FAKULTA INFORMAČNÍCH TECHNOLOGIÍ ČVUT V PRAZE

Posudek oponenta závěrečné práce

Oponent práce:	Ing. Karel Hynek
Student:	Bc. Tibor Engler
Název práce:	Analýza a identifikace chování Tor klientů
Obor / specializace:	Počítačová bezpečnost
Vytvořeno dne:	19. 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

Zadání bylo splněno v celém rozsahu. Podle mého názoru převyšuje tato práce svou kvalitou a hlavně náročností standard kladený na diplomovou práci na FIT. Oceňuji skutečně důkladné nastudování relevantních prací, shrnujících poměrně velké množství útoku na Tor.

2. Písemná část práce 90 /100 (A)

Text práce je psaný v angličtině a je dobře strukturovaný. V rámci typografie jsem zaznamenal jen pár nekonzistencí drobného charakteru. Všiml jsem si, že práce občas používá výrazy, které se do odborného textu nehodí a jsou spíše vypravěčského typu. Dále mi občas chyběly citace, hlavně Alexa TOP list by určitě chtělo minimálně odkázat pod čarou. I přes výše zmíněné nedostatky považuji text práce za velice kvalitní.

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

Nepísemná část přílohy se skládá z python implementace sítě DeepCorr publikované autory Nasr et al. v roce 2018. Student vytvořil vlastní implementaci této sítě v jazyce python, což není ani z textu ani z dokumentace úplně jasné. Celkově by dokumentace nepísemné části mohla být propracovanější. Ačkoliv jsou zdrojové kódy psané čitelně, strukturovaně a obsahují komentáře, není využito žádného nástroje pro automatickou tvorbu dokumentace, což by u závěrečné práce na FIT mělo být samozřejmostí.

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

Využitelnost výsledků je podle mého názoru vysoká. Studentovi se povedlo vylepšit existující útok na Tor pomocí hlubokého učení, který ale používá agregovaná data jako jsou například histogramy a paketovou sekvenci o kratší délce. To výrazným způsobem zjednoduší provedení útoku na reálné síti, protože nebude třeba sbírat ohromné množství dat z každého toku, jako tomu bylo doposud. Doporučuji výsledky diplomové práce publikovat.

Celkové hodnocení 95 /100 (A)

Diplomovou práci považuji za velice kvalitní. Student si pečlivě nastudoval problematiku útoků na Tor a problematiku hlubokého učení. Tyto informace dokázal následně použít a vylepšit jeden z nejpřesnějších korelačních útoků. Oceňuji i tvorbu datové sady, při které byl vytvořen opravdový aktivní Tor uzel, který byl omezen pouze na webové porty. Student tak nechal reálné uživatele generovat data a šum, což mi připadá skvělé a značným způsobem to zvyšuje důvěryhodnost výsledků. Z výše popsaných důvodů nemohu hodnotit práci jinak než stupněm A a doporučit ji k obhajobě.

Otázky k obhajobě

1, Model vyhodnocujete na velkém množství datových sad pocházející z několika poměrně rozdílných zdrojů. Docházelo k přetrénování modelu při vyhodnocování na každé datové sadě, nebo má model opravdu tak vysokou úspěšnost napříč rozdílnými datovými sadami?
2, Jaké jsou zkušenosti s provozováním Tor uzlu? Chodilo vám hodně upozornění ohledně "škodlivé" činnosti? Jak jste se s nimi vypořádal?
3, V textu píšete, že paketové dávky exportované v ipfixprobe mátly klasifikátor. Dokážete vysvětlit, proč tomu tak bylo?

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.