Posudek oponenta závěrečné práce

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

Student: Václav Hrabě
Oponent práce: Ing. Tomáš Zahradnický, Ph.D.
Název práce: Deminimizace a deobfuskace malware v jazyce JavaScript
Obor: Informační technologie

Datum vytvoření: 28. 5. 2017

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:
Téma deobfuskace považuji za náročnější téma vzhledem k tomu, že jde obecně o proces chodu proti jednosměrnosti obfuskace.

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:
Zadání práce považuji za splněné s menší výhradou, kterou spatřuji v nedostatečně provedené analýze, návrhu a zdrojích, ze kterých student čerpal.

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:
Přestože má práce závěr na straně 43, považuji její rozsah za hraničně naplňující požadavky na bakalářskou práci. Důvod spatřuji v tom, že práce obsahuje mnoho popisu, který je práci irelevantní (např. text na straně 5-14), a naopak relevantní části jsou rozpracovány jen mělce (např. sekce 2.5, 2.6 a 2.7).

Hodnotící kritérium:	*Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):*
4. Věcná a logická úroveň práce	*50 (E)*

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:
Logická stránka práce není zcela vpořádku. Práci považuji za obsahově nevyváženou - text na straně 5-14, část stran 20-22, jsou de-facto irelevantní. Práce se zabývá deminifikací a deobfuskací. Měla by proto zaměřit svoji analytickou část tímto směrem, ne na popis antivirových společností a druhů škodlivého softwaru. Pokud chtěl student zdůraznit, že mnoho škodlivého softwaru je dnes psáno v javascriptu, stačila by na to jedna stránka textu.

V práci nacházím věcné chyby typu: "Win32 - tento programovací jazyk ..." (str. 20, sekce 3.1.1.2), "skriptovacího jazyku shell neboli bash" (str. 21, sekce 3.1.1.5), "Knihovny jsou algoritmy ..." (str. 22 sekce 3.3).

Hodnotící kritérium:	*Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):*
5. Formální úroveň práce	*60 (D)*

Popis kritéria:
Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 14/2015, článek 3.

Komentář:
Typografická stránka práce je průměrná. Nacházím vytékající slova ze zrcadla stránky (např. str. 30, str. 31), občas špatně rozdělená slova (ma-lware namísto mal-ware např. na str. 6, sekce 2.1.1.4).

Jazyková stránka práce je podprůměrná. Nacházím hrubé chyby např. "programy ..., které byli výše vyjmenovány" (str. 8, sekce 2.3.1.1).

Úroveň formalizmů v práci bych očekával na daleko vyšší úrovni i vzhledem k tomu, že jde o manipulace s datovými strukturami. Pokud by se tímto student v práci zabýval hlouběji, bylo by možné provádět deobfuskaci daleko účinněji a například přejmenovávat třídy a lokální proměnné lidsky čitelnými názvy, jak tomu bývá u obdobného softwaru, anebo formálně popsat transformace abstraktního syntaktického stromu.

Hodnotící kritérium:	*Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):*
6. Práce se zdroji	*40 (F)*

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:
Práci se zdroji považuji za nedostatečnou. Student cituje sice 42 zdrojů, z nichž 16 je na webové stránky společnosti AVAST Software, spol. s r.o., 8 na webové stránky společnosti ESET spol. s r.o., apod. Chybí odkazy na základní literaturu v podobě vědeckých článků a knih. Například článek E. Collberg et al.: "Taxonomy of Obfuscating Transformations", který považuji za jeden ze základních článků o obfuskační taxonomii, by měl tvořit jeden ze základů, kterých by se měl student držet. Takto to vypadá, že čerpal jen ze stránek antivirových společností, avšak z žádné vědecké literatury k tématu.

Hodnotící kritérium:	*Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):*
7. Hodnocení výsledků, publikační výstupy a ocenění	*60 (D)*

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:
Výstupy práce považuji za podprůměrné. Proto hodnotím splnění zadání s menší výhradou.

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:
Výstupem práce je nástroj s omezenou použitelností.

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:
Sekce 4.6.2 na straně 33 uvádí binární operátory {+,-,*,/,%}. Kam se poděly binární operátory pro logický součin, součet, neekvivalenci a posun v jejich jednoduché (&, |, ^, >>, >>>, <<) a zdvojené formě (&& a ||), bez nichž nelze zpracovávat podmínky pro větvení programu? Kam se poděly operátory *=, /= a %= a operátory pro porovnávání <, >, <=, >=, ==, !=, ===?

V názvu práce se vyskytuje termín deminimizace, který by měl pocházet z anglického termínu minification. Proč deminimizace a ne deminifikace?

Hodnotící kritérium:	*Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):*
10. Celkové hodnocení	*60 (D)*

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:
I přes uvedené výtky bakalářskou práci pana Václava Hraběte doporučuji k obhajobě a hodnotím ji stupněm D (uspokojivě).

Podpis oponenta práce: