UNIVERZA V LJUBLJANI
FAKULTETA ZA RAČUNALNIŠTVO IN INFORMATIKO

Arnela Jušić

Skrivanje podatkov v slikovne datoteke z uporabo steganografskih metod

DIPLOMSKO DELO

VISOKOŠOLSKI STROKOVNI ŠTUDIJSKI PROGRAM PRVE STOPNJE RAČUNALNIŠTVO IN INFORMATIKA

MENTOR: prof. dr. Aleksandar Jurišić

Ljubljana, 2012

I Z J A V A O A V T O R S T V U

diplomskega dela

Spodaj podpisana Arnela Jušić, z vpisno številko **63040408**, sem avtor/-ica diplomskega dela z naslovom:

Skrivanje podatkov v slikovne datoteke z uporabo steganografskih metod

S svojim podpisom zagotavljam, da:

- sem diplomsko delo izdelal/-a samostojno pod mentorstvom (naziv, ime in priimek) prof. dr. Aleksandra Jurišića

- so elektronska oblika diplomskega dela, naslov (slov., angl.), povzetek (slov., angl.) ter ključne besede (slov., angl.) identični s tiskano obliko diplomskega dela

- soglašam z javno objavo elektronske oblike diplomskega dela v zbirki »Dela FRI«.

V Ljubljani, dne _____ Podpis avtorja/-ice:_____

Univerza *v Ljubljani*
Fakulteta *za računalništvo in informatiko*

Tržaška 25
1001 Ljubljana, Slovenija
telefon: 01 476 84 11
01 476 83 87
faks: 01 426 46 47
01 476 87 11
www.fri.uni-lj.si
e-mail: dekanat@fri.uni-lj.si

Št. naloge: 00205/2012
Datum: 02.04.2012

Univerza v Ljubljani, Fakulteta za računalništvo in informatiko izdaja naslednjo nalogo:

Kandidat: **ARNELA JUŠIĆ**

Naslov: **SKRIVANJE PODATKOV V SLIKOVNE DATOTEKE Z UPORABO STEGANOGRAFSKIH METOD**

IMAGE DATA HIDING USING STEGANOGRAPHIC METHODS

Vrsta naloge: Diplomsko delo visokošolskega strokovnega študija prve stopnje

Tematika naloge:

Delo predstavi osnove, potrebne za razumevanje najpomembnejših metod steganografije za skrivanje podatkov v slikah, npr. metoda vstavljanja, metoda zamenjave najmanj pomembnih bitov - LSB (Least Significan bit). Opravi naj se še primerjava metode zamenjave v različnih slikovnih formatih glede na ocenjevanje kakovosti slike s postopkom povprečne kvadratne napake - MSE (Mean Square Error) in postopkom razmerje med signalom in šumom - PSNR (Peak Signal to Noise Ratio).

Glavni cilji so (a) sistematizirati steganografske in stegoanalitske metode, (b) narediti primerjavo steganografskih metod za skrivanje podatkov v slikovnih datotekah, (c) podrobna analiza zgoraj omenjene LSB metode.

Mentor:

prof. dr. Aleksandar Jurišić

Dekan:

prof. dr. Nikolaj Zimic

Kazalo

Seznam ključnih besed

IT – Information Technology

LSB – Least Significant Bit

BMP - Bitmap

JPEG – Joint Photographic Experts Group

GIF – Graphics Interchange Format

PNG – Portable Network Graphics

MSE – Mean Squared Error

PSNR – Peak Signal-To-Noise Ratio

RGB – Red-Green-Blue

HTML – HyperText Markup Language

DCT – Discrete Cosine Transform

DFT – Discrete Fourier Transform

DWT – Discrete Wavelet Transform

SLSB – Selected Least Significant Bit

OPAP – Optimal Pixel Adjustment Process

TCP/IP – Transmission Control Protocol/Internet Protocol

PV – Pair of Values

Povzetek

Digitalna komunikacija je postala bistven del današnjega življenja in zato je v nekaterih primerih zaželeno, da komunikacija poteka skrivno. Na voljo sta dva načina: kriptografija in steganografija. Kriptografija je tehnika za zagotavljanje tajnosti komunikacije s pomočjo šifriranja. Včasih ni dovolj, da je skrita sama vsebina sporočila, ampak želimo prikriti tudi obstoj sporočila. Rezultat te potrebe je razvoj steganografije. V tej diplomski nalogi je najprej podan kratek pregled zgodovine steganografije in primeri uporabe v sodobnem svetu. Nato je predstavljena shema steganografske komunikacije in njeni deli. Cilj naloge je podrobneje predstaviti najbolj pogoste steganografske metode, ki se uporabljajo pri skrivanju podatkov v slikovne datoteke ter njihovo medsebojno primerjavo, da bi pokazali prednosti in slabosti vsake izmed njih. Primerjava predstavljenih metod je izvedena s pomočjo metod za ocenjevanje kakovosti slike srednja kvadratna napaka - MSE (Mean Square Error) in razmerje med signalom in šumom - PSNR (Peak Signal to Noise Ratio).

Abstract

Digitial communication has become an essential part of modern life and therefore, in some cases, its use is desired to stay secret. There are two available techniques: cryptogtaphy and steganography. Cryptography is a techique that provides secret communication with encryption. Sometimes hiding just the message content is not enough, so we try to conceal entire existence of the message. The need for this resulted in the development of steganography. The explanation how steganography has been developed throughout the history and today's application of it is explained at the beginning of thesis. Then the steganography communication scheme and it's components are presented. The main goal of this thesis is to thoroughly explain the most common image-hiding techniques and comparative analysis is made to demonstrate the strong and weak points of the proposed methods. The comparison of the presented methods is performed by using methods for image quality evaluation: MSE (Mean Square Error) and PSNR (Peak Signal To Noise Ratio).

1 Uvod

Potreba po zakrivanju sporočil obstaja praktično odkar obstaja človeštvo. Z razvojem in napredkom informacijsko-komunikacijskih tehnologij se je še povečala potreba po zasebnosti posameznika in po upravljanju z njegovimi osebnimi informacijami v digitalnem svetu. Tako sta se razvili dve sorodni tehniki za skrivno komuniciranje: *kriptografija* in *steganografija*. V tej diplomski nalogi se bomo osredotočili na slednjo. Cilj kriptografije je varovati skrivne podatke s šifriranjem, medtem ko je cilj steganografije prikriti obstoj teh podatkov. Steganografija je umetnost nevidnega komuniciranja, ki se izvaja s skrivanjem informacij v druge informacije. Ima različne uporabne namene, kot recimo prepričati nasprotnike, ki budno nadzirajo naše komunikacijske kanale (npr. internetne povezave, radijske in kabelske zveze), da sporočila sploh ni, ali pa vsaj, da ne gre za sporočilo, na katero bi morali biti pozorni (dnevno vremensko poročilo, družinske fotografije, ipd.). Da bi to uspešno storili, je pomembno izbrati ustrezno metodo za skrivanje, glede na naše zahteve.

V prvem delu diplomske naloge je predstavljen razvoj steganografije skozi zgodovino ter njena uporaba v sodobnem svetu. V glavnem delu diplome se bomo ukvarjali z različnimi steganografskimi metodami za skrivanje, s poudarkom na skrivanju v slikovne datoteke. Cilj je sistematizirati steganografske in stegoanalitske metode, približati terminologijo uporabnikom ter narediti primerjavo steganografskih metod, ki se uporabljajo pri skrivanju podatkov v slikovne datoteke. Podali bomo kriterije, na podlagi katerih se uporabnik lažje odloči za ustrezno steganografsko metodo.

Glede na to, da je metoda LSB zamenjave najbolj razširjena pri slikovnih datotekah, jo bomo analizirali in primerjali za slikovne formate BMP, GIF, PNG in JPG s pomočjo metod srednja kvadratna napaka - MSE (Mean Square Error) in razmerje signala in šuma - PSNR (Peak Signal to Noise Ratio), ki sta metodi za ocenjevanje kakovosti slike.

1.1 Kaj je steganografija

Termin *steganografija* je vpeljal Trithemius v knjigi Poligrafija in Steganografija [1], ki je ena zgodnejših del na področju kriptografije. Beseda izvira iz grškega jezika in je sestavljena iz dveh grških besed: steganos – prikrit in graphein – pisanje. Steganografija je kriptološka veda, ki se ukvarja s prenašanjem tajnih sporočil, na način, da prikrije oziroma zakrije obstoj sporočila. Steganografija v računalniškem svetu je mlada veja, saj se je začela intenzivneje razvijati šele leta 1996, ima pa široko polje uporabnosti.

Za razliko od kriptografije oz. šifriranja podatkov, katere cilj je zaščititi podatke, se steganografija ukvarja s prikrivanjem obstoja podatkov. S stališča uporabljenih tehnik sta kriptografija in steganografija precej sorodni tehniki. V kontekstu prenosa podatkov v IT svetu bi s pomočjo kriptografije podatke pretvorili v neberljivo obliko, medtem ko bi s steganografijo prikrili obstoj podatkov.

1.2 Steganografija skozi zgodovino

Kot smo že omenili, začetki steganografije segajo daleč v zgodovino. Prvi zabeleženi zapisi uporabe steganografije segajo v 440 leto pred našim štetjem. Herodot v svojemu delu »Histories« omenja dva primera steganografije :

- Demaratus je opozoril Šparto o prihajajočem napadu na Grčijo tako, da je napisal skrivno sporočilo neposredno na leseno podlago voščene ploščice, pred nanosom voščenega sloja. Takšna voščena ploščica se je zdela prazna in neuporabljena, tako da ni pritegnila pozornost med inšpekcijskim pregledom. Voščene ploščice so se uporabljale kot površine za pisanje za večkratno uporabo.
- Zgodba [2], ki pripoveduje o sužnju, ki ga je njegov gospodar Histacijus poslal v jonsko mesto Milet s tajnim sporočilom tetoviranim na glavo. Počakali so, da mu lasje zrastejo in prikrijejo sporočilo, nato pa so ga

poslali nazaj v Milet, da prenese sporočilo mestnemu vladarju Aristagorasu.

Naštetih je še nekaj najbolj znanih primerov iz zgodovine:

- Kitajci so svoja sporočila skrivali tako, da so koščke svile, na katere so napisali besedilo, pomočili v vroč vosek iz katerega so nato oblikovali majhne kroglice. Te kroglice je nato pogoltnil osel [3].
- Eden izmed najbolj znanih načinov za skrivanje besedila, ki se uporablja tudi danes, je uporaba skrivnega črnila, ki je pri normalni sobni temperaturi nevidno. Pri segrevanju papirja pa se besedilo obarva rjavo.
- Med drugo svetovno vojno so Nemci razvili tehnologijo mikropik (angl. Microdot) [4], ki fotografijo zmanjša na velikost 1mm. Postopek ustvarjanja je preprost, ampak je potrebna specializirana oprema.
- Leta 1550 je italijanski fizik, matematik in astrolog Girolamo Cardano uvedel preprosto mrežo za pisanje skrivnih sporočil. Vsak prejemnik ima svoj list papirja z luknjami. Ko kardansko mrežo (list papirja z luknjami) položimo na list papirja z na videz nedolžnim besedilom in zapolnimo izrezane dele, dobimo skrivno sporočilo. Kardanska mreža predstavlja enega izmed prvih primerov steganografskega ključa, ki si ga morata sogovornika izmenjati [5].
- Obstajajo govorice, da je v 80-ih letih Margaret Thatcher, tedanja predsednica vlade v Veliki Britaniji, jezna zaradi uhajanja zaupnih informacij iz predsedniške pisarne, v urejevalnike besedil dala programirati identiteto pisca v razmike in na ta način odkrila nelojalne ministre [6].

1.3 Steganografija danes

Danes se steganografija uporablja tako iz zakonitih kot tudi iz nezakonitih razlogov.

Zakonito uporabo v veliki meri predstavlja uporaba digitalnega vodnega žiga za zaščito avtorskih pravic in lastništva multimedijskih datotek. Poleg tega je mogoče s pomočjo steganografije dodati različne opombe mulimedijskim datotekam, brez spreminjanja njihovega formata. Še ena pomembna uporaba steganografije je vstavljanje podatkov o pacientu v medicinske slike, ker na ta način ne prihaja do težav z usklajevanjem pacientovih zapisov in slik. Nazadnje je najbolj logična uporaba steganografije za varovanje zaupnih podatkov in tajnosti pomembnih informacij (kot smo omenili že v primeru Thatcherjeve), ter njihova zaščita pred morebitno sabotažo, tatvinami ali nepooblaščenim dostopom.

Zaradi svoje posebnosti kot sredstvo skrivnega komuniciranja, steganografija pogosto najde mesto onstran zakona. Nezakonita uporaba steganografskih metod je pogosto povezana s krajo zaupnih podatkov (npr. v industriji in v poslovnem sektorju), finančnimi goljufijami, izmenjavo otroške pornografije, krajo identitete, tihotapljenjem ter terorizmom. Tako so bili na primer, po terorističnih napadih 11. septembra 2001, napisani številni članki, v katerih je predstavljena teorija o komunikaciji med člani teroristične organizacije s pomočjo steganografije.

Napredek v razvoju steganografije je spodbudil razvoj komplementarnega področja – stegoanalize, ki se je začelo razvijati hitreje po terorističnih napadih 11. septembra 2001. Stegoanaliza se ukvarja z razvojem metod za ugotavljanje prisotnosti skrivnih sporočil in na koncu, z odkrivanjem vsebine teh sporočil. Več o stegoanalizi bomo povedali v poglavju 5.

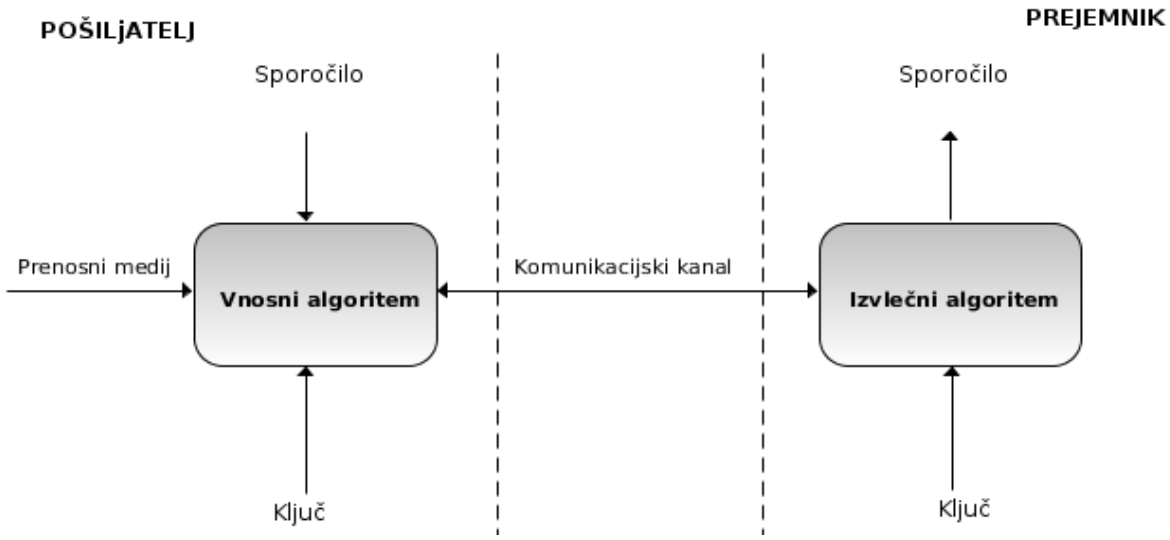
2 Shema steganografske komunikacije

Prvo neformalno definicijo steganografske sheme je podal Simmons [7] pri opisu t.i. problema zapornikov. Dva zapornika, Alice in Bob, sta pod nadzorom upravnika Eve. Upravnik bo dovolil komunikacijo med njima, pod pogojem, da lahko preverja sporočila. Če se upravniku sporočilo zdineškodljivo, ga lahko posreduje naprej, lahko pa obliko sporočila spreminja, z namenom odstranjevanja morebitnega skritega sporočila. Ta model je prikazan na sliki 1.

Po Coxu [8] steganografsko komunikacijo razdelimo na:

- komunikacijski kanal,
- prenosni medij,
- vnosni in izvlečni algoritem,
- steganografski ključ,
- skrivno sporočilo.

Omenjene enote bomo podrobneje opisali v nadaljevanju.



Slika 1: Shema steganografske komunikacije.

2.1 Komunikacijski kanal

Mediju, preko katerega se izmenjujejo informacije od pošiljatelja do prejemnika, bomo rekli *komunikacijski kanal*. Pri uporabi steganografije v spletu fizični komunikacijski kanal večinoma ne vsebuje šumov oziroma napak pri prenosu, saj to preprečujejo internetni protokoli. Recimo da, obstaja tretja oseba, ki prestreza naše podatke in skuša ugotoviti, ali poteka med sogovornikoma izmenjava skrivnih informacij. Glede na pristojnosti in odločitve poznamo tri vrste posegov v komunikacijo, ki jih bomo v nadaljevanju podrobneje opisali.

Pasivni poseg v komunikacijo

Pri *pasivnem posegu* v komunikacijo, nadzornik oz. napadalec nima pristojnosti za spreminjanje vsebine sporočila, lahko pa samo prepreči ali dovoli njegovo dostavo. V primeru pasivnega poseganja napadalec testira vsako posredovano sporočilo in išče sledi skrite komunikacije. Če je sporočilo sumljivo, lahko prekine komunikacijo med sogovornikoma.

Aktivni poseg v komunikacijo

O *aktivnem posegu* govorimo takrat, ko napadalec vsa sporočila, ki potekajo med sogovornikoma, spreminja. Pri komuniciranju preko svetovnega spleta pride do aktivnega posega takrat, ko tretja oseba ne zaupa v učinkovitost stegoanalitskega orodja. Ne glede na negativne rezultate o prisotnosti skritega sporočila, bo napadalec vseeno spremenil vsebino celotnega sporočila in na ta način uniči potencialno skrito sporočilo. Če steganografski algoritem ni prilagojen na aktivne posege v komunikacijo, obstaja velika verjetnost, da bo pri spreminjanju sporočila prišlo tudi do popačenja ali izgube skrivnega sporočila.

Zahrbtni poseg v komunikacijo

Pri *zahrbtnem posegu* v komunikacijo gre predvsem za poskus odkritja oseb, ki med seboj skrivno komunicirajo. To lahko vključuje napadalčevo oponašanje enega izmed sogovornikov. Takšna vrsta napada je prisotna predvsem v steganografiji z javnim

ključem. Steganografski ključ je pri tem postopku znan vsem in vsakdo lahko pridobi skrito sporočilo, ki je šifrirano s pomočjo javnega ključa. Vendar lahko poslano sporočilo dešifrirajo samo tisti, ki imajo pošiljateljev zasebni ključ. Čeprav je steganografski ključ znan, je težko razlikovati med šifriranim sporočilom in pridobljenim zaporedjem naključnih bitov.

2.2 Prenosni medij

Prenosni medij je nosilec skrivnega sporočila. Vanj se skrijejo podatki, tako da pride do komaj opaznih sprememb. V praksi to pomeni, da mora prenosni medij izgledati in delovati enako kot prvotni nespremenjen medij ter ne sme biti sumljiv. Mediji, ki jih najpogosteje uporabljamo pri tem, so slikovne datoteke, zvočne datoteke, tekstovne datoteke in omrežni promet [9]. Slikovne datoteke so najbolj primeren medij za skrivanje zaradi svoje velikosti in pogostosti uporabe. V poglavju 3 bomo podrobneje razložili potek skrivanja pri uporabi različnih slikovnih formatov.

2.3 Vnosni in izvlečni algoritem

Steganografski algoritem ni nič drugega kot matematična formula, ki skrivno sporočilo zapiše v izbran prenosni medij in ga iz medija zna ponovno izločiti. Steganografski algoritem pove na kakšen način se bo skrivno sporočilo zapisalo v izbran prenosni medij. Glavni cilj novih steganografskih algoritmov je razviti statistično nezaznavne metode z visoko steganografsko kapaciteto.
Po Coxu [10] lahko kvaliteto in uspešnost algoritma ocenimo na podlagi lastnosti, ki ji bomo opisali v nadaljevanju.

Kapaciteta in učinkovitost vnosnega algoritma

Kapaciteta vnosnega algoritma je največje število bitov, ki jih je mogoče skriti v prenosni medij. Na primer, če želimo skriti skrivno sporočilo z uporabo metode LSB

(zamenjava najmanj pomembnega bita) v sivinsko sliko, je zmogljivost vstavljanja (v bitih) število slikovnih pik.
Pomemben koncept v steganografiji je *učinkovitost vnosnega algoritma*, ki je definiran kot število bitov skrivnega sporočila, ki so skriti na enoto popačenja.

Možnost odkrivanja skritih podatkov

Pri *odkrivanju skritih podatkov* je potrebno omeniti pojem *steganografska kapaciteta*. Je največje število bitov, ki jih lahko skrijemo v prenosni medij na določeno enoto podatkov, tako da je verjetnost odkrivanja skritih podatkov zanemarljiva.
Najlažji način, da se izognemo odkrivanju skritih podatkov, je zmanjšanje količine podatkov, ki jih želimo skriti. Zmanjšanje količine podatkov bo zmanjšalo vpliv vgrajevanja in posledično nastane bolj varna steganografska shema.
Učinkovitost vnosnega algoritma za skrivanje se meri glede na težavnost odkrivanja skritih podatkov. Kakovost skrivanja je močno odvisna od vnosnega algoritma za skrivanje podatkov ter od lokacije skritih podatkov.

Učinkovitost pridobivanja skritih podatkov

Če so skriti podatki zaznani, oziroma če je zaznan njihov obstoj, še ne pomeni, da je že znana vsebina skritega sporočila. Pomembno je izbrati močan stego-ključ, ker bi v nasprotnem primeru napadalec poskušal prebrati skrito sporočilo z uporabo vseh možnih stego-ključev. Ko pride do smiselnega sporočila, je pravilen ključ odkrit. Takšen tip napada ne bo deloval, če je bilo sporočilo pred skrivanjem še šifrirano. Na splošno velja, da je dobra praksa šifriranje sporočila pred skrivanjem. Kriptografski ključ lahko izpeljemo iz stego-ključa ali pa se izbere samostojno.

2.4 Steganografski ključ

Steganografski ključ je skrivni ključ s pomočjo katerega vnašamo skrivno sporočilo v prenosni medij. Glavna naloga steganografskega ključa je kreiranje naključnih vrednosti, ki določajo način, kako se biti skrivnega sporočila zapišejo v prenosni

medij. Pri večkratni uporabi steganografije, moramo paziti, da ne uporabimo istega stego-ključa večkrat, saj le-ta vsakič določa isto pot za zapis skrivnega sporočila v prenosni medij. Najvarnejši ukrep za izognitev takšni situaciji je periodična menjava steganografskega ključa. Ključ z določeno dobo veljavnosti imenujemo sejni ključ, ki je različen za vsako skrivanje [11]. Glede na izbiro steganografskega ključa lahko steganografijo delimo na naslednje tri vrste: klasično steganografijo, steganografijo s skrivnim ključem ter steganografijo z javnim ključem.

Klasična steganografija

Klasična steganografija (pure steganography) je definirana kot steganografski sistem, ki ne zahteva izmenjave stego-ključa pred začetkom komunikacije. Ta tip steganografije je najmanj varna metoda skrivanja, saj pošiljatelj in prejemnik samo domnevata, da se nobena druga stranka ne zaveda skrivnega sporočila. Pri uporabi odprtega sistema kot je medmrežje, ta domneva ne drži.

Steganografija s skrivnim ključem

Steganografski sistem, ki temelji na uporabi enega skrivnega ključa (stego-ključa), ki ga morajo sogovorniki izmenjati pred začetkom komunikacije, je definiran kot s*teganografija s skrivnim ključem.* Vnos skrivnega sporočila se izvede z uporabo skrivnega ključa. Samo stranke, ki ga poznajo lahko izvedejo obraten proces in preberejo skrivno sporočilo. Za razliko od čiste steganografije, za katero je značilno nevidno komuniciranje zaradi odsotnosti stego-ključa, steganografija s skrivnim ključem uporablja stego-ključ in je zato bolj občutljiva na prestrezanje. Prednost takšne vrste steganografije je da, tudi če je sporočilo prestreženo, lahko skrivno sporočilo preberejo samo pooblaščene stranke.

Steganografija z javnim ključem

Steganografija z javnim ključem uporablja koncepte iz kriptografije z javnimi ključi, ki je opisana v razdelku 7.1. Definirana je kot steganografski sistem, ki uporablja dva stego-ključa – javnega in zasebnega. Pošiljatelj uporabi javni ključ za šifriranje

skrivnega sporočila pri vnosu v prenosni medij, skrivno sporočilo pa lahko dešifrira samo prejemnik z zasebnim ključem.

2.5 Skrivno sporočilo

Po Coxu *skrivno sporočilo* predstavljajo podatki, ki jih želimo poslati sogovorniku tako, da o njihovem obstoju ne posumi nihče. Pri najpomembnejšem delu steganografskega procesa, vnosu skrivnega sporočila v prenosni medij, je potrebno izbrati ustrezno metodo vnosa sporočila v prenosni medij.

3 Steganografske metode vnosa

Naslednja formula podaja zelo splošen opis delov steganografskega procesa [12]:

prenosni_medij + skrivni_podatki + stego_ključ = stego_medij

V tem kontekstu bomo privzeli, da je *prenosni medij* datoteka, v katero bomo skrili podatke, ki so lahko šifrirani tudi s pomočjo stego-ključa. Tako nastali datoteki bomo rekli *stego-medij*. Prenosni medij so ponavadi avdio datoteke in slike. V tej diplomski nalogi se bomo osredotočili na slednje.

Najpomembnejši del steganografskega procesa, način vnosa skrivnega sporočila v prenosni medij, je odvisen od formata datoteke. Glede na različne načine spreminjanja datotek, da bi znotraj njih skrili podatke, lahko steganografske metode razdelimo na:

- metode vstavljanja,
- metode zamenjave najmanj pomembnih bitov,
- metode generiranja.

Preden začnemo obravnavati načine vnosa skrivnih sporočil, si je vredno na hitro pogledati različne slikovne formate in načine shranjevanja.

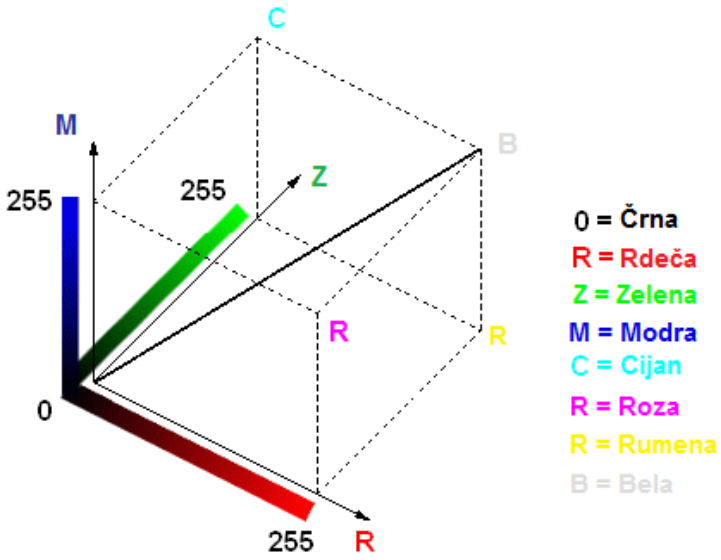
3.1 Slikovne datoteke

Slikovna datoteka je binarna datoteka, ki vsebuje predstavitev barve ali jakosti svetlobe vsakega slikovnega elementa (piksel), ki tvori sliko [13].

Slike običajno uporabljajo bodisi 8-bitni bodisi 24-bitni zapis barve. Pri uporabi 8-bitnega barvnega zapisa je število različnih barv 256, ki tvorijo paleto za to sliko. Vsaka barva je določena z 8-bitno vrednostjo. 8-bitni sistem je značilen za format GIF (Graphics Interchange Format).

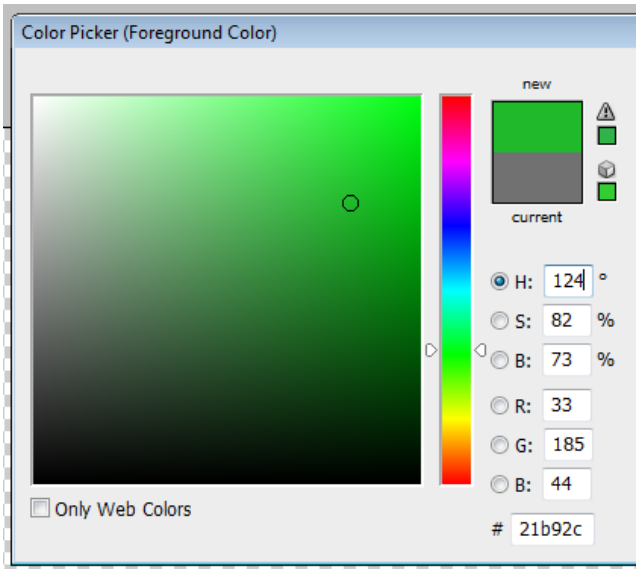
24-bitna barvna shema uporablja 24-bitne barve in omogoča veliko boljši nabor barv (2^{24} možnih barv). V tem primeru je vsak piksel predstavljen s tremi bajti, vsak bajt pa

predstavlja intenzivnost treh osnovnih barv: rdeče, zelene in modre (RGB), v tem zaporedju (slika 2). Če je vrednost vseh komponent RGB 0%, nastane črna barva, medtem ko pri vrednosti 100% vseh komponent nastane bela barva.



Slika 2: RGB kocka.

Na sliki 3 vidimo, da je izbran odtenek zelene barve, ki ima vrednosti rdeče barve 33, zelene 185 in modre 44.



Slika 3: Odtenek zelene barve.

Binarni zapis navedenih RGB komponent je 0010001 10111001 00101100. Najbolj levi biti se imenujejo *najbolj pomembni biti* in sprememba njihove vrednosti zelo vpliva na celotno vrednost. Analogno so najbolj desni biti *najmanj pomembni biti* ali LSB (Least Significant Bit) in najmanj prispevajo k vrednosti števila [14]. Sprememba njihovih vrednosti ne vpliva močno na vrednost števila oziroma spremembo barve in ravno to lastnost izkorišča metoda zamenjave najmanj pomembnih bitov, ki jo bomo opisali v nadaljevanju.

Velikost slikovne datoteke je neposredno povezana s številom pik in z zrnatostjo barvne definicije. Zato moramo biti pri steganografiji pozorni, da ne pride do naglega povečanja velikosti slikovne datoteke, ker bi to povzročilo sume. Najbolj priljubljeni formati slikovnih datotek so BMP (Bitmap), GIF (Graphic Interchange Format) in JPEG (Joint Photographic Experts Group) oz JPG, vendar zaradi svojih lastnosti pri stiskanju niso vsi enako primerni za uporabo steganografije. Glede na to, ali pri stiskanju pride do izgube podatkov, lahko stiskanje opredelimo kot stiskanje brez izgube ali pa z izgubo.

Stiskanje brez izgube

Stiskanje brez izgube zmanjša velikost slike z uporabo različnih algoritmov in možno jo je vrniti nazaj v prvotno stanje. Predstavnika te vrste stiskanja sta formata GIF in BMP. Glede na to, da se pri stiskanju brez izgube ohranijo vse informacije, so slikovni formati, ki uporabljajo ta način stiskanja, veliko primernejši za steganografijo. Pri tovrstnih formatih se uporabljajo steganografske metode vnosa, ki delujejo znotraj *prostorske domene* [15]. To pomeni, da se izvaja neposredna manipulacija pikslov v sliki. Algoritmi prostorske domene so dokaj enostavni in hitri, glavni predstavnik pa je metoda zamenjave najmanj pomembnih bitov, ki jo bomo opisali v nadaljevanju.

Stiskanje z izgubo

Stiskanje z izgubo tudi zmanjša originalno velikost slike, vendar na račun kakovosti slike. Temelji na tem, da se lahko nekatere manj pomembne informacije zavrže. Zaradi narave človeškega očesa je za samo kvaliteto slike veliko bolj pomembna

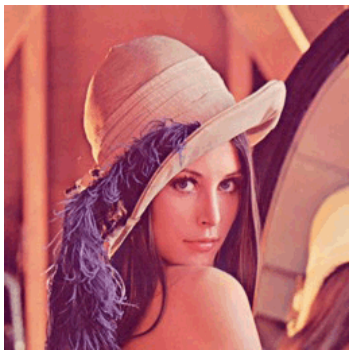
informacija o svetlosti kot o barvi. Takšno sliko je nemogoče vrniti v prvotno stanje, ker so deli slike trajno odstranjeni. Predstavnik stiskanja z izgubo je format JPG.
Za stiskanje uporablja *diskretno kosinusno transformacijo*, ki jo bomo podrobneje razložili v razdelku 3.3.3.1. Format JPG se smatra za slab format za steganografijo, ker je slika, ki je obnovljena iz stiskanja sicer blizu originalni sliki, nikoli pa ni identična. Zaradi lastnosti izgubljanja informacij, delujejo steganografske metode vnosa znotraj frekvenčne domene in so zato bolj robustne in odporne na napade kot steganografske metode znotraj prostorske domene.

3.2 Metoda vstavljanja

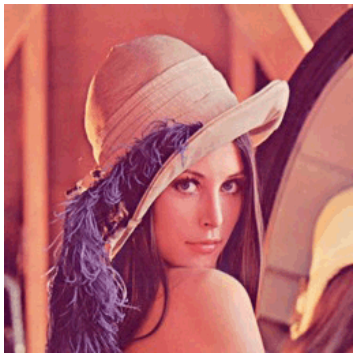
Metoda vstavljanja je dokaj preprosta metoda, ki neposredno vstavlja skrivne podatke v nosilno datoteko. Vsaka datoteka vsebuje oznake, ki določajo začetek in konec datoteke, kar pomeni, da lahko vstavimo poljubno mnogo skrivnih podatkov pred začetek oz. za konec datoteke, ne da bi vplivali na delovanje datoteke. Glavna težava pri tej metodi je, da lahko znatno poveča velikost nosilne datoteke, kar hitro povzroči sume.
Najenostavnejši primer je uporaba skrivnih atributov v urejevalniku besedil Microsoft Word, ki dovoljuje skrivanje besedila z uporabo skrivne pisave [16]. Ta enostavna metoda se uporablja za shranjevanje opomb in referenc pri kreiranju dokumenta.
Še en dober primer metode vstavljanja je skrivanje sporočila znotraj komentarja v dokumentu HTML. Komentarji so zanimivi in uporabni ker so obiskovalcu nevidni, brskalnik pa prepozna obliko komentarja in ga ne prikaže. Ne smemo pretiravati z dolžino vstavljenega besedila, ker so datoteke HTML običajno majhne velikosti.
Tudi številne oblike zlonamerne kode uporabljajo metodo vstavljanja kot npr. trojanski konj in računalniški virusi.
Za uporabo metode vstavljanja v slikovne datoteke ne potrebujemo nobene aplikacije. Vstavljanje lahko izvedemo z enostavnim sistemskim ukazom za kopiranje (primer za operacijski sistem Windows):

copy LenaOriginal.gif /b + sporocilo.txt /b LenaStego.gif

kjer */b* pomeni binarni zapis. Na sliki 4 vidimo, da sta prvotna in stego-slika enaki, kar pomeni, da vstavljeno sporočilo ne vpliva na kakovost slike. Metoda je odporna na vizualne napade in napade s histogramom, ker je sporočilo skrito za oznako EOF. Ni pa odporna na katerekoli načine spreminjanja slike in če primerjamo velikost prvotne in stego-slike, takoj opazimo, da se je velikost stego-slike povečala za toliko, kolikor je velikost datoteke sporocilo.txt (24KB).



LenaOriginal.gif
(velikost 47KB)



LenaStego.gif
(velikost 71 KB)

Slika 4: Slika Lene pred in po uporabi metode vstavljanja.

3.3 Metoda zamenjave najmanj pomembnih bitov

Metoda zamenjave najmanj pomembnih bitov, v nadaljevanju *LSB zamenjava*, spremeni LSB (najmanj pomemben bit) v bit iz skrivnega sporočila.
Najpomembnejši vprašanji pri LSB zamenjavi sta izbira slike in izbira formata: 8, 16 ali 24 bitna slika. Večina strokovnjakov predlaga uporabo 8-bitnih sivinskih slik, saj je njihova paleta precej manj raznolika kot barvna, tako da je LSB zamenjavo zelo težko odkriti s človeškim očesom. To je najbolj pogosta metoda skrivanja pri slikovnih datotekah, saj so le-te zelo razširjene na medmrežju.

3.3.1 LSB zamenjava in format BMP

Na primeru pikslov 24-bitne slikovne datoteke bomo pokazali, kako poteka zamenjava manj pomembnih bitov. Na sliki 5 so predstavljeni trije sosednji piksli, v katere želimo skriti število 365 (binarno: 101101101).

10010101	00001101	11001001
10010110	00001111	11001010
10011111	00010000	11001011

Slika 5: Binarni zapis treh pikslov.

10010101	00001100	11001001
10010111	00001110	11001011
10011111	00010000	11001011

Slika 6: Zamenjava manj pomembnih bitov.

Rezultat skrivanja je predstavljen na sliki 6. Opazimo, da smo uspešno skrili 9 bitov in pri tem je prišlo do spremembe samo pri 4 najmanj pomembnih bitih (v krepki pisavi). V povprečju je potrebno spremeniti približno polovico bitov ker je verjetnost, da je LSB pravilno postavljen 50% (vrednost je lahko ali 0 ali 1). Iz zgornjega primera je razvidno, da lahko v vsakem pikslu zamenjamo tri bite. Na primer, če imamo sliko velikosti 800 x 600 pikslov (480 000 pikslov), lahko shranimo 1 440 000 bitov (480 000 x 3) oziroma 180 000 bajtov. Prednost LSB zamenjave je, da se velikost datoteke ne spremeni.

Najlažja implementacija LSB zamenjave je pri formatu BMP, ker uporablja stiskanje brez izgube. Velika slabost uporabe tega formata je, da se dandanes preko medmrežja redko prenašajo slike BMP velikosti 800 x 600 in bi takšna dejanja lahko povzročila sume. Zato so se razvile metode LSB zamenjave, ki uporabljajo tudi druge slikovne formate, kot so GIF ali JPG.

3.3.2 LSB zamenjava in format GIF

Format GIF je rastrski slikovni format. Podpira do 8 bitov na piksel, kar pomeni, da lahko prikaže največ 256 barv iz 24-bitne barvne palete RGB. GIF stisne podatke v dveh korakih. V prvem koraku poišče vse barve, s katerimi so obarvani piksli na sliki, v drugem pa dobljene barve razporedi v barvno paleto in jih oštevilči. Barve so v paleti običajno sortirane padajoče glede na pogostost uporabe, da bi se zmanjšal čas iskanja barve.

Pri uporabi LSB zamenjave v formatu GIF lahko pride do težav, ker lahko sprememba najmanj pomembnega bita povzroči spremembo barve, glede na to, da se je spremenil indeks barve v barvni paleti. Ena izmed možnih rešitev je, da se razvrsti paleto tako, da se barvne razlike med zaporednimi barvami zmanjšajo. Druga rešitev je, da se dodajo nove barve, ki so zelo podobne obstoječim barvam v paleti. To zahteva, da ima slika čim manj edinstvenih barv, zato je zelo pomembna izbira primerne slike. Katerokoli spreminjanje barvne palete pušča sledi in olajša odkrivanje skritih sporočil. Optimalna rešitev je uporaba sivinskih slik, ker se 8-bitna sivinska slika GIF sestoji iz 256 različnih odtenkov sive barve. Spremembe med barvami so zelo postopne in jih je zato težko odkriti.

3.3.3 LSB zamenjava in format JPG

Pri formatu JPG se slika razdeli na 8x8 bloke, za vsako komponento barve posebej. Cilj je najti blok, v katerem bo sprememba vrednosti piksla najmanjša. Če je vrednost prevelika, se blok deli na naslednjih 8x8 podblokov, dokler vrednost ni dovolj nizka. Vsak blok ali podblok, je transformiran v 64 diskretnih kosinusnih transformacijskih koeficientov, ki so približek osvetlitvi in signalu barve tega dela slike.

Kot smo že omenili v poglavju 2, se metode vnosa v frekvenčni domeni uporabljajo pri obdelavi signalov in slik. Sestavljene so iz matematičnih postopkov, ki pretvorijo podatke iz ene domene v drugo. V novi domeni se podatki lažje obdelujejo, po obdelavi pa se pretvorijo nazaj v prvotno domeno. Pri slikah se frekvenčne transformacije uporabljajo predvsem za stiskanje, ostrenje, za vstavljanje vodnega žiga in njihovo obdelavo.

Obstajajo trije osnovni tipi frekvenčne transformacije, ki se uporabljajo pri skrivanju podatkov [17]:

- Diskretna kosinusna transformacija - DCT (*Discrete Cosine Transformation*),
- Diskretna Fourierjeva transformacija - DFT (*Discrete Fouirier Transformation*),
- Valčna transformacija (*Discrete Wavelet Transformation*).

3.3.3.1 Diskretna kosinusna transformacija

Diskretna kosinusna transformacija, v nadaljevanju DCT, je ena izmed metod, ki pretvorijo sliko iz prostorske v frekvenčno domeno [18]. DCT je pravzaprav posplošitev Diskretne Fourierjeve transformacije, s to razliko, da uporablja samo realni del. Najpogostejša uporaba je pri stiskanju slik. Sestavljata jo dva postopka: transformacija signalov iz prostorske v frekvenčno domeno, ter kvantizacija. Slikovni format JPG uporablja dvodimenzionalno diskretno kosinusno transformacijo, ki pretvori signal v frekvenčno predstavitev, tako da, združi piksle v 8x8 slikovnih blokov in jih transformira v 64 diskretnih kosinusnih koeficientov. Vsak koeficient DCT *F(u,v)* iz 8x8 bloka pikslov *f(x,y)* je podan s formulo:

$$F(u,v)=\frac{1}{4}C(u)C(v)\left[\sum_{x=0}^{7}\sum_{y=0}^{7}f(x,y)*\cos\frac{(2x+1)u\pi}{16}\cos\frac{(2y+1)v\pi}{16}\right]$$

kjer

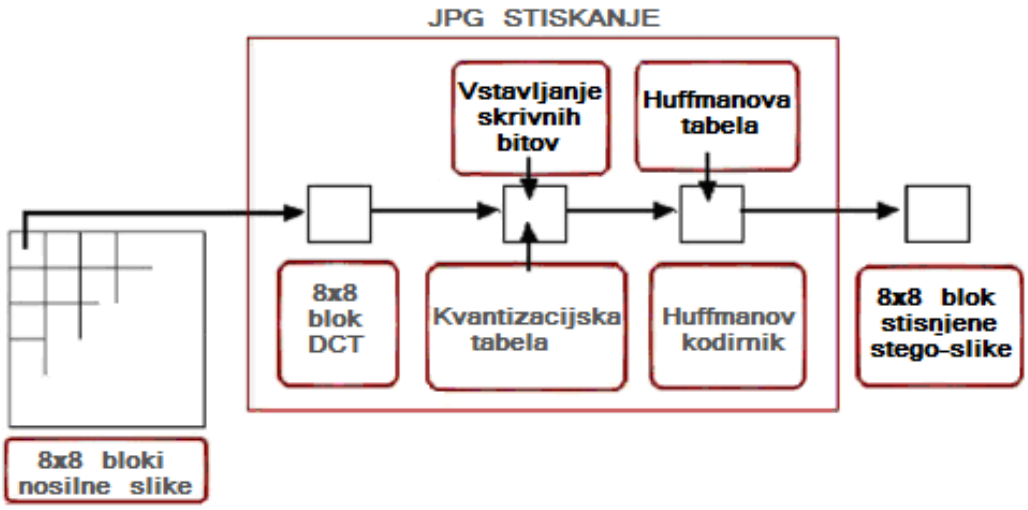
$$C(u,v)=\begin{cases}\frac{1}{\sqrt{2}} \rightarrow u,v=0\\ 1, \rightarrow u,v\neq 0\end{cases}$$

f(x,y) je diskretna vrednost elementov slike v bloku s koordinatami *xy.* Rezultat DCT-ja je matrika z vrednostmi transformacijskih koeficientov *F(u,v)* s koordinatami *uv*. Ko smo izračunali koeficiente, se izvede kvantizacijska operacija:

$$F^{Q}(u,v)=\left\lfloor\frac{F(u,v)}{Q(u,v)}\right\rfloor$$

kjer je *Q(u,v)* kvantizacijska tabela s 64 elementi. *Kvantizacija koeficientov DCT* je proces zmanjševanja števila možnih vrednosti količine, s čimer se zmanjša število bitov, potrebnih za njihovo predstavitev. Preprost primer kvantizacije je zaokroževanje realnih števil v cela števila. Tako bi realno število 2.765423 zaokrožili na število 3 in bi s tem zmanjšali število bitov, potrebnih za predstavitev, na račun izgube podatkov. Pri stiskanju slik JPG, je vsak koeficient DCT kvantiziran z uporabo uteži, ki je odvisna od frekvence za določen koeficient (sprememba višjih frekvenc manj vpliva na opazno slabšanje kvalitete slike). Koeficienti DCT se delijo s kvantizacijsko matriko, rezultat pa se zaokroži navzdol na najbližje celo število. Koeficienti DCT so nato še obdelani s Huffmanovo metodo kodiranja za zmanjšanje velikosti datoteke.

Z uporabo enakih načel kot pri LSB zamenjavi je sporočilo mogoče skriti v najmanj pomembne bite kvantiziranih koeficientov DCT, pred uporabo Huffmanovega kodiranja za stiskanje. Pomembno je, da se skrivanje izvede pred Huffmanovim kodiranjem, ker se v tej fazi odvečni podatki in šumi izgubijo. Ta postopek je prikazan na sliki 7. Prednost omenjene metode je, da poteka znotraj frekvenčne domene slike in zato spremembe ne bodo vidne na nosilni sliki in jih bo zelo težko odkriti.



Slika 7: Postopek vstavljanja skrivnih bitov med stiskanjem JPG.

Primer algoritma za skrivanje sporočila v sliko JPG:

Vhodni podatki: sporočilo, nosilna slika.

Izhodni podatki: stego-slika, ki vsebuje skrito sporočilo.

While *obstajajo podatki za skrivanje* ***do***

pridobi naslednji DCT koeficient iz nosilne slike

if *DCT ≠ 0 in DCT ≠ 1* ***then***

pridobi naslednji LSB iz sporočila

zamenjaj DCT LSB z bitom sporočila

end if

vstavi DCT v stego-sliko

end while.

3.3.3.2 Diskretna Fouirierjeva transformacija

Diskretna Fourierjeva transformacija, v nadaljevanju DFT, je linearna preslikava, ki pretvori signal ali sliko iz prostorske domene v frekvenčno domeno. Glede na to, da je vhodna funkcija končno zaporedje realnih in kompleksnih števil, je DFT idealna metoda za obdelavo podatkov, shranjenih v računalniku. Pogosto se uporablja pri obdelavi signalov (obdelava slik, obdelava zvoka) in za reševanje parcialnih diferencialnih enačb [19].

Pri DFT je vsak vektor *F(u,v)* predstavljen iz dveh delov: realnega in imaginarnega, kjer realni del pomeni kosinusno amplitudo, imaginarni pa sinusno. Če formulo za DFT uporabimo na bloku MxN pikslov dobimo naslednjo enačbo:

$$X(u,v)=\frac{1}{MN}\left[\sum_{k=0}^{M-1}\sum_{j=0}^{N-1} f(x,y)\,e^{-j2\pi(kx/M+vy/N)}\right]$$

Uporaba metode LSB se izvaja samo na realnem delu transformirane slike po podobnem postopku kot pri diskretni kosinusni transformaciji.

Omenjena valčna transformacija ima podobne koncepte kot Fourierjeva transformacija, za razliko od Fourierjeve transformacije pa hkrati prikazuje signal v času in prostoru [20]. Signal predstavimo s t.i. valčnimi funkcijami.

3.3.4 Izboljšave metode LSB zamenjave

Kot smo že omenili, so posledica metode LSB zamenjave specifični vzorci na histogramu. Vrsta stegoanalize – napad z histogramom, se ukvarja z razporeditvijo najmanj pomembnih bitov. Statistika pravi, da najmanj pomembni biti niso povsem naključni, ampak obstaja povezava med njimi. Naslednje izboljšave LSB metode preprečujejo statistične stegoanalitske napade.

3.3.4.1 Metoda SLSB zamenjave

Metoda SLSB zamenjave spremeni določeni najmanj pomemeben bit (*SLSB – Selected Least Significat Bit*) v bit iz skrivnega sporočila [21]. Ta metoda zamenjuje najmanj pomemben bit samo ene barvne komponente piksla. Poleg tega so spremenjeni tudi drugi biti, da bi dobili najbolj podobno barvo originalni.
Na primer, če imamo barvo #A8A8A8 (binarni zapis – 10101000 10101000 10101000) in če želimo skriti sporočilo 111, bo rezultat: 1010100**1** 1010100**1** 1010100**1** (tabela 1).

	Šestnajstiško	Desetiško	Rdeča	Zelena	Modra
Originalni piksel	A8A8A8	11053224	168	168	168
Spremenjeni piksel	A9A9A9	11119017	169	169	169

Tabela 1: Primerjava vrednosti barvnih komponent pred in po uporabi LSB metode.

V teoriji so se spremenili samo trije najmanj pomembni biti, vendar je razlika od stare do nove barve 65793 barv na barvni lestvici. Metoda SLSB povzroča manj popačenj in se na ta način poveča učinkovitost. Pri skrivanju treh bitov v zeleno barvno komponento pri istemu primeru, dobimo rezultat: 10101000 10101**111** 10101000 (tabela 2).

	Šestnajstiško	Desetiško	Rdeča	Zelena	Modra
Originalni piksel	A8A8A8	11053224	168	168	168
Spremenjeni piksel	A8AFA8	11055016	168	175	168

Tabela 2: Primerjava vrednosti barvnih komponent pred in po uporabi SLSB metode.

V tem primeru je razlika med staro in novo barvo 1792 barv na barvni lestvici. Če bi se odločili za skrivanje v modro barvno komponento piksla, bi bila razlika samo v sedmih barvah. Prednosti SLSB metode v primerjavi z LSB metodo:

- odporna je na vizualne napade,
- odporna je na napade s histogramom, ker je pogostost pojavljanja barv na stego-sliki zelo podobna originalni sliki,
- odporna je na statistične napade, ker sta dve barvni komponenti enaki tistim iz originalne slike.

3.3.4.2 Postopek optimalne prilagoditve pikslov

Postopek optimalne prilagoditve pikslov ali OPAP (*Optimal Pixel Adjustment Procedure*) je metoda, ki poskuša zmanjšati popačenja, ki nastanejo po uporabi LSB zamenjave [22]. Metoda OPAP popravi vrednosti pikslov, tako da izboljša kakovost stego-slike, pri tem pa ne vpliva na skrite podatke.

Primer uporabe metode OPAP

- Naj bo *n* število najmanj pomembnih bitov zamenjano v vsakem pikslu.
- Naj bo *d* decimalna vrednost piksla po zamenjavi.
- *d1* decimalna vrednost zadnjih *n* bitov piksla.
- *d2* decimalna vrednost *n* skritih bitov v pikslu.

1. Najprej se najmanj pomembni biti zamenjajo s skrivnimi podatki.
2. Če je potrebno, popravi bite v pikslu, ki se nahajajo pred skritimi biti
3. ***If*** *(d1~d2) <= (2^n) / 2*

then *popravek se ne zgodi v tem pikslu*

else

if *(d1 < d2)*

d = d – 2^n

if *(d1 > d2)*

d = d + 2^n

4. *d* se pretvori nazaj v binarni zapis in se zapiše nazaj v piksel.

3.4 Metoda generiranja

Metoda generiranja ni tako pogosta metoda za skrivanje, ampak ima eno pomembno značilnost. Ta metoda ne potrebuje prenosnega medija za prenos skrivnega sporočila. Algoritem datoteko generira na podlagi matrike, ki je lahko vzorec ponavljanja besed, zvokov, fraktalov, ki se pojavljajo v datotekah. Pri tej metodi je potrebno paziti, da je algoritem dovolj dober, da generira zvok ali besedilo, ki ni sumljivo. Znan primer generiranja je opisan v Simmonsevem 'Problemu zapornikov', kjer dva zapornika, Alice in Bob komunicirata preko upravnika Eve, ki ne bo posredovala sporočil, ki se ji zdijo sumljiva. Bob nariše navidezno neškodljivo sliko, ki vsebuje posebne barve in vzorce, Alice pa takoj prepozna skrivno sporočilo in ga zna interpretirati.

Še en znan primer metode generiranja je spletna aplikacija Spam Mimic [23], ki generira nezaželeno pošto, ki jo srečujemo vsakodnevno na medmrežju. Razlika je v tem, da aplikacija skrije sporočilo znotraj nezaželene pošte, ki jo pošljemo na čim več naslovov, med naslovniki pa bo seveda tudi tista oseba, ki ji je sporočilo namenjeno.

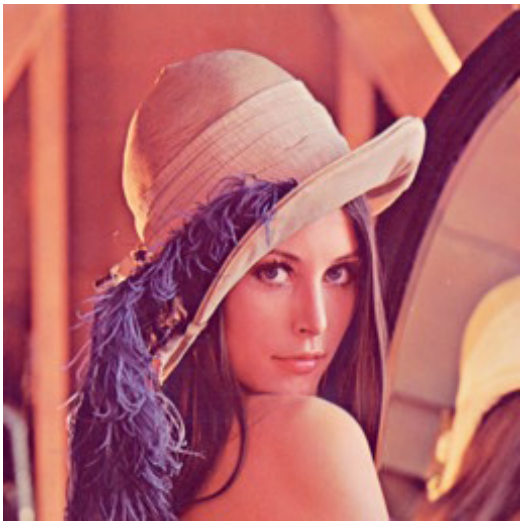
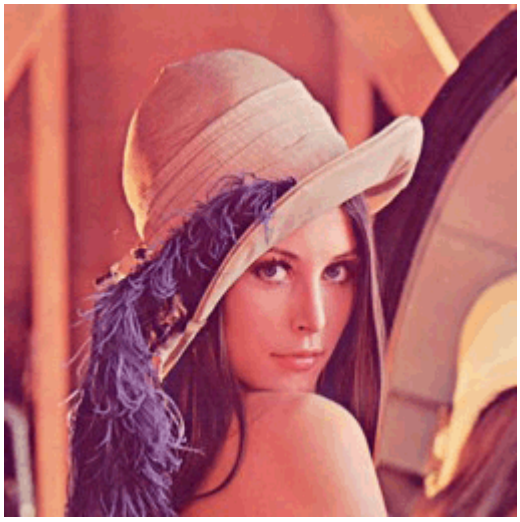
4 Primerjava metode LSB zamenjave v različnih slikovnih formatih

Čeprav je uporaba steganografije razširjena na različne prenosne medije, kot so slikovne datoteke, avdio zapisi, video zapisi in zaglavja TCP/IP paketov, se bomo osredotočili na skrivanje podatkov znotraj slikovnih datotek, ker je ta vrsta steganografije najbolj razširjena. Glede na to, da je steganografska metoda zamenjave najmanj pomembnega bita najbolj pogosta metoda za skrivanje, se bomo osredotočili na njeno delovanje v različnih slikovnih formatih.

Uporabili bomo eno izmed najbolj znanih standardnih slik za testiranje, sliko Lene (slika 8). Slika je velikosti 256 x 256 pikslov in je shranjena v formatih BMP, PNG, GIF in JPG. V vsako izmed slik bomo poskusili skriti besedilno datoteko velikosti 24kB s pomočjo metode zamenjave najmanj pomembnih bitov, razen pri formatu JPG, pri katerem bomo podatke skrivali v najmanj pomembne bite koeficientov DCT (razdelek 3.3.3.1). Z metodama za ocenjevanje kakovosti slike MSE in PSNR, ki sta predstavljeni v razdelku 5.2.2, bomo ocenili kakovost slik po uporabi metode LSB zamenjave. Rezultat je prikazan v tabeli 3.

	Format BMP	Format GIF	Format PNG	Format JPG
MSE	19,93	592,93	6,99	3
PSNR	44 dB	28,8 dB	51 dB	54,30 dB

Tabela 3: Meritve kakovosti slik po LSB zamenjavi.

 LenaOriginal.bmp	 LenaStego.bmp
 LenaOriginal.gif	 LenaStego.gif

Slika 8: Primerjava originalnih in stego-slik.

Metode bomo analizirali na podlagi dobljenih rezultatov in zahtev, ki jih bomo opisali v nadaljevanju. Vsaka izmed omenjenih metod ima tako močne kot šibke točke, najpomembnejša lastnost pa je nezaznavnost. Našteli bomo zahteve, ki definirajo nezaznavnost algoritma:

- Nevidnost: je prva in najpomembnejša zahteva, glede na to, da je moč steganografije v tem, da je nezaznavna človeškemu očesu.

- Kapaciteta: cilj steganografije je skriti čim več podatkov pri skrivnem komuniciranju.
- Odpornost na statistične napade: statistična stegoanaliza poskuša odkriti skrite podatke na podlagi statističnih testov, ker mnogi steganografski algoritmi pri skrivanju naredijo specifične vzorce.
- Odpornost na spreminjanje slike: v primeru aktivnega napada se lahko na sliki naredijo spremembe s ciljem, da se odstranijo skriti podatki.
- Neodvisnost od slikovnega formata: glede na to, da danes obstaja veliko različnih slikovnih formatov, ki se pošiljajo po medmrežju, stalno pošiljanje istega formata med sogovornikoma lahko postane sumljivo. Najbolj zmogljivi steganografski algoritmi imajo sposobnost skrivanja podatkov v katerekoli formate slik.
- Nesumljive datoteke: ta zahteva vključuje vse neobičajne lastnosti slike, ki jih lahko povzroči steganografski algoritem. Nenormalna velikost datoteke npr. bi lahko vzbudila sum.

Spodnja tabela prikazuje primerjavo metode LSB zamenjave za različne slikovne formate, na podlagi zgornjih zahtev in rezultatov MSE in PSNR analize.

	LSB v formatu BMP	**LSB v formatu GIF**	**LSB v formatu PNG**	**LSB v formatu JPG**
Nevidnost	Visoka	Nizka	Visoka	Visoka
Kapaciteta	Visoka	Srednja	Srednja	Nizka
Odpornost na stat. napade	Nizka	Nizka	Nizka	Srednja
Odpornost na spreminjanje slike	Nizka	Nizka	Nizka	Srednja
Neodvisnost od formata	Nizka	Nizka	Srednja	Visoka
Nesumljive datoteke	Nizka	Nizka	Nizka	Visoka

Tabela 4: Primerjava steganografskih metod.

Vrednosti, ki določajo, kako dobro algoritmi izpolnjujejo zahteve, so opredeljene kot visoka, srednja in nizka. Visoka vrednost pomeni, da algoritem v celoti izpolnjuje zahtevo, medtem ko nizka vrednost pomeni, da ima algoritem pri tej zahtevi pomanjkljivost. Kot je razvidno iz tabele 4, noben algoritem ne izpolnjuje vseh zahtev, tako da se odločimo za primeren algoritem glede na pomembnost zahtev.

Zaključki primerjave:

Glavna pomanjkljivost LSB zamenjave v formatu BMP je zagotovo sum, ki izhaja iz zelo velikih slikovnih datotek, ki se prenašajo med sogovornikoma, saj se format BMP ne uporablja pogosto. LSB zamenjava v formatu BMP je najbolj primerna v primerih, kjer je poudarek na količini podatkov, ki jih je potrebno poslati, in ne na tajnosti teh podatkov.

LSB zamenjava v formatu GIF ima podobne prednosti in pomanjkljivosti kot pri formatu BMP. Glavna razlika je, da imajo slike GIF globino 8 in je zato količina podatkov, ki jih lahko skrijemo, manjša kot pri formatu BMP. Kot vidimo v tabeli 3, je pri skrivanju samo 24kB podatkov kakovost slike dvakrat manjša kot pri drugih formatih. Na stego-sliki so bile spremembe v pikslih takoj opazne, kar pomeni, da je format GIF zelo občutljiv na statistične in vizualne napade. LSB zamenjava je lahko učinkovita za skrivanje razumne količine podatkov v sivinsko sliko.

LSB zamenjava v formatu PNG – format PNG (*Portable Network Graphics*) je rastrski slikovni format, uporablja pa stiskanje brez izgube. Nastal je kot izboljšava formata GIF. Ker se uporablja pogosto, ne bo povzročil sumov kot nosilni medij za skrivanje. V format PNG se lahko skrije veliko podatkov. Kakovost slike formata PNG je precej boljša kot pri formatu GIF, saj uporablja 24-bitno barvno lestvico.

Skrivanje v format JPG ima visoko stopnjo nevidnosti, saj skrivanje poteka v frekvenčni domeni. Njegova prednost je, da je najbolj pogost format na medmrežju in je zato najmanj sumljiv, vendar je implementacija algoritma za skrivanje bolj zapletena. Za razliko od ostalih formatov je bilo v sliko formata JPG velikosti 256 x 256 pikslov mogoče shraniti samo 11kb podatkov.

5 Stegoanaliza

Stegoanaliza je za steganografijo to, kar je kriptoanaliza za kriptografijo. Stegoanaliza ne pomeni samo odkrivanje skritih informacij, temveč tudi pridobivanje skrite vsebine, onemogočanje pregleda vsebine – uničevanje, ter spreminjanje vsebine, da bi se prejemniku poslala napačna informacija. Temelji na različnih matematičnih, predvsem pa statističnih analizah. Osnovna razdelitev stegoanalize je na aktivne in pasivne napade.

5.1 Aktivni napadi

Aktivni napadi vključujejo uničenje skritega sporočila in so bolj razširjeni v tehnologijah, kot je digitalni vodni žig, kjer je glavni cilj odstranjevanje žiga. Uporabni so tudi v primerih, kjer sumimo, da skrito sporočilo obstaja, vendar odkrivanje njegove vsebine ni pomembno. Aktivni napad deluje na način, da spremeni stego-medij, vendar spremembe niso opazne, skrito sporočilo pa ni več veljavno.
V kontekstu slikovnih datotek, obstajajo številne metode aktivnih napadov, ki uničijo skrivno sporočilo:

- Zameglitev: gladi prehode in zmanjša kontrast slike.
- Šum: naključni šumi dodajajo na sliko piksle naključnih barv.
- Zmanjšanje: zmanjša šum na sliki s prilagoditvijo barve in povprečevanjem vrednosti pikslov.
- Ostrenje: nasprotni proces zameglitvi. Poveča kontrast med sosednjima piksloma, običajno na robovih objektov.
- Vrtenje: premakne sliko okrog središča v določeni smeri.
- Ponovno vzorčenje: vključuje interpolacijski postopek z namenom zmanjšanja nepravilnosti.

- Mehčanje: zameglitev cele slike zaradi glajenja robov in zmanjšanja kontrasta, povzroča manj izkrivljanja kot zameglitev.

Steganografija se zelo težko upre navedenim metodam, zlasti če je več metod uporabljenih zaporedoma. Znano orodje za tovrstne aktivne napade je aplikacija StirMark [24].

5.2 Pasivni napadi

Pasivni napadi temeljijo na odkrivanju in dešifriranju skrivnih sporočil. Glede na to, koliko informacij je na voljo napadalcu, lahko pri pasivnem napadu uporablja naslednje tehnike:

- Napad na steganografijo: ko imamo za analizo na voljo samo stego-medij.
- Napad na znan prenosni medij: ko sta nam dostopna prenosni in stego-medij, lahko naredimo primerjavo med njima.
- Napad na znano sporočilo: ko sta nam znana skrito sporočilo in stego-medij, želimo pa ugotoviti uporabljeno metodo za skrivanje.
- Napad na izbrano steganografijo: ko sta nam znana medij in orodje ali algoritem za skrivanje.
- Napad na izbrano sporočilo: ko sta nam znana sporočilo in algoritem.
- Napad na znano steganografijo: ko so nam znani nosilec, medij in algoritem.

V kontekstu skrivanja v slikovne datoteke, obstajata dve vrsti pasivnih napadov:

- vizualni napadi,
- statistični napadi.

5.2.1 Vizualni napadi

Več avtorjev domneva, da so najmanj pomembni biti vrednosti svetilnosti popolnoma naključni in so zato lahko zamenjani z biti iz skrivnega sporočila. *Vizualni napadi* temeljijo na tem, da je ta domneva napačna. Pravzaprav je zelo težko razlikovati

naključnost in vsebino slike s pomočjo računalnika - zaradi človeške sposobnosti za prepoznavanje znanih oblik je pri vizualnih napadih potreben človek. Če omogočimo prikaz plasti manj pomembnih bitov stego-slike in jo primerjamo z originalom, bodo razlike takoj opazne.

5.2.2 Statistični napadi

Pri *statističnih napadih* se uporablja statistična analiza slik, ki s pomočjo matematičnih formul zaznava prisotnost skritih podatkov. Statistični napadi izkoriščajo predpostavko, da skrito sporočilo v sliko vnaša naključnost. Obstajajo trije načini za primerjanje kakovosti originalne in stego-slike: povprečna kvadratna napaka, PSNR in histogram.

Povprečna kvadratna napaka

Povprečna kvadratna napaka ali MSE *(Mean Squared Error*) med dvema slikama (originalno in stego) je definirana kot:

$$MSE = \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \left[I(i,j) - K(i,j) \right]^2$$

kjer sta *I(i,j)* in *K(i,j)* sliki velikosti *M*×*N* pikslov. Manjša vrednost MSE pomeni manjšo napako, medtem ko za PSNR velja obratno [25].

PSNR

Pojem *Peak Signal-to-Noise Ratio* ali *PSNR* pomeni razmerje med maksimalno vrednostjo signala in šumom, ki popači izvorni signal. Izraža se v decibelni logaritemski skali [26].

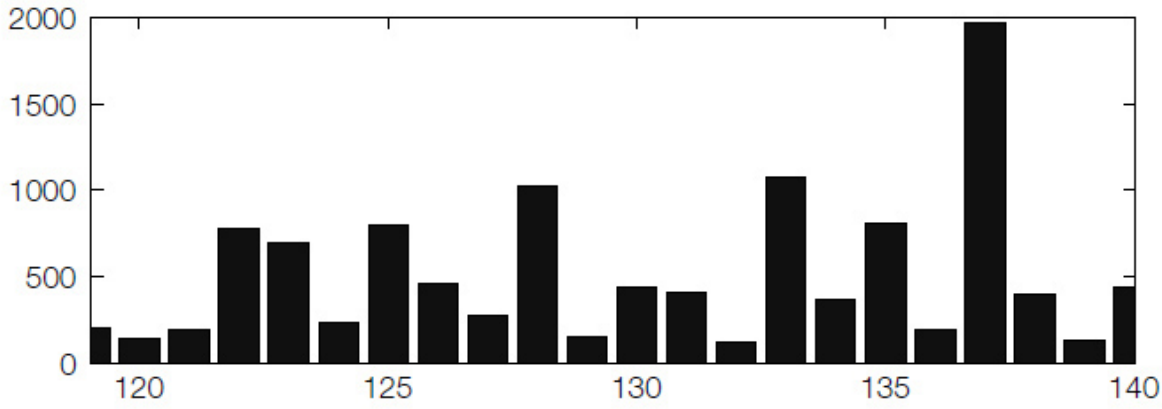
$$PSNR = 20 \log_{10} \frac{255}{\sqrt{MSE}}$$

Večja vrednost PSNR pomeni boljšo kvaliteto. Ko sta dve sliki identični, je MSE nič. Iz tega sledi, da je PSNR nedefiniran. PSNR se pogosto uporablja kot merilo za kakovost rekonstrukcije slik pri algoritmih stiskanja z izgubo.

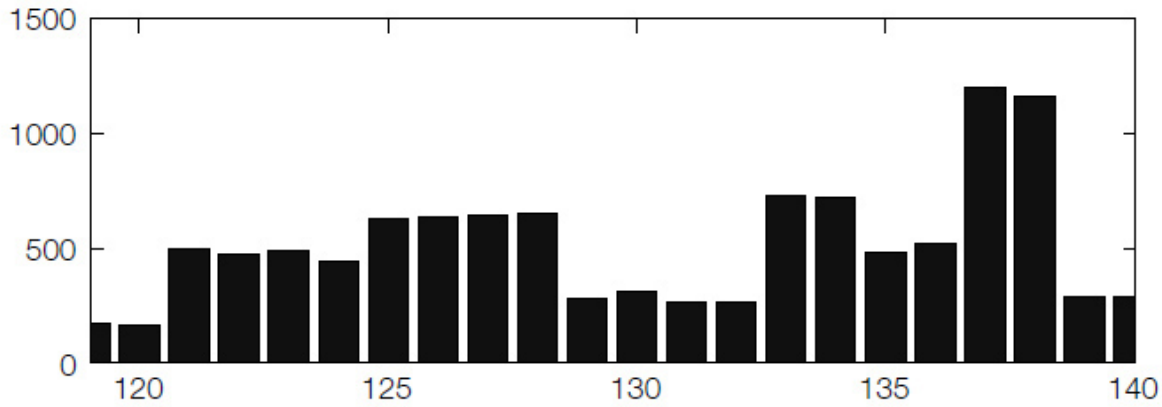
Histogram

V praksi je zelo težko oceniti ali je slika »naravna«, vendar smo lahko prepričani, da ni, če vsebuje nepravilnosti, ki so značilne za določene steganografske metode, kot je metoda zamenjave najmanj pomembnih bitov [27].

Zaradi asimetrije vnosne metode lahko vidimo specifične vzorce v histogramu, ki se sicer nikoli ne pojavijo v nespremenjenih slikah. Na slikah 9a in 9b vidimo del histograma sivinske slike pred in po uporabi LSB zamenjave (1 bit na vsak piksel).



Slika 9a: Histogram sivinske slike pred LSB zamenjavo.



Slika 9b: Histogram sivinske slike po LSB zamenjavi.

Očitno je, da se je vsak stolpec histograma para vrednosti PV, (2i, 2i+1) izenačil pri uporabi LSB zamenjave. To je intuitivno jasno, ker so se najmanj pomembni biti naključno spremenili in je tako enaka verjetnost da bo vsak piksel od PV imel vrednost 2i ali 2i+1.

6 Posebne oblike steganografije

Na koncu je pomembno opisati še dve metodi, ki nista manj pomembni in imata drugačen namen kot steganografija, ampak zaradi svojih lastnosti sodita v isto panogo. V današnjem svetu se pogosto srečujemo z nezakonitimi dejanji na medmrežju. Nelegalne namestitve programov, prenos filmov, glasbe prek omrežja in podobna dejanja so danes bolj pravilo kot izjema. Da bi takšna dejanja preprečili, so nastali načini za zaščito avtorskih pravic. Uporabljata se predvsem digitalni vodni žig in digitalni prstni odtis.

6.1 Digitalni vodni žig

Digitalni vodni žig je povezan s steganografijo, vendar se uporablja v različne namene. Skupna točka vodnemu žigu in steganografiji je skrivanje informacij, vendar cilj steganografije je skriti dejstvo, da skrito sporočilo obstaja, medtem ko je digitalni vodni žig namenjen prenosu skritih podatkov in ni nujno neviden [28]. Obstajata dve vrsti vodnega žiga: krhek in robusten. Krhek vodni žig se uniči že pri majhnih spremembah in na ta način pokaže, da je bil izveden poskus napada. Robusten vodni žig je lahko viden ali neviden in ga je zelo težko odstraniti ali poškodovati. Za razliko od steganografskih vnosnih metod je robustni vodni žig odporen na različne spremembe slike, kot so rezanje, rotacije, vrtenje, stiskanje slike ipd.
Uporaba nevidnega vodnega žiga je najbolj razširjena na naslednjih področjih:

- Zaščita avtorskih pravic: preprečevanje kraje digitalnih multimedijskih datotek, kjer je pomembno, da se žig ne more odstraniti in da je odporen na različne spremembe nosilne datoteke.
- Zaščita pred kopiranjem: nadzoruje napravo za kopiranje podatkov in prepreči kopiranje zaščitenih multimedijskih vsebin.
- Preverjanje avtentičnosti: žig je dodan na celoten signal prenosnega medija, da bi lahko zaznali morebitne spremembe medija.

- Zdravstvene kartoteke: glede na to, da medicinska industrija vse bolj uporablja digitalne zapise, bi lahko vodni žig postal orodje za preprečevanje zamenjave bolniških kartotek. Večina slikovnih formatov še vedno ločuje slikovne podatke in besedilo (ime bolnika, datum slikanja, ime zdravnika) in hitro lahko pride do izgube ali zamenjave podatkov. To lahko preprečimo s vstavljanjem bolnikovih podatkov v sliko.
- Digitalna glasba: s pomočjo vodnega žiga lahko v datoteko zapišemo, da je zaščitena z avtorskimi pravicami in tudi podatke o izvajalcu. Vodni žig bi lahko prilagodili, da izsledi prvotnega lastnika piratske datoteke.

6.2 Digitalni prstni odtis

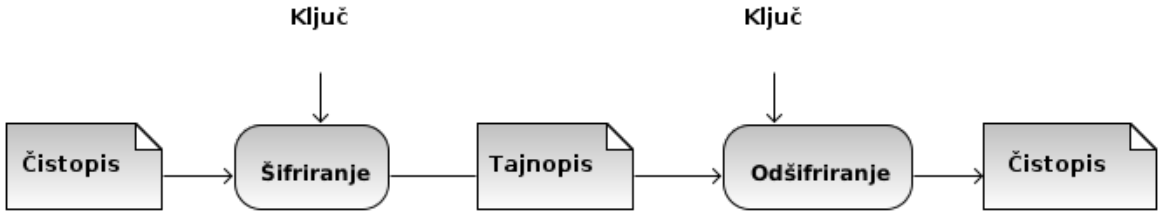
Digitalni prstni odtis je posebna oblika vodnega žiga, ki se nanaša predvsem na informacije o ustvarjalcu ali o prejemniku. Njegov namen je preprečevanje nezakonitega razširjanja kupljenih vsebin na medmrežju (slike, dokumenti, glasba, itd). Za razliko od digitalnega vodnega žiga, pri katerem ima vsaka zakonita kopija izdelka isti vodni žig, ima pri digitalnem prstnem odtisu vsaka kopija lasten prstni odtis, ki je edinstven, skrit in težko uničljiv. Vse nezakonite kopije bodo imele isti prstni odtis kot original in pri iskanju krivca za nedovoljeno širjenje preberemo prstni odtis nezakonite kopije, nato pa v bazi odtisov originalov poiščemo kupca.

7 Razlike in povezave med steganografijo in kriptografijo

Čeprav sta steganografija in kriptografija sorodni tehniki, sta njuna cilja različna. Cilj kriptografije je narediti podatke neberljive, medtem ko je cilj steganografije prikriti obstoj podatkov [28]. S pomočjo kriptografije lahko zaščitimo vsebino med prenosom, ampak ko je enkrat dešifrirana, nima nadaljnje zaščite.

Šifrirana sporočila so očitna, in ko so prestrežena, je jasno, da pošiljatelj in prejemnik skrivno komunicirata. Steganografija je mlajša sestra kriptografije. Namesto da bi sporočila šifrirali, jih skrijemo v druge navidezno neškodljive vsebine in tako njihova prisotnost ni razkrita. Zato je steganografija alternativna možnost v državah, kjer je uporaba šifriranja nezakonita ali v zatiralskih režimih, kjer bi z uporabo kriptografije privabili nezaželeno pozornost.

Steganografija in kriptografija skupaj izboljšujeta učinkovitost varnostnega sistema. Tudi če je zaznan obstoj skritega sporočila, je sporočilo še vedno šifrirano in napadalcu nerazumljivo. Pred prenosom se kriptografija uporabi za šifriranje sporočil s pomočjo ključa. Ko sprejemnik prejme skrito sporočilo, ga mora dešifrirati z uporabo istega ali sorodnega ključa, da bi razkril vsebino skritega sporočila. Ta ureditev je prikazana na sliki 10.



Slika 10: Model komunikacijskega kanala s šifriranjem.

Obstajata dve uporabi kriptografije. Prva je, da se preprečijo pasivni napadi v obliki nepooblaščenega branja sporočila. Druga je, da se preprečijo aktivni napadi v obliki

nepooblaščenega pisanja. Vendar kriptografija ne bo nujno preprečila nasprotniku, da zazna prenos skritega sporočila. Poleg tega kriptografija ne zagotavlja nobene zaščite pred nasprotnikovo namero o motenju ali odstranitvi sporočila pred dostavo le-tega prejemniku.

7.1 Simetrični in asimetrični algoritmi

Simetrične algoritme imenujemo tudi algoritmi z zasebnim ključem. Pri tem za šifriranje in dešifriranje potrebujemo le en ključ. Omogočajo nam hitro šifriranje, težava pa je v izmenjavi ključa, ki jo je težko izvesti na varen način. Težava je tudi v številu ključev, saj mora vsak uporabnik imeti za vsakega dopisovalca svoj ključ. Zaradi zgoraj omenjenih pomanjkljivosti so se razvili *asimetrični algoritmi* oz. *algoritmi z javnim ključem*. Sistem deluje na podlagi dveh povezanih ključev, javnega in zasebnega. Tvorita par ključev na način, da je javni ključ nepovratno pridobljen zasebnega. Podatke, ki smo jih šifrirali z javnim ključem, lahko dešifriramo izključno s pripadajočim zasebnim ključem. Postopek je iz matematičnega vidika precej bolj zapleten in zato počasnejši od simetričnega.

8 Sklepne ugotovitve

Steganografija je zelo zanimiva tema in ni samo predmet raziskovanja v laboratorijih, ampak je prisotna tudi v vsakdanjem življenju.

V tej diplomski nalogi smo predstavili zgodovino steganografije in primere uporabe v današnjem času, za boljše razumevanje samega termina. Glavni namen diplomske naloge je bil podati presek najpomembnejših steganografskih metod in razložiti njihovo delovanje. Poudarek je bil na skrivanju podatkov v slikovne datoteke, saj je ta način skrivanja najbolj razvit. Podani so bili kriteriji oz. zahteve steganografskih metod pri različnih slikovnih formatih, s ciljem, da se uporabnik lažje odloči za ustrezno metodo glede na njegove zahteve. Učinkovitost določene steganografske metode je močno odvisna od slikovnega formata in izbire same slike. Rezultati so pokazali, da za prenosni medij niso primerne slike, ki vsebujejo manjše število barvnih odtenkov ali slike, kjer so barvni kontrasti zelo veliki. Niso vsi slikovni formati enako primerni za skrivanje, izbira je odvisna predvsem od naših potreb. Zelo težko je izpostaviti določeno metodo kot najboljšo, vsaka ima določene prednosti kot tudi pomanjkljivosti. Če lahko metoda skrije veliko količino podatkov, verjetno ne bo zelo robustna in obratno. Na podlagi rezultatov testiranja s pomočjo metod za ocenjevanje kakovosti slike MSE in PSNR vidimo, da smo pri skrivanju s pomočjo metode LSB v format JPG dobili najboljše rezultate, kar je tudi logično, saj se podatki skrivajo znotraj frekvenčne domene. Slabost tega načina pa je manjša kapaciteta skrivanja kot pri formatih BMP, GIF in PNG. Čeprav je količina skritih podatkov manjša kot pri ostalih formatih, je skrivanje v format JPG priporočljivo zaradi najmanjših popačenj slike.

Steganografija je zelo učinkovita tehnika, ki omogoča varno in skrivno komuniciranje. V kombinaciji s kriptografijo predstavlja dodatno varnostno plast za zaščito informacij. Glede na to, da je steganografija relativno mlada veda v razvoju, je težko z gotovostjo napovedati njen nadaljni razvoj. Predvideva se, da bo uporaba steganografije postala še pogostejša, steganografske metode pa kompleksnejše in

bolj prefinjene. Ker že obstajajo primeri računalniških virusov in črvov v animacijah Flash, lahko sklepamo, da se bodo razvile tudi steganografske metode, s pomočjo katerih bodo trojanski konji in virusi skriti v slike in avdio zapise, aktivirali pa se bodo z ogledom oz. predvajanjem teh datotek. Na drugi strani lahko pričakujemo razvoj antivirusnih programom s stegoanalitskimi sposobnostmi za odkrivanje virusov v avdio in slikovnih datotekah.

Uporaba steganografije pa seveda ni omejena samo na nezakonite in sumljive aktivnosti. Obstaja veliko možnosti za uporabo steganografije v zakonitem kontekstu in v prihodnosti se je potrebno osredotočiti na njihov razvoj. Ponudniki vsebin si želijo, da bi zaščitili svoja avtorska dela proti nezakoniti distribuciji. Digitalni vodni žig omogoča sledenje lastnikom. Čeprav ta metoda ne bo preprečila same distribucije, bo ponudnikom vsebin v pomoč pri sodnih postopkih zoper kršiteljem avtorskih pravic. Še eno možno področje razvoja so pametne osebne kartice, kjer so podatki zapisani v sliko. Poleg tega je uporaba možna tudi v podjetjih, za zaščito poslovnih skrivnosti. V medicinskih informacijskih sistemih obstajajo aplikacije, kjer je potrebno ločevanje pacientovih slik in podatkov zaradi zaupnosti, po drugi strani pa mora med njima obstajati povezava. Tu bi lahko steganografija precej pripomogla: če bi podatke o pacientu vstavili v njegove slike, bi to predstavljalo koristen varnostni ukrep, ki bi preprečil zamenjavo zdravstvenih kartotek.

9 Literatura in viri

[1] (2012) Johannes Trithemius. Dostopno na: http://en.wikipedia.org/wiki/Johannes_Trithemius

[2] Petitcolas, F. A. P., Anderson, R. J., Kuhn M. G. Information Hiding – A Survey, University of Cambridge,1999.

[3] Singh, S., Knjiga šifer: Umetnost šifriranja od starega Egipta do kvantne kriptografije, Tržič, 2006.

[4] Kipper, G., Investigator's Guide to Steganography, Auerbach Publications, 2004, pogl. 4.

[5] Kipper, G., Investigator's Guide to Steganography, Auerbach Publications, 2004, pogl. 3.

[6] (2011) Stretching the Limits of Steganography. Dostopno na: http://www.cl.cam.ac.uk/~rja14/Papers/stegan.pdf

[7] Simmons, G. J., The Prisoner's Problem and the Subliminal Channel, Plenum Press, 1984, str 51-67.

[8] Cox, I. J., Miller, M. L., Bloom, J. A., Fridrich, J., Kalker, J. Digital Watermarking and Steganography, Morgan Kaufman, 2008, pogl 11.

[9] (2012) Steganogrpahy Tools - Carrier. Dostopno na: http://en.wikipedia.org/wiki/Steganography_tools

[10] Cox & Co., Digital Watermarking and Steganograpy, Morgan Kaufman, 2008, razdelek 2.5.3.

[11] Cox & Co., Digital Watermarking and Steganograpy, Morgan Kaufman, 2008, razdelek 12.1.1.

[12] (2011) Steganography: Hiding Data Within Data. Dostopno na: http://www.garykessler.net/library/steganography.html

[13] (2012) Image File Formats. Dostopno na: http://en.wikipedia.org/wiki/Image_file_formats

[14] (2012) Least Significant Bit. Dostopno na: http://mathworld.wolfram.com/LeastSignificantBit.html

[15] (2005) An Overview of Image Steganography. Dostopno na: http://martinolivier.com/open/stegoverview.pdf

[16] (2010) Steganography: An overview. Dostopno na: http://www.ijest.info/docs/IJEST10-02-10-100.pdf

[17] (2002) Transformacijsko kodiranje. Dostopno na: http://www.vcl.fer.hr/dtv/jpeg/mj2.htm

[18]] (2001) The Discrete Cosine Transform (DCT). Dostopno na: http://www.cs.cf.ac.uk/Dave/Multimedia/node231.html

[19] Wayner, P., Disappearing Cryptography, Morgan Kaufman, 2009, pogl. 14.5.1.

[20] (2001) The Wavelet tutorial. Dostopno na: http://users.rowan.edu/~polikar/WAVELETS/WTtutorial.html

[21] (2011) SLSB: Improving the Steganographic Algorithm LSB. Dostopno na: http://www.fing.edu.uy/inco/eventos/cibsi09/docs/Papers/CIBSI-Dia3-Sesion9(1).pdf

[22] (2004) Hiding data in images by simple LSB substitution. Dostopno na: citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.158.8300.pdf

[23] (2010) Spam Mimmic. Dostopno na: http://www.spammimic.com/

[24] (1997) StirMark – Image-Watermarking robustness test. Dostopno na: http://www.cl.cam.ac.uk/~mgk25/stirmark.html

[25] (2012) Mean Squared Error. Dostopno na: http://en.wikipedia.org/wiki/Mean_squared_error

[26] (2012) PSNR. Dostopno na: http://sl.wikipedia.org/wiki/PSNR

[27] Kipper, G., Investigator's Guide to Steganography, Auerbach Publications, 2004.

[28] Hölbl, M., Skrivanje podatkov – steganografija, Monitor, 2008, str. 100-101. Dostopno na: http://www.monitor.si/clanek/skrivanje-podatkov-steganografija