Posudek oponenta závěrečné práce

České vysoké učení technické v Praze **Fakulta informačních technologií**

Student: Bc. Petr Lessner
Oponent práce: Ing. Tomáš Zahradnický, Ph.D.
Název práce: Detekce anomálií síťového provozu pomocí data miningové analýzy síťových toků
Obor: Počítačová bezpečnost

Datum vytvoření: 27. 1. 2016

Hodnotící kritérium:	*Způsob hodnocení - následující škálou 1 až 5:*
1. Náročnost a další komentář k zadání	*1=mimořádně náročné zadání,* *2=náročnější zadání,* ***3=průměrně náročné zadání,*** *4=lehčí, ale ještě dostatečně náročné zadání,* *5=nedostatečně náročné zadání*

Popis kritéria:
Podrobněji charakterizujte diplomovou (bakalářskou) práci a její případné návaznosti na předchozí nebo běžící projekty. Dále posuďte, čím je zadání této ZP náročné. (U obtížnější ZP lze dále tolerovat některé nedostatky, které by u ZP standardní obtížnosti tolerovány nebyly; a naopak u jednoduché ZP mohou být zjištěné nedostatky hodnoceny přísněji.)

Komentář:
Cílem práce bylo návrh, implementace a testování selekčních a detekčních metod ve formě zásuvného modulu pro sondu NfSen. Takové zadání považuji za středně obtížné.

Hodnotící kritérium:	*Způsob hodnocení - následující škálou 1 až 4:*
2. Splnění zadání	***1=zadání splněno,*** *2=zadání splněno s menšími výhradami,* *3=zadání splněno s většími výhradami,* *4=zadání nesplněno*

Popis kritéria:
Posuďte, zda předložená ZP splňuje zadání. V komentáři uveďte body zadání, které nebyly zcela splněny, případně rozšíření ZP oproti původnímu zadání. Nebylo-li zadání zcela splněno, pokuste se posoudit závažnost, dopady a případně i příčiny jednotlivých nedostatků.

Komentář:
Student navrhl a implementoval modul pro NfSen a otestoval ho na reálných síťových datech. Proto považuji zadání za splněné.

Hodnotící kritérium:	*Způsob hodnocení - následující škálou 1 až 4:*
3. Rozsah písemné zprávy	***1=splňuje požadavky,*** *2=splňuje požadavky s menšími výhradami,* *3=splňuje požadavky s většími výhradami,* *4=nesplňuje požadavky*

Popis kritéria:
Zhodnoťte přiměřenost rozsahu předložené ZP vzhledem k obsahu, tj. zda všechny části ZP jsou informačně bohaté a ZP neobsahuje zbytečné části.

Komentář:
Práce svým rozsahem splňuje požadavky na diplomovou práci.

Hodnotící kritérium:	*Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):*
4. Věcná a logická úroveň práce	*85 (B)*

Popis kritéria:
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.

Komentář:
Faktickou a logickou stránku považuji za velmi dobrou. Kapitola návrhu by měla více obsahovat zdůvodnění, proč se student rozhodl tak či onak, ne jen výběr balíčků na základě toho, že jsou cit. „široce doporučované“ - kým, citace chybí.

Hodnotící kritérium:	*Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):*
5. Formální úroveň práce	*65 (D)*

Popis kritéria:
Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 12/2014, článek 3.

Komentář:
Formální úroveň práce považuji jen za uspokojivou. Pokud jsou navrhovány a implementovány detekční metody, očekával bych jejich popis prostřednictvím formalizmu. Bohužel takové popisy v práci chybí. Nejméně u algoritmů DBSCAN a LOF, které jsou pro práci kritické, to považuji za nutnost. Popis algoritmů na stranách 12-13 nepovažuji za dostatečný. Také způsob vyhodnocení a experimenty s nastavováním hodnot parametrů algoritmů nenacházím.

Nacházím nepřesnosti např. TAP na str. 7 není zkratka Test Access Point, ale síťový odposlech, z angl. tap, přestože to uvádí citace [9].

Věty typu „Různé pomocné funkce zde nebudu zbytečně rozepisovat, vše je vždy alespoň stručně okomentováno v kódu“ (str. 33) anebo „Nevím, jak přesně IPython notebook ukládá historii výstupů, nicmého jeho paměť se postupně zvyšovala až na úroveň, kdy byly jeho reakce velmi pomalé a bylo nutné ho restartovat“ (str. 34) považuji za nevhodné. „Toto zadání bylo splněno...“ (str. 43) - splnění zadání hodnotí vedoucí práce a oponent, nikoliv student.

