FAKULTA INFORMAČNÍCH TECHNOLOGIÍ ČVUT V PRAZE

Posudek oponenta závěrečné práce

Student:	Ondřej Voronecký
Oponent práce:	Ing. Josef Kokeš
Název práce:	Utajená komunikace použitím steganografie
Obor:	Bezpečnost a informační technologie

Datum vytvoření: 6. 6. 2020

Hodnotící kritérium:	*Způsob hodnocení – následující škálou 1 až 4:*
1. 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 dostatečně a v souladu se zadáním obsahově vymezuje cíle, správně je formuluje a v dostatečné kvalitě naplňuje. V komentáři uveďte body zadání, které nebyly splněny, posuďte závažnost, dopady a případně i příčiny jednotlivých nedostatků. Pokud zadání svou náročností vybočuje ze standardů pro daný typ práce nebo student případně vypracoval ZP nad rámec zadání, popište, jak se to projevilo na požadované kvalitě splnění zadání a jakým způsobem toto ovlivnilo výsledné hodnocení.

Komentář:
Zadání bylo splněno. Student provedl rešerši steganografie i samoopravných kódů, naimplementoval aplikaci pro ukrytí zprávy do obrázku a provedl experimenty, jak se jím ukrytá zpráva vyrovná s poškozením. Moje výhrady se týkají toho, že rešerše samoopravných kódů je velice stručná a ne zcela vyargumentovaná a testy schopností opravy podle mě nejsou spolehlivě provedené.

Hodnotící kritérium:	*Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):*
2. Písemná část práce	*80 (B)*

Popis kritéria:
Zhodnoťte přiměřenost rozsahu předložené ZP vzhledem k obsahu, tj. zda všechny části ZP jsou informačně bohaté a ZP neobsahuje zbytečné části. Dále 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. Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 26/2017, článek 3. Posuďte, zda student využil a správně citoval relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami. Zhodnoťte, zda převzatý software a jiná autorská díla, byly v ZP použity v souladu s licenčními podmínkami.

Komentář:
Textová stránka práce je nevyrovnaná. Kapitole o steganografii není mnoho co vytknout, je přiměřeně detailní, pěkně popsaná a dobře srozumitelná; jediné, co mi chybí, je aspoň několik odstavců o stegoanalýze. Naproti tomu kapitola o samoopravných kódech je nesmírně stručná, pro čtenáře neobeznámeného s problematikou obtížně srozumitelná a naprosto nedostatečně vysvětluje důvody pro volbu konvolučních kódů. Uvítal bych, kdyby z kapitoly o steganografii byla odstraněna podkapitola o formátu PNG, která je podle mě jak pro text tak pro implementaci zcela zbytečná, a ušetřený prostor byl věnován detailnějšímu popisu opravných kódů. Nedostatky v kapitole o implementaci a o testování popíšu v dalších sekcích.
Z hlediska logiky textu by bylo lepší, kdyby kapitola o opravných kódech byla první a na ni navázala kapitola o steganografii, která ji využívá.
Po jazykové stránce jsem narazil na více chyb, než by se mi líbilo. Jde zejména o psaní čárek mezi větami, psaní i/y ve spojeních jako "bity se rozmístili" a překlepy. Nepříjemné je míchání češtiny a angličtiny bez zjevného důvodu, např. název kapitoly 2.1 v angličtině (proč?) nebo výrazy jako "cover-objekt" a "stego-key".
Co se týče technické stránky, poměrně systematicky je odkazováno na různé typy objektů bez uvedení typu objektu a je ponecháno na čtenáři, aby se sám rozhodl, jestli "stegosystém 2.2" odkazuje na kapitolu, obrázek, definici nebo jestli to dokonce vůbec není odkaz. Seznam použitých zkratek mohl být seřazen podle abecedy. A jsem si dost jistý, že by se našlo i autoritativnější vysvětlení rušení ve fotografiích než odkaz na StackExchange.com.

Hodnotící kritérium:	*Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):*
3. Nepísemná část, přílohy	*80 (B)*

Popis kritéria:
Dle charakteru práce se případně vyjádřete k nepísemné části ZP. Například: SW dílo – kvalita vytvořeného programu a vhodnost a přiměřenost technologií, které byly využité od vývoje až po nasazení. HW – funkční vzorek – použité technologie a nástroje, Výzkumná a experimentální práce – opakovatelnost experimentů

Komentář:

