FAKULTA INFORMAČNÍCH TECHNOLOGIÍ ČVUT V PRAZE

Posudek oponenta závěrečné práce

Oponent práce:	Mgr. Martin Jureček
Student:	Bc. Mikuláš Formánek
Název práce:	Automatická detekce nebezpečných aktivit ve vnitřní síti
Obor / specializace:	Počítačová bezpečnost
Vytvořeno dne:	31. 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

Všechny body popsané v pokynech pro vypracování práce považuji za splněné.

2. Písemná část práce 60 /100 (D)

V práci se vyskytuje několik desítek drobných chyb, avšak čitelnost práce se tím výrazněji nesnižuje. Rozsah práce je na dolní hranici a z textu je cítit, že ho student psal na poslední chvíli (např nedokončená věta na str. 28). Popis state of art algoritmů je velmi stručný (někdy jen v 1 až 2 odstavcích) a bez předchozích znalostí nelze z popisu diplomanta pochopit, jak algoritmy fungují. Není sjednocené značení (např. TadGAN, kap. 2.1.6). V textu je nadměrné množství citovaných pasáží, které nejsou v uvozovkách (jsou ale označeny šikmým písmem). Vzorce nejsou očíslovány a nejsou v nich vysvětleny všechny pojmy (např. integrál na str. 28, navíc ve vzorci chybí "dx"). V textu je nerovnoměrná velikost obrázků, které jsou často v horší kvalitě (např. Fig. 1.3, 2.1,2.9, ...). Nakonec členění práce je zbytečně jemné (např. kap. 1.3.3.0.1).

3. Nepísemná část, přílohy 90 /100 (A)

Implementace byla provedena v jazyce Python a pro experimentální analýzy byl využit nástroj Jupyter Notebook. Použitý dataset není k dispozici, proto není možné ověřit experimentální výsledky.

4. Hodnocení výsledků, jejich využitelnost 90 /100 (A)

Dosažené výsledky jsou vzhledem k tomu, že se využily neoznačené data, poměrně zajímavé a jako diplomant v práci navrhuje, mohli by se ještě dále vylepšit a tak případně mít potenciál nasazení do produkce.

Celkové hodnocení 70 /100 (C)

Diplomant úspěšně aplikoval metody strojového učení na problém detekce anomálií, konkrétně detekce brute-force útoků. Avšak vzhledem k slabší písemné části celkově hodnotím diplomovou práci na dolní hranici známky C.

Otázky k obhajobě

1. Jaká byla velikost datasetu a na základě čeho se vybíraly příznaky?
2. Mohl by diplomant detailněji vysvětlit výpočet finálního skóre u NAB algoritmu (str. 24)?
3. Mohl by diplomant přesněji vysvětlit Dynamic time warping u TadGAN algoritmu (str. 28)?
4. V názvu práce je "Automatická detekce". Do jaké míry je detekce popsána v práci automatická?

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.