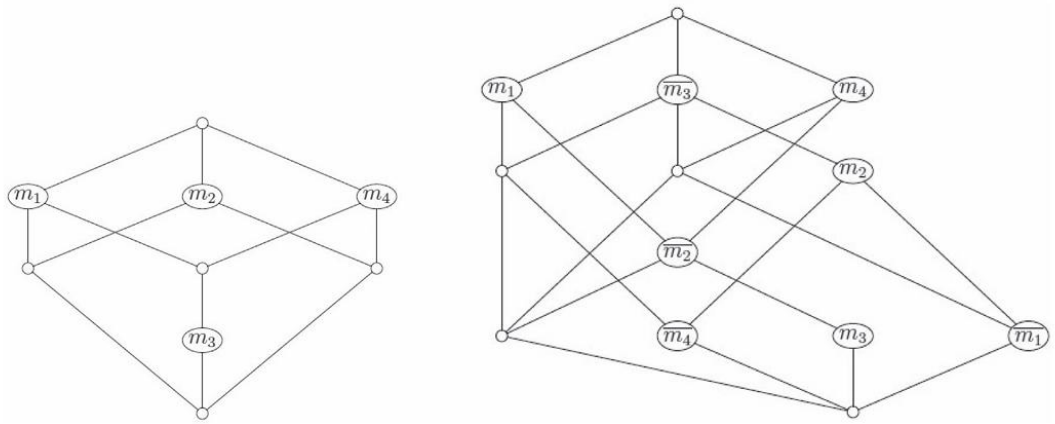


FIGURE 2 Concept lattice and mixed concept lattice from Table 1

Formálny kontext, konceptový zväz a zmiešaný konceptový zväz (príklad)

Zmiešaný konceptový zväz:

- Vznikol pridaním **negácie** každého atribútu do formálneho kontextu
- Umožňuje rozlišovať rôzne podpisy falšovateľov a pri nových falzifikátoch rozhodnúť, či zodpovedajú existujúcim podpisom alebo vytvárajú nové vzory

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

Formal concept analysis with negative attributes for forgery detection

M. Ojeda-Aciego, J. M. Rodríguez-Jiménez (2013)

Autori zavádzajú miery na hodnotenie a výber relevantných atribútov a objektov (odporúčanie pre políciu):

- Presnosť atribútu (accuracy) – na určenie bezpečnostných prvkov, ktoré majú najvyššiu informačnú hodnotu pri rozlišovaní falzifikátov a mali by byť prioritizované pri kontrole
- Správnosť objektu (correctness) – na identifikáciu falzifikátov, ktoré sa vymykajú očakávaným vzorom (anomálie)

Navrhnutá metóda dokáže zefektívniť kontrolu dokladov tým, že umožní nájsť najinformatívnejšie bezpečnostné prvky a znížiť tak počet tých, ktoré je potrebné fyzicky kontrolovať

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

E-fraud forensics investigation techniques with formal concept analysis

W. V. Onomza, A. Umar, M. Olalere (2014)

Motivácia:

- Organizácie sú v kyberpriestore terčom podvodných/trestných aktivít
- Páchateľmi môžu byť jednotlivci alebo organizované skupiny z rôznych geografických území čo sťažuje vyšetrovanie podvodov

Výskumné ciele:

- Aplikácia FCA na štruktúrovanie a reprezentáciu podvodných aktivít získaných z digitálnych zariadení
- Podpora detekcie a vyhľadávania podvodných aktivít v kyberpriestore

Využitie FCA:

- Konceptový zväz ako vizualizačný a analytický nástroj, ktorý zachytáva spoločné a rozdielne atribúty digitálnych stôp a umožňuje analyzovať ich vzájomné súvislosti

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

Social Engineering Cybercrime Evidence Analysis Using Formal Concept Analysis

I. B. Senkyire, Q.-A. Kester (2021)

Motivácia:

- Sociálne inžinierstvo predstavuje jednu z hlavných hrozieb v kybernetickej bezpečnosti (tvorí až 84% kyberútokov)

Výskumné ciele:

- Analýza dôkazov z kybernetických trestných činov súvisiacich s phishingovými útokmi – konkrétne zo správ (e-mail/SMS)

Použitie FCA:

- Štruktúrovanie digitálnych dôkazov a odhaľovanie vzťahov medzi prípadmi a ich vlastnosťami pomocou konceptového zväzu

Dataset:

- Phishingové správy + extrakcia charakteristických znakov, napr. pocit naliehavosti, vydávanie sa za známeho odosielateľa, aktualizácia údajov účtu, kliknutie na odkaz v e-mailovej správe, ...

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

Detecting phishing websites by using a hybrid method of Page Content and Formal Concept Analysis

S. Chen, H. Wu, X. Cheng, L. Mao (2023)

Motivácia:

- S narastajúcim počtom phishingových webových stránok sa otázka ich rozpoznania stala intenzívne skúmanou témou
- Len málo výskumov je zameraných na phishingové weby patriace ku konkrétnemu typu webovej stránky

Výskumný cieľ:

- Detekcia phishingových webov, ktoré napodobňujú konkrétne známe weby (napr. banky, e-commerce)

Metodológia:

- Hybridná metóda kombinuje Page Content (TF-IDF) a FCA
- Navrhnutá metóda robí na danej webovej stránke TF-IDF analýzu obsahu a extrahuje slová, ktoré ju reprezentujú – na tomto základe sa potom konštruuje konceptový zväz

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

Detecting phishing websites by using a hybrid method of Page Content and Formal Concept Analysis

S. Chen, H. Wu, X. Cheng, L. Mao (2023)

- **TF-IDF analýza obsahu webovej stránky** slúži na výber kľúčových slov/fráz charakterizujúcich daný web (TF = term frequency, IDF = inverse document frequency)
- V konceptovom zväze sú objekty webové stránky a atribúty extrahované frázy
- Pre (nový) podozrivý web sa vykoná TF-IDF analýza obsahu a frázy sa porovnávajú s konceptami vo zväze + porovnanie domény (rovnaká doména – legitímny web, iná – phishing)

Dataset:

- Top 10 bankových webov v Číne + 20 phishingových webov

Experiment ukazuje, že táto metóda dokáže efektívne identifikovať phishingové webové stránky (dosiahnutá presnosť (accuracy) 95%)

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

Analýza malvéru

Financované Európskou úniou
NextGenerationEU

PLÁN OBNOVY

MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

MarCHGen: A framework for generating a malware concept hierarchy

T. B. Nguyen, C. D. Tran, T. T. Quan, M. H. Nguyen, A. T. Le (2019)

Motivácia:

- Manuálne vytváranie hierarchií malvéru už nie je reálne kvôli veľkému počtu variantov
- Klasické prístupy založené na atribútoch alebo signatúrach zlyhávajú (polymorfné, metamorfické alebo obfuskované vzorky)
- Malvér nie je vhodné opisovať jeho atribútmi, ale správaním
- Temporálna logika umožňuje zachytiť sekvenčný charakter vykonávania inštrukcií

Výskumné ciele:

- Automatické generovanie hierarchie malvéru založené na správaní malvéru z veľkých a dynamicky sa meniacich datasetov
- Hierarchia ľahko interpretovateľná pre bezpečnostných analytikov

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

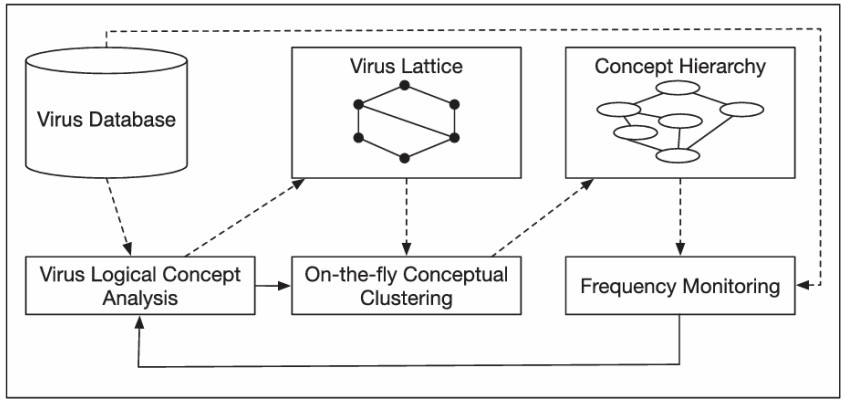
MarCHGen: A framework for generating a malware concept hierarchy

T. B. Nguyen, C. D. Tran, T. T. Quan, M. H. Nguyen, A. T. Le (2019)

Metodológia:

Rámec **MarCHGen** (Malware Concept Hierarchy Generation):

- Rozšírenie FCA na **V-LCA** (Virus Logical Concept Analysis), kde objekty – vzorky malvéru, atribúty – temporálne logické formuly, koncepty – zovšeobecnené správanie skupiny vírusov
- Konceptové zhlukovanie **OCC** (On-the-Fly Conceptual Clustering) – koncepty vznikajú priebežne počas generovania zväzu
- **Monitorovanie frekventovaných konceptov** – hierarchia sa priebežne optimalizuje – uvažuje iba koncepty, ktoré sú dostatočné časté (na základe zvoleného prahu)
  → redukcia konceptového zväzu



MarCHGen rámec

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

MarCHGen: A framework for generating a malware concept hierarchy

T. B. Nguyen, C. D. Tran, T. T. Quan, M. H. Nguyen, A. T. Le (2019)

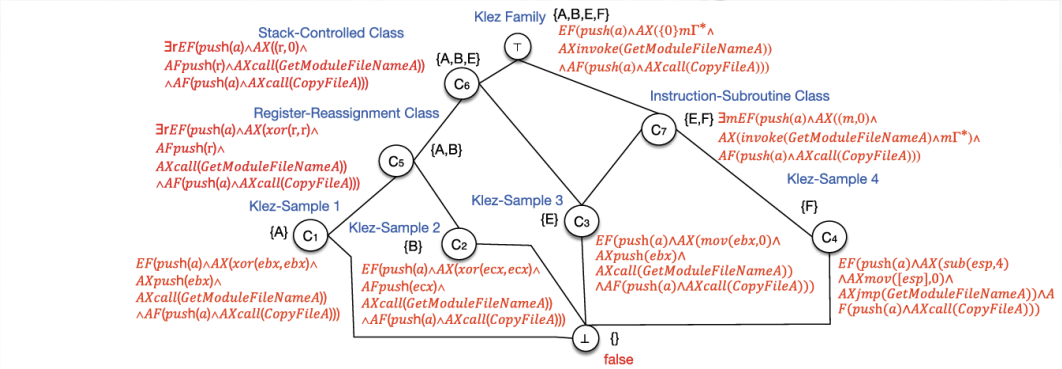


FIGURE 4 The concept lattice generated by virus logical concept analysis

TABLE 2 Some code segments illustrating the virus behaviours

ID	Sample pattern	Logic formulas	Meaning
A	*pusha*		Klez worm
	xorebx, ebx		
	pushebx	$EF(push(a) \wedge AX(xor(ebx, ebx) \wedge AXpush(ebx) \wedge$	
	callGetModuleFileNameA	$AXcall(GetModuleFileNameA))$	
	pusha	$\wedge AF(push(a) \wedge AXcall(CopyFileA)))$	
	callCopyFileA		
B			
	pusha		
	xorecx, ecx		
	decebx	$EF(push(a) \wedge AX(xor(ecx, ecx) \wedge AFpush(ecx) \wedge$	Klez variant with
	pushecx	$AXcall(GetModuleFileNameA))$	register reassignment
	callGetModuleFileNameA	$\wedge AF(push(a) \wedge AXcall(CopyFileA)))$	and junk code
	pusha		
	callCopyFileA		

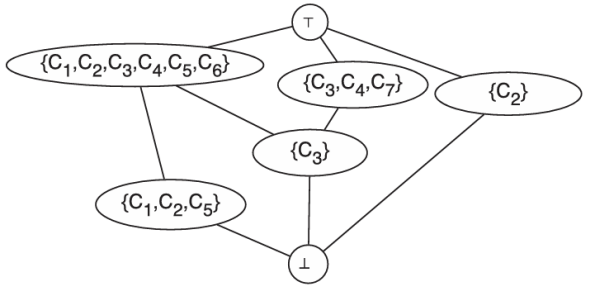


FIGURE 5 The concept hierarchy generated by on-the-fly conceptual clustering

Financované Európskou úniou NextGenerationEU

PLÁN OBNOVY

MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

MarCHGen: A framework for generating a malware concept hierarchy

T. B. Nguyen, C. D. Tran, T. T. Quan, M. H. Nguyen, A. T. Le (2019)

Dataset:

- 5 000 vzoriek malvéru z databáz (VXHeaven,VirusShare, MALICIA)
- Dáta sú do datasetu pridávané priebežne

Navrhnutý prístup **Frequent OCC** je porovnaný s inými prístupmi (FFCA, OCC, HAC (Hierarchical Agglomerative Clustering), k-means):

- Najefektívnejší z hľadiska výkonu
- Dosahuje najlepšie výsledky z pohľadu hodnotenia kvality hierarchie a hodnotenia kvality zhlukov

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

A Formal Concept Analysis approach to hierarchical description of malware threats

M. Ojeda-Hernández, D. López-Rodríguez, Á. Mora (2024)

Motivácia:

- Problém nejednotného názvoslovia malvéru v rôznych antivírusových systémoch
- Rôzne identifikátory od rôznych antivírusových výrobcov môžu označovať tú istú hrozbu

Výskumné ciele:

- Konštrukcia hierarchie malvéru, ktorá je nezávislá od konkrétneho antivírusového výrobcu
- Modelovanie závislostí medzi kategóriami malvéru

Využitie FCA:

- Konceptový zväz ako nástroj na analýzu vzťahov medzi malvérovými hrozbami a ich identifikátormi v rôznych antivírusových systémoch
- Objekty – malvérové súbory, atribúty – identifikátory malvéru ako dvojice ⟨antivirus, label⟩

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

A Formal Concept Analysis approach to hierarchical description of malware threats

M. Ojeda-Hernández, D. López-Rodríguez, Á. Mora (2024)

Dataset:

- 183 JSON súborov z platformy VirusTotal, ktorá agreguje výsledky detekcie malvéru z rôznych antivírusových systémov

Zistenia z konceptového zväzu:

- Identifikované vzťahy generalizácie a špecializácie medzi malvérmi
- Identifikované ekvivalentné identifikátory, ktoré označujú identickú množinu súborov
- Niektoré malvérové kategórie sú presne charakterizovateľné ako prienik iných kategórií

Výsledky môžu slúžiť ako podklad pre štandardizáciu názvoslovia

$$(A, B) \quad \text{with} \quad \begin{cases} A = \{\text{cd581c0251f2b1e7559c4b5830.json}, \\ \quad \text{fb1bd5c6486f120268c0803901.json}, \\ \quad \text{44ed7c95e37adfa1e90cc55847.json}\} \\ B = \{\langle \text{Kaspersky, not-a-virus:heur:.os.adlo.b}\rangle, \\ \quad \langle \text{Avast-Mobile, :evo-gen}\rangle, \\ \quad \langle \text{Avira, /agent.mer.gen}\rangle, \\ \quad \langle \text{Microsoft, :script/wacatac.b!ml}\rangle\} \end{cases}$$

formálny koncept (príklad)

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

Feature-Driven Formal Concept Analysis for Malware Hierarchy Construction

T. B. Nguyen, C. D. Tran, T. T. Quan, M. H. Nguyen (2015)

Motivácia:

- Podobne ako v predošlých výskumoch – automatická klasifikácia malvéru je kľúčovým problémom vo výskume malvéru
- Novšie prístupy sa zameriavajú na správanie malvéru namiesto signatúr – využívajú temporálnu logiku

Výskumný cieľ:

- Generovanie hierarchie malvéru založené na správaní malvéru

Metodológia:

- Rozšírenie FCA na F-FCA (Feature-Driven FCA) – atribúty ako logické formuly
- Návrh algoritmu FOCA (Feature-Driven On-the-Fly Conceptual Clustering) – efektívne generovanie hierarchie pre veľké datasety

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

Feature-Driven Formal Concept Analysis for Malware Hierarchy Construction

T. B. Nguyen, C. D. Tran, T. T. Quan, M. H. Nguyen (2015)

Porovnanie konceptových zväzov FCA a F-FCA na rodine vírusu Avron

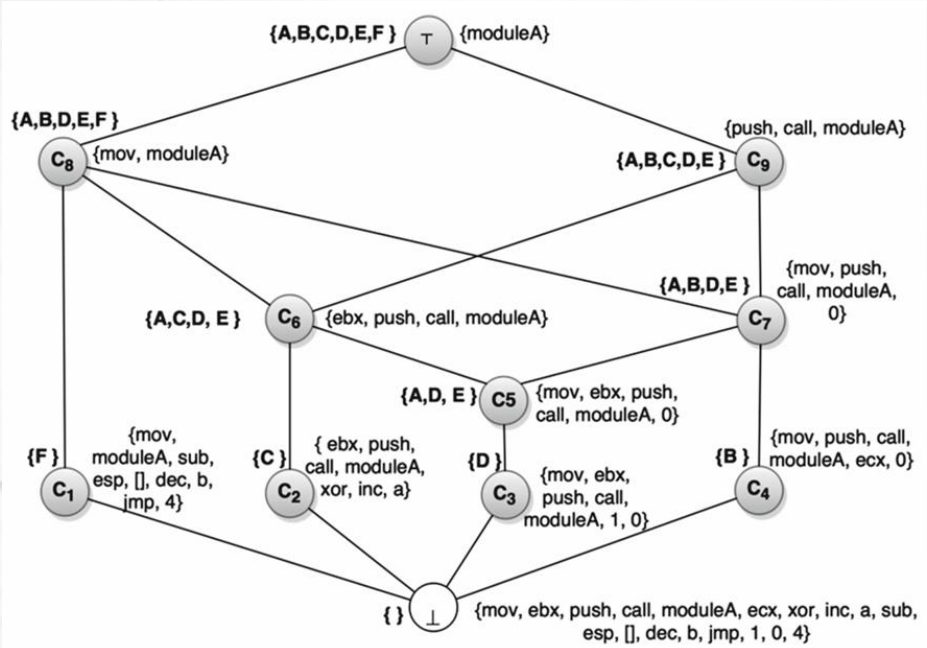


Fig. 2. The concept lattice of virus is generated by FCA method

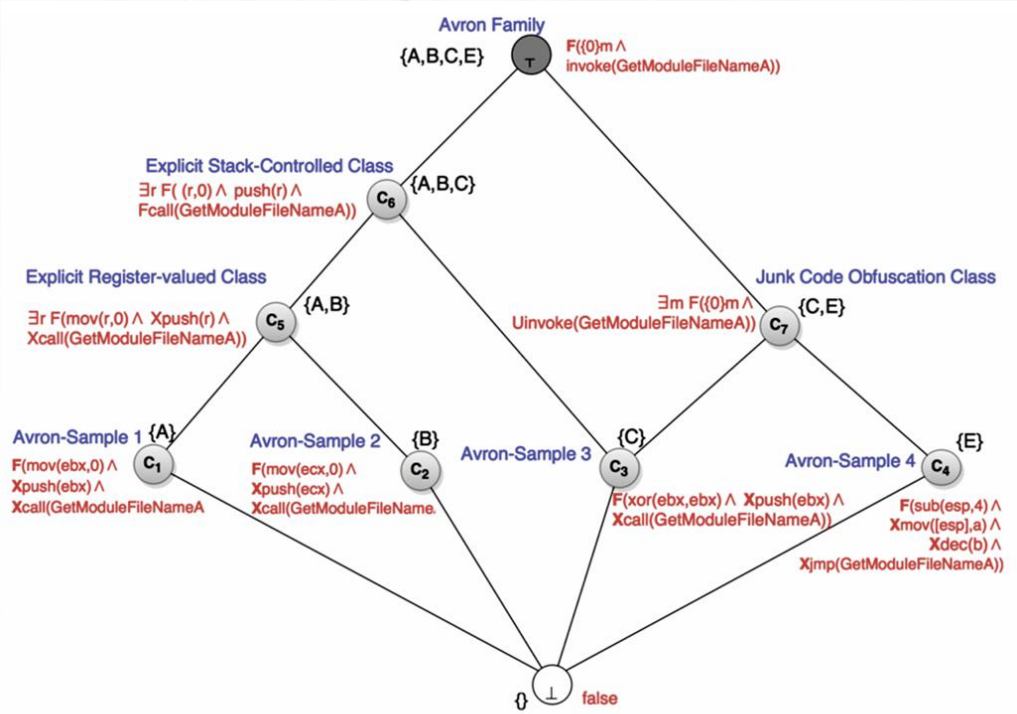


Fig. 3. The concept lattice generated by F-FCA

Financované Európskou úniou
NextGenerationEU

PLÁN OBNOVY

MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

Feature-Driven Formal Concept Analysis for Malware Hierarchy Construction

T. B. Nguyen, C. D. Tran, T. T. Quan, M. H. Nguyen (2015)

Algoritmus FOCA:

- Vstup – F-FCA formálny kontext
- Každý objekt je inicializovaný ako samostatný koncept, koncepty sú inkrementálne porovnávané a zlučované
- Kľúčový je object-joining operátor (operátor spájajúci objekty) – zlučuje dva koncepty pomocou widening (rozširujúceho) operátora, ktorý vytvára abstraktnejšiu logickú formulu implikujúcu správanie oboch konceptov
- Výstup – hierarchia malvéru

Dataset:

- 3 000 vírusových vzoriek z VXHeaven

F-FCA (s použitím FOCA) dosahuje lepšie výsledky z pohľadu hodnotenia kvality hierarchie a hodnotenia kvality zhlukov než FCA

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

Modelovanie hrozieb

Financované Európskou úniou
NextGenerationEU

PLÁN OBNOVY

MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

A Social Dimensional Cyber Threat Model with Formal Concept Analysis and Fact-Proposition Inference

A. Sharma, R. Gandhi, Q. Zhu, W. Mahoney, W. Sousan (2013)

Motivácia:

- Mnohé kybernetické útoky sú reakciou na sociálne konflikty a predstavujú formu „občianskej kybernetickej vojny“ (napr. porušovanie ľudských práv, cenzúra informácií, ...)
- Neexistuje formálny model, ktorý by systematicky prepájal sociálne udalosti, ľudské úmysly, kybernetické útoky a ich dôsledky

Výskumný cieľ:

- Navrhnúť sociálno-dimenzionálny model, ktorý integruje sociálne udalosti do analýzy kybernetických hrozieb a slúži na podporu kybernetického situačného povedomia

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

A Social Dimensional Cyber Threat Model with Formal Concept Analysis and Fact-Proposition Inference

A. Sharma, R. Gandhi, Q. Zhu, W. Mahoney, W. Sousan (2013)

Metodológia:

- Hlavné entity modelu – sociálne udalosti, motívy útočníkov, útočníci, typy útokov, ciele a obete útokov, dôsledky útokov
- **FCA** – formálny nástroj na štruktúrovanie znalostí o útokoch
- Konceptový zväz umožňuje identifikovať všeobecné a špecifické typy hrozieb a analyzovať vzťahy medzi útokmi
- **Inferencia v priestore faktov a propozícií** (Fact-Proposition Inference) – logické odvodzovanie potenciálnych kybernetických hrozieb na základe pozorovaných sociálnych udalostí

Príklad: prebiehajúci politický protest (fakt) – ak dôjde k politickému napätiu, zvyšuje sa riziko DDoS útokov (propozícia) – odvodzujú sa potenciálne hrozby, hodnotí sa ich závažnosť a pravdepodobnosť

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

A Social Dimensional Cyber Threat Model with Formal Concept Analysis and Fact-Proposition Inference

A. Sharma, R. Gandhi, Q. Zhu, W. Mahoney, W. Sousan (2013)

Webová aplikácia FPS ako demonštrácia navrhnutého modelu

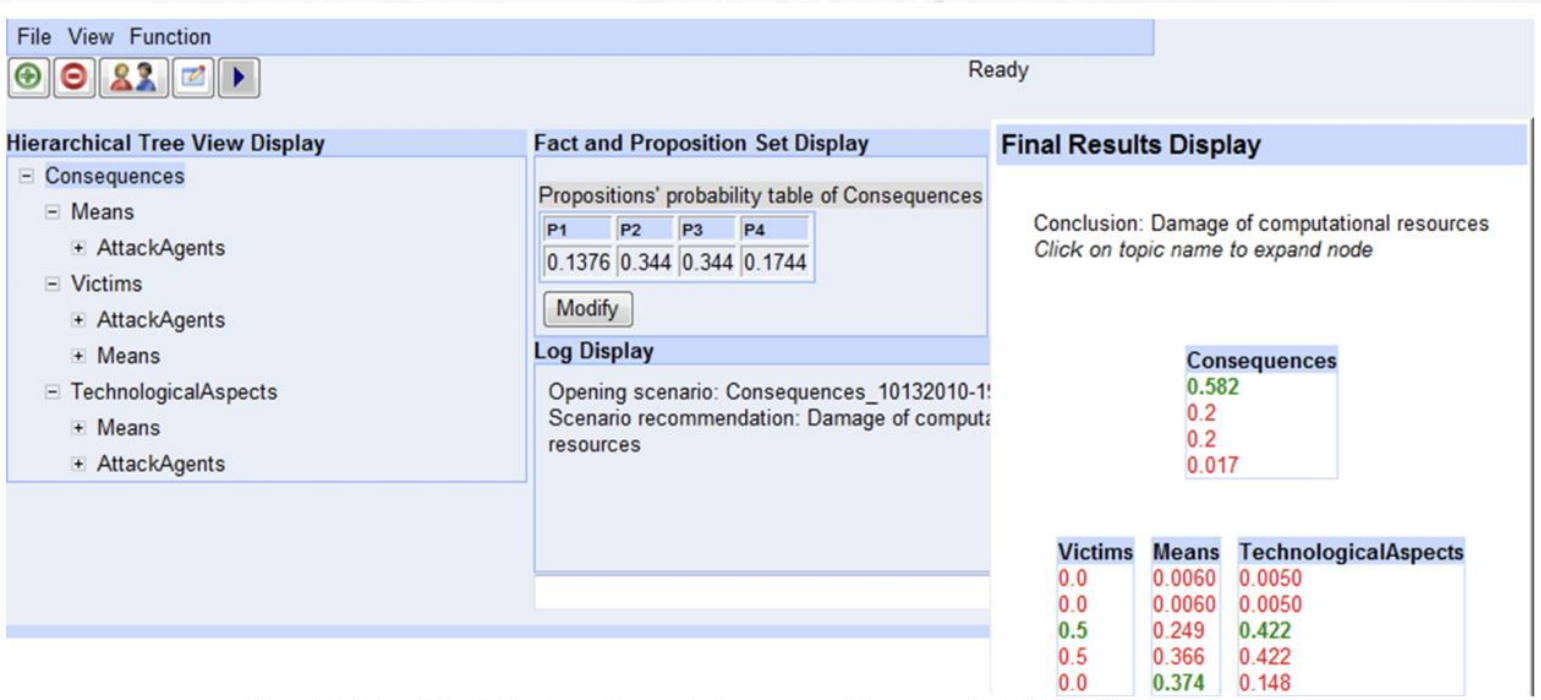


Fig. 13. Graphical display of result the propositions probability of consequences

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

Event-Based Threat Intelligence Ontology Model

P. Wang, G. Dai, L. Zhai (2023)

Motivácia:

- Fragmentácia a nízka interoperabilita existujúcich threat intelligence riešení
- Nedostatočná formalizácia threat intelligence dát
- Potreba jednotnej znalostnej reprezentácie založenej na **udalostiach** – udalosti poskytujú rámec na modelovanie vzťahov medzi aktérmi, technikami, cieľmi a následkami útokov

Výskumné ciele:

- Navrhnúť ontologický threat intelligence model založený na udalostiach na reprezentáciu cyber threat intelligence (CTI) jednotným a formalizovaným spôsobom
- FCA – na semi-automatickú konštrukciu ontológie a podporu spolupráce človek-stroj

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

Event-Based Threat Intelligence Ontology Model

P. Wang, G. Dai, L. Zhai (2023)

Metodológia:

Udalosť ako zákl. jednotka modelu – kto, čo, na čo, ako, kedy, s akým následkom

- Zber threat intelligence dát (opisy útokov, existujúce CTI zdroje, ...)
- Tvorba formálneho kontextu (objekty – konkrétne udalosti, atribúty – typ útoku, aktér, cieľ, ...)
- Konštrukcia konceptového zväzu (zhlukovanie podobných udalostí)
- Extrahovanie ontologických pojmov (triedy ontológie, ontologické vzťahy)

Ontológia zjednocuje terminológiu + vytvára formálnu sémantiku

Vďaka ontológii je možné:

- Korelovať udalosti (napr. rovnaký aktér)
- Inferovať nové poznatky (napr. určitý vzorec správania naznačuje konkrétny typ hrozby)

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

Construction of domain ontology utilizing formal concept analysis and social media analytics

R. Jindal, K. R. Seeja, S. Jain (2020)

Motivácia:

- Vývoj doménových ontológií je časovo náročný problém závislý od experta
- Existujúce korpusy nezachytávajú aktuálne udalosti
- Sociálne médiá poskytujú bohatý a aktuálny zdroj znalostí

Výskumné ciele:

- Navrhnúť semi-automatický prístup na konštrukciu doménovej ontológie
- Použitie dát zo sociálnych médií (Twitter/X) – dynamický a aktuálny zdroj znalostí v rýchlo sa meniacej doméne (terorizmus)

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

Construction of domain ontology utilizing formal concept analysis and social media analytics

R. Jindal, K. R. Seeja, S. Jain (2020)

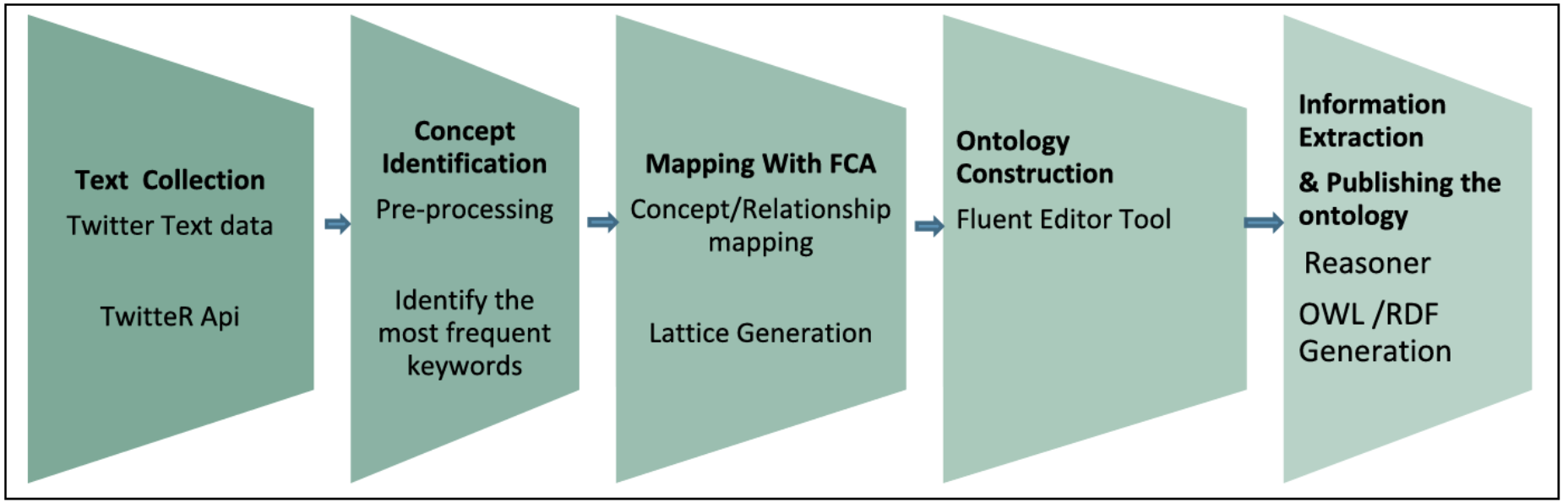


Fig. 2. Process flow diagram for ontology construction.

Financované Európskou úniou
NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

Ďalšie oblasti kybernetickej bezpečnosti

Financované Európskou úniou
NextGenerationEU

PLÁN OBNOVY

MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

Formal concept analysis model for static code analysis

S. Motogna, D. Cristea, D. Șotropa, A.-J. Molnar (2022)

Motivácia:

- **Statická analýza kódu** umožňuje odhaliť chyby pred spustením programu (detekcia bugov, code smells, zlých praktík, bezpečnostných zraniteľností)
- Nástroje ako Pylint, SonarQube, PMD sú bežne používané v praxi a schopné zachytiť významnú časť chýb už v raných fázach vývoja
- Statické analyzátory produkujú veľké množstvo hlásení rôznych typov a závažností – ich výstupy sú pre vývojárov často ťažko interpretovateľné

Výskumné ciele:

- Preskúmať, ako môže FCA poskytnúť matematický základ pre analýzu problémov identifikovaných nástrojom Pylint
- Preskúmať, ako možno model založený na FCA využiť na analýzu rozloženia, frekvencie a vzájomných korelácií medzi hláseniami nástroja Pylint

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

Formal concept analysis model for static code analysis

S. Motogna, D. Cristea, D. Șotropa, A.-J. Molnar (2022)

Metodológia:

- Zber dát zo statickej analýzy kódu (Pylint hlásenia)
- Kritériá výberu (typ hlásenia, kategória pravidla, ...)
- Na základe zvolených kritérií – výber konkrétnych atribútov, ktoré budú uvažované vo formálnom kontexte

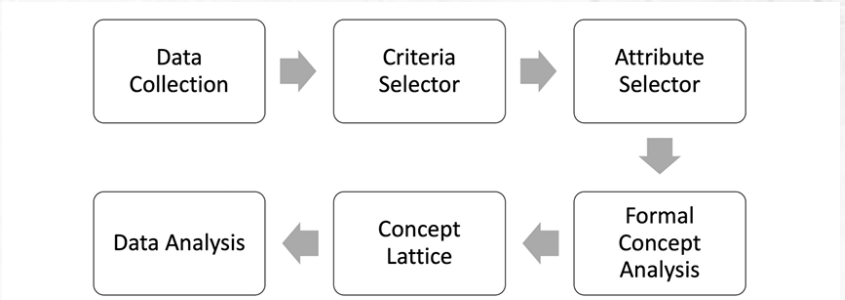


FIGURE 1. Stages of applying FCA to code review results

- Konštrukcia formálneho kontextu a formálnych konceptov (objekty – Pylint hlásenia, atribúty – vybrané vlastnosti)
- Konštrukcia konceptového zväzu + analýza
  - Analýza rozloženia (distribution) – kde sa chyby v kóde koncentrujú
  - Analýza frekvencie (frequency) – ktoré typy hlásení sa najčastejšie vyskytujú
  - Analýza korelácií (correlations) – ktoré chyby sa často vyskytujú spolu

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

Formal concept analysis model for static code analysis

S. Motogna, D. Cristea, D. Șotropa, A.-J. Molnar (2022)

Dataset:

- Študentské riešenia programátorských úloh v Pythone + ich automatická analýza v Pylint

Case studies:

- Kritérium „kontrola dizajnu“ – identifikácia problémov v kóde súvisiacich s návrhovými nedostatkami
  - Zistenia: študenti bez problémov s návrhom kódu, študenti s príliš zložitým kódom, funkcie/metódy s príliš veľkým počtom príkazov
- Kritérium „prevencia chýb“ – na základe pozorovania najčastejších chýb
  - Napr. nešpecifikovanie typu výnimky, odchytenie príliš všeobecnej výnimky

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

A Novel Method for Network Intrusion Detection

H. Wang, Q. Wei, Y. Xie (2022)

Motivácia:

- **Systémy na detekciu prienikov (IDS)** zohrávajú kľúčovú úlohu pri identifikácii neoprávnených alebo škodlivých aktivít v sieťovej prevádzke
- Existujúce IDS prístupy sú založené na strojovom učení (ML), neurónových sieťach (NN), podporných vektorových strojoch (SVM)
- Avšak často trpia vysokou mierou falošných poplachov, majú obmedzenú schopnosť detekcie nových/neznámych útokov, fungujú ako „black-box“ modely s nízkou interpretovateľnosťou

Výskumné ciele:

- Navrhnúť IDS model založený na fuzzy triadickej FCA (FCTA)
- Modelovať vzťahy medzi sieťovými spojeniami, atribútmi a typmi útokov
- Experimentálne overiť a porovnať navrhovaný prístup s existujúcimi metódami

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

A Novel Method for Network Intrusion Detection

H. Wang, Q. Wei, Y. Xie (2022)

Metodológia:

- FCTA: **objekty** – sieťové spojenia, **atribúty** – charakteristiky spojení (počet paketov, dĺžka spojenia, ...), **podmienky** – typy útokov (DoS, DDoS, Brute Force, ...)
- Fuzzifikácia – hodnoty z jednotkového intervalu (TF-IDF váhovanie + Z-score normalizácia)
- Fuzzy triadické koncepty – typické vzory útokov
- Transformácia konceptov na atribútové vektory (profily útokov)

Klasifikácia nového spojenia:

- Vypočíta sa jeho atribútový vektor
- Porovná sa s existujúcimi vektormi (na základe Euklidovskej vzdialenosti)
- Spojenie sa priradí k najbližšiemu útoku/konceptu

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

A Novel Method for Network Intrusion Detection

H. Wang, Q. Wei, Y. Xie (2022)

Dataset:

- CIC-IDS2018

Porovnanie a vyhodnotenie FCTA modelu:

- Metriky: Accuracy – celková presnosť klasifikácie, Detection Rate (DR) – miera správne detegovaných útokov, False Alarm Rate (FAR) – miera falošných poplachov
- Model je porovnávaný s klasickými ML metódami: SVM, KNN (k-Nearest Neighbors), BP Neural Network (neurónová sieť so spätným šírením chyby)

Výsledky:

- FCTA dosahuje vyššiu presnosť než porovnávané metódy
- FCTA má lepšiu schopnosť zachytiť útoky
- FCTA generuje menej falošných poplachov

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

A novel outlier detection approach based on formal concept analysis

Q. Hua, Z. Yuan, K. Qin, J. Zhang (2023)

Motivácia:

- **Outlier** (anomália, odľahlý objekt) – zriedkavý, neobvyklý alebo potenciálne významný jav
- Detekcia outlierov je kľúčová v mnohých oblastiach (dolovanie dát, kyberbezpečnosť, ...)
- Limity tradičných prístupov založených na štatistike, vzdialenosti, hustote, zhlukovaní
- Novšie prístupy sú založené na drsných množinách a Granular Computing

Výskumné ciele:

- Navrhnúť nový prístup na detekciu outlierov založený na FCA
- Experimentálne overiť navrhnutý prístup na rôznych datasetoch a porovnať ho s existujúcimi metódami detekcie outlierov

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

A novel outlier detection approach based on formal concept analysis

Q. Hua, Z. Yuan, K. Qin, J. Zhang (2023)

Algoritmus Granular concept-based outlier detection (GCOD):

- Výpočet granulárnych konceptov z formálneho kontextu
- Výpočet vzdialeností medzi granulárnymi konceptami
- Výpočet stupňa odľahlosti granulárnych konceptov (outlier degree of granular concept (GOD)) – súčet vzdialeností medzi daným granulárnym konceptom a ostatnými granulárnymi konceptami, normalizovaný počtom granulárnych konceptov
- Výpočet faktora odľahlosti objektov (granular concept-based outlier factor (GCOF)) – vypočítaný agregáciou hodnôt GOD všetkých granulárnych konceptov, do ktorých objekt patrí
- Zoradenie objektov podľa GCOF – objekty s najvyššími hodnotami sú identifikované ako outliery

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

A novel outlier detection approach based on formal concept analysis

Q. Hua, Z. Yuan, K. Qin, J. Zhang (2023)

Vyhodnotenie navrhnutého algoritmu GCOD:

- 15 datasetov rôznych veľkostí, s rôznym charakterom a s rôznymi atribútmi (numerické, kategorické, zmiešané)
- GCOD je porovnávaný s viacerými reprezentatívnymi prístupmi na detekciu outlierov
- Hodnotiace metriky – precision, recall, F skóre, ROC krivka, AUC

Experimentálne výsledky ukazujú, že GCOD dosahuje porovnateľné alebo lepšie výsledky než konkurenčné metódy a je stabilný naprieč rôznymi datasetmi

- Prístup založený na FCA navyše poskytuje lepšiu interpretovateľnosť výsledkov
- Použitie granulárnych konceptov namiesto konštrukcie konceptového zväzu znižuje výpočtovú zložitosť

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

Bimorphisms and attribute implications in heterogeneous formal contexts

Ľ. Antoni, P. Eliaš, J. Guniš, D. Kotlárová, S. Krajči, O. Krídlo, P. Sokol, Ľ. Šnajder (2024)

Motivácia:

- Klasická a fuzzy FCA predpokladajú jednotnú štruktúru hodnôt pre všetky objekty a atribúty
- V mnohých reálnych aplikáciách sú však údaje heterogénne a využívajú rôzne škály a typy hodnôt
- Na modelovanie takýchto dát boli zavedené heterogénne formálne kontexty

Výskumný cieľ:

- Definovať alternatívnu formuláciu heterogénnych formálnych kontextov pomocou bimorfizmov a rozšíriť atribútové implikácie do heterogénneho prostredia

Ilustračné aplikačné domény:

- GDPR/ochrana osobných údajov
- Zmluvy z oblasti kyberbezpečnosti

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

Bimorphisms and attribute implications in heterogeneous formal contexts

Ľ. Antoni, P. Eliaš, J. Guniš, D. Kotlárová, S. Krajči, O. Krídlo, P. Sokol, Ľ. Šnajder (2024)

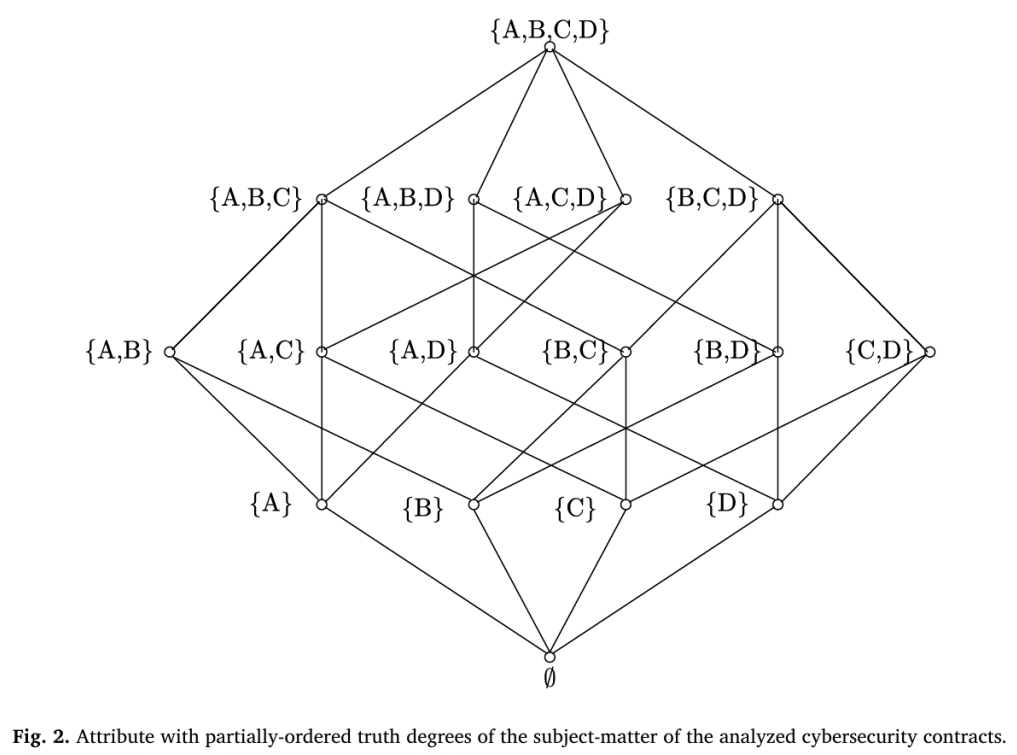


Fig. 2. Attribute with partially-ordered truth degrees of the subject-matter of the analyzed cybersecurity contracts.

Čiastočne usporiadaná množina pravdivostných stupňov pre atribút „predmet zmluvy“ (kyberbezpečnostné zmluvy)

- A: Poskytovanie kvalifikovaného odborného personálu
- B: Poskytovanie služieb (proaktívnych a reaktívnych opatrení) súvisiacich s riešením bezpečnostných udalostí a incidentov
- C: Poskytovanie služieb súvisiacich s implementáciou bezpečnostných opatrení
- D: Poskytovanie služieb týkajúcich sa kybernetického auditu a kontroly

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY

Formálna konceptová analýza v kyberbezpečnosti

Február 2026

Financované Európskou úniou
NextGenerationEU

PLÁN OBNOVY

MINISTERSTVO INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA A INFORMATIZÁCIE SLOVENSKEJ REPUBLIKY