FAKULTA INFORMAČNÍCH TECHNOLOGIÍ ČVUT V PRAZE

Posudek oponenta závěrečné práce

Oponent práce:	Ing. Matouš Kozák
Student:	Bc. Pavla Louthánová
Název práce:	Porovnání adversariálních učících technik pro detekci malwaru
Obor / specializace:	Počítačová bezpečnost
Vytvořeno dne:	3. června 2023

Hodnotící kritéria

1. Splnění zadání

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

Zadání považuji za splněné. Jedná se o kvalitní diplomovou práci s množstvím experimentální výsledků.

2. Písemná část práce 87 /100 (B)

Diplomová práce má dostatečný rozsah a kvalitní formu. Práce obsahuje rozsáhlý úvod do adversariálního strojového učení s důrazem na oblast detekce malware a podrobný popis PE file formátu spolu s aktuálně používanými modifikacemi binárních souborů. V rešeršní části jsou popsány relevantní a aktuální práce z oblasti adversariálních malware útoků. Popis experimentů a dosažených výsledků je podrobný a dobře členěný. V závěrů práce jsou shrnuty dosažené výsledky, ale ocenil bych i zahrnutí konkrétních návrhů pro budoucí práce (viz. otázky k obhajobě). Seznam použité literatury je aktuální a relevantní k tématu.

Z formální stránky se jedná o práci povedenou, ale s několika výhradami:
- u některých obrázků chybí odkazy z hlavního textu (např. u Figure 1.4 nebo Figure 1.5)
- některé zkratky jsou zavedeny vícekrát (např. ACER na straně 31 a 39)
- pro lepší čitelnost Tabulky 5.4 by bylo vhodné seřadit řádky (.data, .rdata, .text)
- na stránce 56 je použit obrat "level of leakage", ale ve zbytku práce se používá označení "evasion rate" (doporučoval bych používat pouze "evasion rate", tak jak je běžné v aktuální literatuře)
- práce obsahuje několik překlepů a chyb, které ale nebrání pochopení textu

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

Všechny zdrojové kódy jsou součástí příloh, experimenty jsou tedy reprodukovatelné. Použitý dataset není součástí příloh, ale v diplomové práci je zmíněno, jak byl dataset vytvořen, tedy je možné dataset získat vlastní cestou. Součástí přílohy je také dokumentace, která usnadňuje orientaci a obsahuje návod pro spuštění jednotlivých skriptů. Skripty jsou napsány v jazyce Python, který se běžně v této oblasti využívá.

4. Hodnocení výsledků, jejich využitelnost 95 /100 (A)

Práce obsahuje experimentální porovnání 5 generátorů adversariálního malwaru s důrazem na praktické využití. Experimenty jsou provedeny na 9 antivirových programech, které zahrnují aktuální špičku v antivirovém průmyslu. Podle mých znalostí, se jedná o jednu z mála prací, která neporovnává generátory adversariálního malwaru pouze teoreticky, ale přináší i experimentální vyhodnocení na realných antivirových programech.

K vyhodnocení jsou použity relevantní metriky, přičemž některé jsou nové a nevyskytují se v aktuální literatuře (např. "evasion rate cmp"). Pro ještě lepší porovnání generátů bych v budoucnu místo metriky "size increase" doporučoval použít např. "perturbation size", která by reflektovala velikost modifikace daného PE souboru. Tím by nedocházelo k nulovým výsledkům u metod jako jsou Partial DOS nebo Full DOS.

Celkové hodnocení 94 /100 (A)

Práci hodnotím jako zdařilou, zadání bylo splněno bez výhrad. Oceňuji kvalitní výběr testovaných metod, který pokrývá všechny 3 aktuální směry generování adversariálního malwaru (gradientní / reinforcement learning / evoluční útoky). Dále oceňuji zavedení nových metrik, které dobře reflektují úspěch/neúspěch kombinace více generátorů.

Práci hodnotím známkou A a doporučuji k obhajobě.

Otázky k obhajobě

Existuje nějaké rozšíření/zlepšení této práce, které by mohlo přínést další zajímavé výsledky?

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. 52/2021, 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.