VŠB TECHNICKÁ UNIVERZITA OSTRAVA | FAKULTA ELEKTROTECHNIKY A INFORMATIKY | KATEDRA TELEKOMUNIKAČNÍ TECHNIKY

Oponentský posudek dizertační práce na téma:

Reputace zdrojů škodlivého provozu

Dizertační práci vypracoval: Ing. Václav Bartoš
Fakulta informačních technologií
Vysoké učení technické v Brně

Oponent: prof. Ing. Miroslav Vozňák, Ph.D.
Fakulta elektrotechniky a informatiky
VŠB – Technická univerzita Ostrava

Předložená dizertační práce se zabývá způsobem zhodnocení míry bezpečnostních hrozeb a svým tématem spadá do studijního oboru Výpočetní technika a informatika. **Téma práce je vysoce aktuální** a tuto skutečnost rovněž dokládá výčet souvisejících publikací z posledních let, na které se dizertant odkazuje.

Po nezbytném úvodu do tématu a upozornění na související práce týkajících se ohodnocování škodlivosti síťových entit a predikce jejich chování se autor věnuje popisu své myšlenky hodnocení reputace, návrhu systému pro analýzu síťového provozu, návrhu a implementaci reputační databáze, analýze získaných dat a nakonec experimentům a jejich vyhodnocením. Prakticky vše od čtvrté kapitoly, tzn. str. 25, je již věcnou částí dizertace a dizertant v těchto kapitolách popisuje výzkum, na kterém pracoval. Autor srozumitelně vysvětlil jednotlivé kroky v dizertaci a dosažené výsledky, nicméně mi při studiu práce chyběl rejstřík zkratek, který by čtenáři usnadnil orientaci. Drobnou výtku mám rovněž k umístění cílů dizertace, které jsou zanořeny v úvodu na straně třetí, nicméně by si zasloužily vlastní podkapitolu.

Hlavním cílem disertační práce je návrh metody výpočtu reputačního skóre přiřazeného každé jednotlivé podezřelé entitě, jako je např. IP adresa. Skóre má

zahrnout informace, jak o předchozím chování této entity, tak i o chování blízkých či jinak podobných entit. Konstatuji, že **dizertant cíl dizertace splnil**.

Jádro a přínos dizertační práce lze spatřovat jednak v kapitole čtvrté, kde je vlastní návrh obecné metody hodnocení reputace síťových entit na základě předpovědi jejich budoucího chování, a v kapitole osmé a deváté, kde je ověření navržené metody nad reálnými daty o škodlivých IP adresách. Dizertace přinesla i další vedlejší výsledky, které jsou zajímavé pro komunitu a jedná se jednak o návrh systému pro analýzu síťového provozu NEMEA (Network Measurements Analysis) a dále o návrh nových metod pro detekci škodlivého síťového provozu na základě analýzy dat o síťových tocích, viz pátá kapitola. Přínosný je rovněž návrh a následná implementace reputační databáze síťových entit popsaná v šesté kapitole a analýzy charakteristik bezpečnostních hlášení a zdrojů škodlivého provozu, kterým je věnována sedmá kapitola. Stěžejní myšlenka práce spočívá v metodě výpočtu reputačního skóre.

V práci je představena nová metoda pro číselné vyjádření reputace síťových entit z hlediska bezpečnostních hrozeb ve formě FMP (Future Misbehavior Probability) skóre, které představuje skalár zahrnující dostupné bezpečnostně relevantní informace dané entity s predikcí budoucího chování. Jak autor výstižně popisuje v závěru, FMP skóre vyjadřuje pravděpodobnost, že daná entita bude v příštích 24 hodinách detekována jako zdroj určitého nežádoucího chování. Dizertant provedl řadu experimentů vyhodnocujících určování FMP skóre IP adres s využitím strojového učení. Konkrétně použil neuronové sítě a GBDT (Gradient Boosted Decision Trees), kde prokázal možnost odhadovat skutečnou pravděpodobnost budoucích hlášení.

Dosažené výsledky dizertace mají potenciál praktického využití, což dokládá i reálná implementace výpočtu FMP skóre do systému NERD (Network Entity Reputation Database) a zamýšlené využití sdružením CESNET. Navíc je plánováno využití FMP skóre i v rámci evropského projektu GN4 a v evropském projektu PROTECTIVE3 (H2020).

Předložená **dizertační práce přináší původní poznatky** a jedná se o vědeckou práci. Jádro dizertace bylo opublikováno a uchazeč uvádí výstupy, které dokládají jeho erudici a schopnost dosahovat kvalitních výsledků vědecké práce. Zde se lze opřít především o článek s názvem “Network entity characterization and attack prediction“ publikovaném v prvosledovém časopise “Future Generation Computer Systems,“ který patří k nejlepším v oboru (7/103 dle WoS v oboru Computer Science, Theory & Methods).

K předložené práci mám následující dotazy:

1. Objasněte, jakým způsobem jste nalezl optimální strukturu neuronové sítě a stejně tak v případě rozhodovacích stromů. Z popisu návrhu v kap. 9.1 není zřejmé, zda jste postupoval empiricky anebo jste použil nějaký systematický přístup.

2. Na str. 87 zmiňujete experimenty s využitím rekurentních neuronových sítí typu LSTM (Long Shortterm Memory). V posledních letech je zřejmý odklon od LSTM a spíše se namísto LSTM využívají TCN (Temporal Convolutional Network). Je nějaký důvod nepoužít TCN?

Dizertace **splňuje** podmínky samostatné tvůrčí vědecké práce, obsahuje původní a autorem dizertační práce publikované výsledky vědecké práce, a proto

doporučuji

předloženou dizertační práci k obhajobě v souladu s § 47 zákona č. 111/1998 Sb. a po úspěšné obhajobě udělit uchazeči akademický titul Ph.D.

V Ostravě, 6. 5. 2019

.....
prof. Ing. Miroslav Vozňák, Ph.D.