Podobná nevyrovnanost je znát i z implementace. Student vytvořil aplikaci pro ukrývání zpráv do obrázků. Zvolil jazyk Python, protože "se v něm dají psát přenositelné programy", načež z něj volá binární knihovnu pro 64bitový Linux. Implementace je dle kapitoly 4.2 psaná s ohledem na bezpečnost (ovšem používá kryptograficky ne-bezpečný pseudonáhodný generátor pro zabezpečení zprávy před útočníkem) a přehlednost a budoucí rozšiřitelnost s využitím objektově orientovaného přístupu, ale s větvením na základě posloupnosti IFů namísto využití polymorfismu a vhodného návrhového vzoru. Chápu aplikaci jako nástroj pro demonstraci a ne hotové řešení a v tomto smyslu svůj účel plní, ale i tak mohla být napsána lépe.
Za nešťastné považuji to, že aplikace používá pevné parametry opravného kódu a z toho odvozuje délku dat. Praktičtější by bylo z velikosti obrázku, velikosti dat a bezpečnostního parametru (max. procento pozměněných pixelů) dopočítat co nejodolnější zabezpečovací kód.
Licence použitých komponent se zdají být v pořádku, jen bych měl obavy z použití Qt5.

Hodnotící kritérium:	*Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):*
4. Hodnocení výsledků, jejich využitelnost	*70 (C)*

Popis kritéria:
Dle charakteru práce zhodnoťte možnosti nasazení výsledků práce v praxi nebo uveďte, zda výsledky ZP rozšiřují již publikované známé výsledky nebo přinášející zcela nové poznatky.

Komentář:

Třetí nevyrovnaná komponenta. Provedená rešerše je odpovídající a vytvořená aplikace funguje dle očekávání. Velké výhrady však mám ke kapitole 5 (Analýza odolnosti metod a samoopravných kódů).
1) *Velice* chybí měření, co se stane, když obrázek z formátu PNG zkonvertujeme do ztrátového JPG. Domnívám se, že toto je zdaleka nejčastější poškození, ke kterému v praxi dochází, a bylo by dobré mít představu, jak si s tím zakódovaná zpráva poradí.
2) Chybí porovnání s blokovými kódy, zejména s běžně používanými kódy Reed-Solomon.
3) Nikde není detailně popsáno, jak přesně měření poškození proběhlo. Z přiložených souborů se zdá, že student vzal vstupní obrázky a ručně v nich poškodil zvolenou část obrázku. Považuji to za neprůkazné, pro jednotlivé míry poškození se liší struktura oblastí - ta měla mezi měřeními zůstat zachována. Použitý test může zdůrazňovat určitý typ chyb a tím znevýhodňovat jeden z algoritmů. Ideálně měly být oblasti generovány náhodně a mělo být provedeno více testů.
4) V celkovém vyhodnocení student tvrdí, že metoda Matrix embedding nabízí vyšší bezpečnost za cenu horší odolnosti. Bezpečnost však provedené testy vůbec nehodnotily. Postrádám měření, kolik bitů obrázku se změnilo při použití té které metody. Nabízí se navíc otázka, jak by výsledky vypadaly, kdyby pro obě metody byly nastaveny stejné podmínky, tzn. kódovala by se zpráva o stejné velikosti - jestli by náhodou metoda +-1 nedosáhla stejné nebo i vyšší bezpečnosti než metoda Matrix embedding.
Z těchto důvodů se obávám, že naměřené výsledky nejsou dostatečně spolehlivé. Aplikace samotná tak může být užitečná pro demonstraci konceptů a algoritmů, pro skutečné utajení zprávy bych ji ale raději nepoužíval.

Hodnotící kritérium:	*Způsob hodnocení – nehodnotí se*
5. Otázky k obhajobě	

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

Otázky:

1) Z jakého důvodu byly z porovnání vyloučeny blokové opravné kódy?
2) Jaký vliv na úspěšnost extrakce zprávy má pro obě metody konverze obrázku do formátu JPG a následně zpět do PNG?

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

Popis kritéria:
Shrňte stránky ZP, které nejvíce ovlivnily Vaše celkové hodnocení. Celkové hodnocení nemusí být aritmetickým průměrem či jinou hodnotou vypočtenou z hodnocení v předchozích jednotlivých kritériích. Obecně platí, že bezvadně splněné zadání je hodnoceno klasifikačním stupněm A.

Text hodnocení:

Práce je nesporně velmi zajímavá a prakticky demonstruje problematiku steganografie. V tomto smyslu ji považuji za jistě přínosnou. Jejím nedostatkem je nevyrovnaná pozornost věnovaná jednotlivým částem a zejména neprůkazné nebo nekompletně popsané vyhodnocení. Také kód nevypadá tak, jak bych si od výborného absolventa FITu představoval. Tyto faktory jsou pro mě hlavním důvodem k hodnocení C-dobře.

Podpis oponenta práce: