Posudek oponenta diplomové práce

Student: Daniš Daniel, Bc.
Téma: Detekce malware pomocí analýzy DNS provozu (id 18696)
Oponent: Ovšonka Daniel, Ing., UITS FIT VUT

1. Náročnost zadání **průměrně obtížné zadání**

Jedná se o zadání standardní obtížnosti.

2. Splnění požadavků zadání **zadání splněno s drobnými výhradami**

Výhrady možno najít k bodu 3. Navrhované řešení zahrnuje tři poměrně triviální a známé metody detekce DNS anomálií bez další přidané hodnoty autora. Analýza spočívá v sekvenčním vykonání jednotlivých algoritmů, přičemž částkové výsledky nemají žádný vplyv na další průběh analýzy.

3. Rozsah technické zprávy **je v obvyklém rozmezí**

Rozsah je v obvyklém rozmezí a činí zhruba 65 normostran.

4. Prezentační úroveň předložené práce **65 b. (D)**

Technická zpráva má vhodnou logickou strukturu, kapitoly na sebe logicky navazují. Teoretická část je zpracovaná přehledně, druhá část práce věnující se samotní implementaci v tomto ohledu zaostává. Kapitola 3 je popsána poněkud zmatečně, použité algoritmy jsou definované pomocí pseudokódu, avšak některé vyskytující se funkce nejsou vůbec v textu definované (napr. GetEdgeIndex v definici Algoritmu 2, strana 29). Bez znalosti referenčního článku jsou proto detaily pro čitatele hůře pochopitelné. Kapitola věnovaná testování se dostatečně nevěnuje analýze získaných výsledků a je popsaná spíše vágně až neformálně. Autor například vůbec neanalyzuje vliv konfigurace analyzátoru pro jednotlivé vzorky.

5. Formální úprava technické zprávy **70 b. (C)**

Po formální stránce obsahuje práce pouze drobné nedostatky, nejčastěji chybné používaní čárek a nesprávné citace. Gramatických chyb a překlepů je pouze minimum. Studen místy zavadí nové pojmy místo pojmů zaužívaných v bezpečnostních kruzích (např. "prístup založený na báze podpisov" v kapitole 1.2.2).

6. Práce s literaturou **60 b. (D)**

Práce s literaturou je na podprůměrné úrovni. Autor sice cituje významné články ze zkoumané oblasti, ale místy je složité určit či se jedná o převzaté části nebo o vlastní přínos autora. Problémem jsou aj citace u obrázku, kde není uváděn zdroj. Například obrázky 3.3, 3.4 jsou zjevně převzaté z literatury. Nevhodně odkazované jsou aj získané vzorky reálného síťového provozu, kde chybí přímé odkazy, přičemž soubory nejsou ani součástí přiloženého CD. Z tohoto důvodu nemožno vůbec vzorky dohledat a ověřit tak fakty uvedené v Kapitole 5.

7. Realizační výstup **55 b. (E)**

Realizační výstup sestává z jednoho scriptu v jazyku Python (cca. 700 řádků, samotná analýza dat tvoří cca 300 řádků). Zdrojový kód je dobře strukturovaný a rozsáhle komentovaný. Program je funkční na přiložených vstupech. Výsledek analýzy je závislý pouze na kvalitě databáze blacklistovaných domén. Při analýze reální vzorky dat zaznamenané za 20 sekund program končí kvůli omezení velkosti dotazu na službu VirusTotal.

8. Využitelnost výsledků

Jedná se o práci kompilačního charakteru, implementované jsou pouze už existující metody. Práce by byla použitelná až po dalším rozšíření při výzkumu nových metod detekce malwaru.

9. Otázky k obhajobě

- Jakou přidanou hodnotu poskytuje vaše řešení v porovnání s běžnými IDS/IPS systémy?
- Bylo by možné aplikovat tento přístup aj na analýzu sítového toku v reálném čase?
- V kapitole 3.2 uvádíte, že původní algoritmus detekce "domain-flux" byl upraven pro potřeby této práce. Můžete uvést hlavní rozdíly, kdeže z textu to není zřejmé?

10. Souhrnné hodnocení **60 b. uspokojivě (D)**

Jedná se o poměrně nevyváženou práci, která by mala potenciál ale hlavním nedostatkem navrhnuté finální řešení a samotná implementace. Implementované jsou pouze základní detekční algoritmy bez další přidané hodnoty přičemž úspešnost detekce závisí na výstupu ze služby VirusTotal. Vzhledem na výše uvedeným faktům navrhuji hodnocení **D (uspokojivě)**.

Prohlášení: Uděluji VUT v Brně souhlas ke zveřejnění tohoto posudku v listinné i elektronické formě.

V Brně dne: 8. června 2016

.....
podpis