FAKULTA INFORMAČNÍCH TECHNOLOGIÍ ČVUT V PRAZE

Posudek oponenta závěrečné práce

Oponent práce:	Ing. Tomáš Čejka, Ph.D.
Student:	Bc. Jaroslav Hlaváč
Název práce:	Framework pro detekci hrozeb z heterogennich dat
Obor / specializace:	Počítačová bezpečnost
Vytvořeno dne:	30. května 2021

Hodnotící kritéria

1. Splnění zadání

- ▸ [1] **zadání splněno**
- [2] zadání splněno s menšími výhradami
- [3] zadání splněno s většími výhradami
- [4] zadání nesplněno

Výsledkem práce je dobře zpracovaný přehled existujících prací, návrh a realizace frameworku pro zpracování dat z provozu síťové infrastruktury a koncových zařízení. Zadání práce bylo splněno.

2. Písemná část práce 85 / 100 (B)

Text práce je vysoce kvalitní v anglickém jazyce, text je dobře členěný a srozumitelně popsaný. Práce obsahuje drobné typografické nedostatky. Seznamy obrázků a tabulek jsou nepřehledné, v tomto případě by bylo vhodné použít krátké popisky. Text obsahuje dostatečné množství citovaných zdrojů, avšak některé záznamy v seznamu referencí se nezdají úplně kompletní, např. u [4] není jasné, odkud zdroj je, podobně u [28] je pouze vydavatel, ale chybí název knihy/sborníku/časopisu.

3. Nepísemná část, přílohy 85 / 100 (B)

Výstupem práce jsou především zdrojové kódy obsažené v "Jupyter sešitech" (notebooks). Vytvořený kód je zdokumentovaný, obsahuje ukázky. Zdrojové kódy - soubory by možná mohly být lépe organizované; chybí návod jak s frameworkem pracovat, např. jak a kde rozšiřovat funkcionalitu. Zdá se, že v odevzdaném archivu zdrojových kódů je pouze jedna komponenta frameworku ze tří, konkr. Event Generation Layer podle Figure 1.1.

4. Hodnocení výsledků, jejich využitelnost 100 /100 (A)

Vytvořené výsledky slouží jako součást existujícího komerčního systému Cognitive Threat Analytics. Z toho se dá předpokládat, že jsou využitelné v praxi. Vytvořené zdrojové kódy navíc mohou sloužit jako základ dalšího výzkumu v oblasti zpracování síťových dat a detekci anomálií.

Celkové hodnocení 89 /100 (B)

Odevzdaná práce je celkově velmi dobře zpracovaná, výsledky jsou použitelné v praxi. Tyto pozitivní vlastnosti závěrečné práce jednoznačně převyšují nedostatky v textové části práce. Celkové hodnocení je na hranici známek A a B; po zodpovězení otázek, které nejsou v textu práce příliš adresované, je vhodné práci klasifikovat známkou A.

Otázky k obhajobě

1. Bylo by možné vytvořený framework rozšířit tak, aby bylo možné data zpracovávat proudově (stream-wise) namísto dávkově (jak popisuje text práce)?
2. Je možné jako zdroj dat do frameworku použít např. data z open source systému NEMEA nebo z jiných open source nástrojů (např. ipfixprobe, yaf, ntop)?
3. Kolik provozu (např. v podobě IP flow záznamů) je vytvořený prototyp frameworku schopen zpracovávat za jednotku času?

Instrukce

Splnění zadání

Posuďte, zda předložená ZP dostatečně a v souladu se zadáním obsahově vymezuje cíle, správně je formuluje a v dostatečné kvalitě naplňuje. V komentáři uveďte body zadání, které nebyly splněny, posuďte závažnost, dopady a případně i příčiny jednotlivých nedostatků. Pokud zadání svou náročností vybočuje ze standardů pro daný typ práce nebo student případně vypracoval ZP nad rámec zadání, popište, jak se to projevilo na požadované kvalitě splnění zadání a jakým způsobem toto ovlivnilo výsledné hodnocení.

Písemná část práce

Zhodnoťte přiměřenost rozsahu předložené ZP vzhledem k obsahu, tj. zda všechny části ZP jsou informačně bohaté a ZP neobsahuje zbytečné části. Dále posuďte, zda předložená ZP je po věcné stránce v pořádku, případně vyskytují-li se v práci věcné chyby nebo nepřesnosti.

Zhodnoťte dále logickou strukturu ZP, návaznosti jednotlivých kapitol a pochopitelnost textu pro čtenáře. Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 26/2017, článek 3.

Posuďte, zda student využil a správně citoval relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami. Zhodnoťte, zda převzatý software a jiná autorská díla, byly v ZP použity v souladu s licenčními podmínkami.

Nepísemná část, přílohy

Dle charakteru práce se případně vyjádřete k nepísemné části ZP. Například: SW dílo – kvalita vytvořeného programu a vhodnost a přiměřenost technologií, které byly využité od vývoje až po nasazení. HW – funkční vzorek – použité technologie a nástroje, Výzkumná a experimentální práce – opakovatelnost experimentů.

Hodnocení výsledků, jejich využitelnost

Dle charakteru práce zhodnoťte možnosti nasazení výsledků práce v praxi nebo uveďte, zda výsledky ZP rozšiřují již publikované známé výsledky nebo přinášející zcela nové poznatky.

Celkové hodnocení

Shrňte stránky ZP, které nejvíce ovlivnily Vaše celkové hodnocení. Celkové hodnocení nemusí být aritmetickým průměrem či jinou hodnotou vypočtenou z hodnocení v předchozích jednotlivých kritériích. Obecně platí, že bezvadně splněné zadání je hodnoceno klasifikačním stupněm A.