Po jazykové stránce práce hýří anglicizmy typu bufferování, pluginu, opensource, notebookem, timeoutem, kompilaci. Nacházím i hrubou chybu (str. 32 poslední věta nad sekcí 3.2.4). Vlastní jména programovacích jazyků a balíčků jsou psána obvykle s malým písmenem (python, perl, atd.). Vypadá to, jako by práce neprošla jazykovou korekturou.

Hodnotící kritérium:	*Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):*
6. Práce se zdroji	*75 (C)*

Popis kritéria:
Vyjádřete se k aktivitě studenta při získávání a využívání studijních materiálů k řešení ZP. Charakterizujte výběr studijních pramenů. Posuďte, zda student využil všechny relevantní zdroje nebo zda se pokoušel řešit již vyřešené problémy. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků a úvah, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami.

Komentář:
Práce obsahuje značné množství odkazů do literatury, avšak mnoho z nich pochybné kvality, nejvíce z webu. Nacházím mnoho webových zdrojů z veřejných encyklopedií a různých návodů. Je škoda, že student čerpal tak málo z odborných knih a časopisů, kterých je na toto téma mnoho.

Strukturu literatury považuji za nekonzistentní. Někdy je jako autor uváděno jméno produktu (např. [5], [6], [7]), jindy jméno televizního kanálu (např. [8]), anebo jméno webové stránky.

Hodnotící kritérium:	*Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):*
7. Hodnocení výsledků, publikační výstupy a ocenění	*85 (B)*

Popis kritéria:
Vyjádřete se k úrovni dosažených hlavních výsledků ZP, např. k úrovni teoretických výsledků, nebo k úrovni a funkčnosti technického nebo programového vytvořeného řešení, apod. Případně také zhodnoťte, zda software nebo zdrojové texty, které nevytvořil sám student, byly v ZP použity v souladu s licenčními podmínkami a autorským právem. Popište případnou publikační činnost a získaná ocenění související s řešením této ZP.

Komentář:
Modul, který student napsal, je funkční. Implementoval existující metody v modulu pro NfSen. V tomto ohledu nejde o nové metody. Nový může být zásuvný modul pro NfSen, avšak bylo by třeba ho otetovat s podobnými moduly a vyhodnotit jeho klady a zápory vůči těmto modulům. Takové srovnání není součástí práce.

Hodnotící kritérium:	*Způsob hodnocení - nehodnotí se*
8. Komentář o využitelnosti výsledků	

Popis kritéria:
Uveďte, zda hlavní výsledky ZP rozšiřují již publikované známé výsledky a/nebo přinášející zcela nové poznatky. Uveďte možnosti využití výsledků ZP v praxi.

Komentář:
Student splnil zadání. Navrhl a implementoval modul pro sondu NfSen. V tomto ohledu konstatuji, že modul je funkční, avšak jeho použitelnost v praxi vidím jako omezenou. Aby mohl být nasazen do reálného provozu, bylo by třeba ho dopracovat na úroveň, ve které nevyžaduje restartování kvůli nedostatku paměti, je kvalitně popsán, zdokumentován a jasně se vymezuje proti obdobným modulům.

Hodnotící kritérium:	*Způsob hodnocení - nehodnotí se*
9. Otázky k obhajobě	

Popis kritéria:
Uveďte případné dotazy, které by měl student zodpovědět při obhajobě ZP před komisí (body oddělte odrážkami).

Otázky:
1. Student na str. 42 píše: „...pro připojení nejspíše potřebuje certifikát serveru.“ Mohl by student objasnit, o jaký certifikát mělo u ssh připojení jít?
2. Testoval student svůj modul s obdobnými moduly? Pokud ano, jakými?

Hodnotící kritérium:	*Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):*
10. Celkové hodnocení	*75 (C)*

Popis kritéria:
Shrňte stránky ZP studenta, které nejvíce ovlivnily Vaše celkové hodnocení. Celkové hodnocení **nemusí** být aritmetickým průměrem či jinou hodnotou vypočtenou z hodnocení v předchozích jednotlivých kritériích 1 až 9.

Text hodnocení:
Při čtení práce nabývám dojmu, že nebyla psána s pečlivostí, jakou si kvalifikační práce zaslouží. Díky tomu obsahuje větší množství menších nepřesností a jazykových prohřešků, které spolu s chybějícím formalizmem snižují její kvalitu. Práci pana Petra Lessnera doporučuji k obhajobě a hodnotím ji stupněm C (dobře).

Podpis oponenta práce: