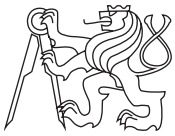
Katedra počítačové grafiky a interakce
Fakulta elektrotechnická
České vysoké učení technické v Praze



Posudek oponenta na diplomovou práci

Student: **Bc. Pavel Novotný**
Název: **Pseudonymizace dat pomocí HSM modulu**
Oponent: **Ing. Jan Kubr, Ph.D.**

1 Náročnost a další komentář k zadání

Hodnocení: 3 – průměrně náročné

Zadání hodnotím jako průměrně náročné. Práce zahrnuje analýzu problému, návrh aplikace a její implementaci.

2 Splnění zadání

Hodnocení: 1 – splněno

Zadání bylo splněno.

3 Rozsah písemné zprávy

Hodnocení: 2 – splňuje požadavky s menšími výhradami

Písemná zpráva splňuje požadavky kladené na diplomovou práci. Rozsah práce je zcela v pořádku. Analýza legislativy je doslovnou citací odstavců zákona. To by samo o sobě nebylo na závadu, ale odstavce jsou, dle autora, citovány namátkově (str. 3). Shrnutí legislativní části nepřinaší žádné zajímavé závěry. Shrnutí porovnání českého a německého zákona chybí. Analýza legislativy se tématu práce dotýká jen okrajově.

Analýza kryptografie také obsahuje řadu informací nesouvisejících s tématem práce. Příkladem je kap. 2.7.

V práci chybí podrobnější analýza rozdílů šifrování s HSM a šifrování prováděném na běžném současném hardware.

4 Věcná a logická úroveň práce

Hodnocení: 70 bodů C

Práce je přehledná. Dlouhé citace zákona jsou čtenářsky nezáživné, zvláště při absenci doprovodného vysvětlujícího textu.

Autor chybuje v popisu převzatých algoritmů. V algoritmu RSA chybně používá exponent (e) pro dešifrování (správně je exponent (d)). Není uvedeno co je (a) v následující rovnici (str. 30).

V případě Diffie-Hellman algoritmu autor zapomíná poslat prvek b druhé straně. Nicméně prvek b druhá strana v následující rovnici využívá.

Autor si v Závěru klade otázku, jaké vlastnosti by pseudonymizace měla mít a odpovídá si, že z pohledu práce to není důležité. S tímto tvrzením zcela nesouhlasím a tuto otázku pro práci považuji za zásadní.

Scénář předávání zpráv uvedený v kap. 7.5.1 mi připadá nesmyslný a nepřináší žádnou přidanou hodnotu. Minimálně podobných výsledků bychom docílili při použití běžného šifrování v e-mailové komunikaci (pro malé soubory).

5 Formální úroveň práce

Hodnocení: 90 bodů A

Práce obsahuje zanedbatelné množství chyb a překlepů.

6 Práce se zdroji

Hodnocení: 85 bodů B

Seznam zdrojů je správně citovaný a bohatý. Autor občas vynechá odkaz na použitou technologii např. Bcrypt a ARGON2 na str. 22.

7 Hodnocení výsledků, publikační výstupy a ocenění

Hodnocení: 75 bodů C

Autor dosáhl všech požadovaných výsledků. Shrnutí kapitol analytické části bohužel nelze využít pro rychlé seznámení s problematikou a přináší jen minimum zajímavých informací.

Autor v práci detekuje výrazný výkonnostní propad při šifrování s použitím HSM oproti šifrování na CPU. V práci chybí podrobná analýza tohoto jevu.

8 Komentář o využitelnosti výsledků

Aplikace ukazuje možnosti začlenění HSM do zpracování informací v big data prostředí.

9 Celkové hodnocení

Práce splňuje zadání s drobnými výhradami.

Celkově se jedná o práci průměrnou a její praktické použití je omezené. Analytická část přináší jen malé množství zajímavých informací.

Implementace je zajímavá a autor se musel vypořádat s řadou problémů.

Autor jen omezeně vyhodnocuje výsledky praktické implementace.

Konstatuji, že autor prokázal schopnost samostatné práce.

Práci hodnotím **75 body** (**C – dobře**).

10 Otázky k obhajobě

- Jaká je teoretická výkonnost HSM modulu (ve vztahu k této práci)?
- Pokud má datový zpracovatel zašifrované sloupce dat s unikátním solením, nebylo by jednodušší mu tyto sloupce nepředat?
- Opravdu je rychlost operací při použití asymetrické kryptografie primárně ovlivněna velikostí klíčů (kap. 2.6)?
- Proč jste pro definici terminologie použil starší standard, změnilo se v novém standardu něco (str. 23)?
- Je druhá varianta pseudonymizace vratná (str. 52)? Pokud ne, jedná se stále o pseudonymizaci?
- Jak si vysvětlujete nelineární časové chování výpočtu podle tab 7.2?

V Praze 15. června 2023 Jan Kubr