SVEUČILIŠTE U ZAGREBU

FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

SEMINAR

Supstitucija LSB u steganografiji

Jakov Dorontić

Voditelj: *Doc. dr. sc. Marin Vuković*

Zagreb, ožujak, 2020

Sadržaj

1. Uvod

Ubrzani razvoj tehnologije doveo je do digitalizacije dokumenata, slika, audio i video zapisa što narušava privatnost i distribuciju sadržaja. Ukoliko netko želi pristupiti podacima koji nisu njemu namijenjeni, nove tehnologije i alati mu omogućavaju pristup i nelegalno kopiranje sadržaja. U svrhu zaštite sadržaja danas se koriste mnoge kriptografske i steganografske metode.

Tajnost informacije se može očuvati njenim kriptiranjem. Kada se informacija kriptira ona tada postaje nerazumljiva svima koji je pokušaju pročitati. Samo osoba kome je informacija bila namijenjena može pravilno interpretirati sadržaj informacije.

Drugi način očuvanja tajnosti informacije je da se na pametan način umetne u objekt koji već nosi nekakvu drugu informaciju. Ovakav pristup očuvanja tajnosti informacije naziva se steganografija. Primjena steganografije najvećim dijelom sačinjava korištenje digitalnog vodenog pečata u svrhu zaštite autorskih prava i vlasništva nad multimedijskim datotekama. Steganografija se također koristi kao supstitut za generiranje jednosmjerne hash funkcije koja nakon obrade varijabilne veličine informacija kao rezultat generira izlazni skup podataka fiksne veličine kojim se potom može utvrditi ukoliko su izvršene ikakve promjene nad izvornim skupom podataka. Naposljetku, daleko najlogičnija primjena steganografije je upravo očuvanje povjerljivosti i tajnosti važnih informacija te njihova zaštita od potencijalne sabotaže, krađe ili neovlaštenog pristupa.

U ovom seminarskom radu opisan je način rada supstitucije LSB, kako se može sakriti tekstualna poruka u digitalni medij i na koje se načine može ostvariti digitalni vodeni pečat. Također, opisani su najpoznatiji napadi na supstituciju LSB s ciljem onemogućavanja detekcije tajne poruke.

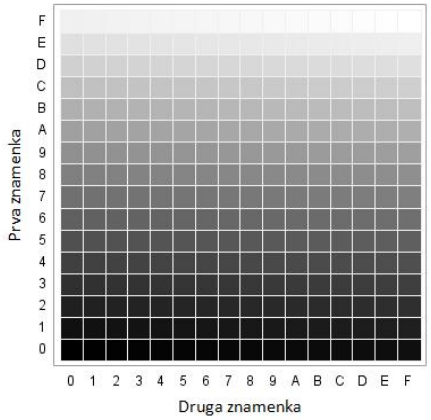
2. Supstitucija LSB

Postoji mnogo različitih načina pomoću kojih se poruka može sakriti unutar digitalnog medija. Naivan pristup bi bio sakriti tajnu poruku u neupotrijebljeni dio datoteke ili nealocirani dio memorijskog prostora.[2] Usprkos velikom broju različitih vrsta medija, slika je ipak najprikladniji digitalni nositelji za kojeg je i najveći broj steganografskih tehnika razvijen. Uobičajeni pristupi su: supstitucija LSB (eng. *least significant bit insertion*), sortiranje paleta, transformacija domena (pomoću diskretne kosinusove transformacije - *DCT* ili pomoću Fourierove transformacije - *DFT*), maskiranje i filtriranje, te mnoge druge.[3]

2.1 Slike u računalu

Slika se u računalu pohranjuje kao niz brojeva koji predstavljaju intenzitet svjetlosti u pojedinoj točki (pikselu) obično u 24-bitnom ili 8-bitnom formatu. Iako je za steganografiju poželjnije koristiti 24-bitni format koji nudi veći prostor za skrivanje informacije, nedostatak mu je veliko opterećenje memorije. Stoga je 8-bitni format češće korišten.[1] Svaka varijacija boje za pojedini piksel izvodi se iz tri osnovne boje: crvene, zelene i plave (RGB). Svaka se osnovna boja u računalu predstavlja s 8 bita. Stoga će primjerice bijela boja imati heksadekatsku vrijednost FF FF FF. Za ovakvo prikazivanje boja u računalu imamo na raspolaganju 2^{24} = 16 777 216 različitih boja.

Mnogi stručnjaci koji se bave steganografijom preporučavaju korištenje slika s 256 nijansi sive boje. Takve slike su vrlo poželjne jer se nijanse postepeno mijenjaju i vrlo je teško uočiti njihovo odstupanje (*Slika 2.1*). Općenito vrijedi da paleta boja mora imati što manje promjena u nijansama kako bi primjena steganografije na njima bila neuočljiva.



Slika 2.1 Paleta od 256 nijansi sive boje

2.2 Implementacija supstitucije LSB

Supstitucija bita najmanje važnosti (eng. *least significant bit substitution*) ili kraće supstitucija LSB je najčešća steganografska tehnika korištena u radu s multimedijskim datotekama. Pojam bita najmanje važnosti vezan je uz numeričku važnost bitova u oktetu. Bit koji najmanje utječe na promjenu broja nalazi se na zadnjoj poziciji okteta i zato ga se naziva najmanje važnim. Ideja steganografske tehnike supstitucije LSB bazira se na rastavljanju tajne poruke na bitove gdje se bitovi tajne poruke pohranjuju na mjesto bita najmanje važnosti u odabranim oktetima slike. Kao jednostavan primjer supstitucije LSB prikazano je skrivanje slova 'G' unutar niza okteta koji predstavljaju nijanse sive boje.[3]

10010101	00001101	11001001	10010110
00001111	11001011	10011111	00010000

1001010**0**	0000110**1**	1100100**0**	1001011**0**
0000111**0**	1100101**1**	1001111**1**	0001000**1**

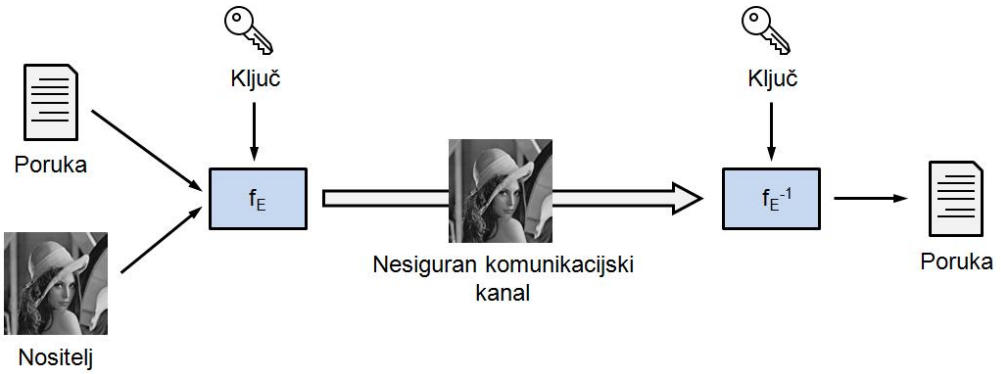
Slovo 'G' se prema ASCII (eng. *American Standard Code for Information Interchange*) standardu zapisuje kao binarni niz 01000111. Ovih 8 bitova zapisuje se na mjesto bitova najmanje važnosti u izvornom skupu okteta. Opisani princip vrlo je djelotvoran zbog činjenice da čovjekov optički sustav nije dovoljno osjetljiv da primijeti takvu malu promjenu u nijansi boje.

Ukoliko se umjesto samo zadnjeg bita najmanje važnosti promijene zadnja dva, tri ili četiri bita najmanje važnosti s ciljem skrivanja veće količine informacija, dolazi do gubljenja originalne informacije unutar multimedijske datoteke. *Slika 2.2* prikazuje kako se originalna multimedijska datoteka mijenja kada joj se zadnji bitovi najmanje važnosti zamijene s 0. Za ovakvu paletu boja (256 nijansi sive) mogu se iskoristiti zadnja tri bita najmanje važnosti u svrhu skrivanja informacije, s time da promjena originalne multimedijska datoteka neće biti opažena.[2]

Slika 2.2 Promjena kvalitete slike [2]

2.3 Skrivanje tajne poruke

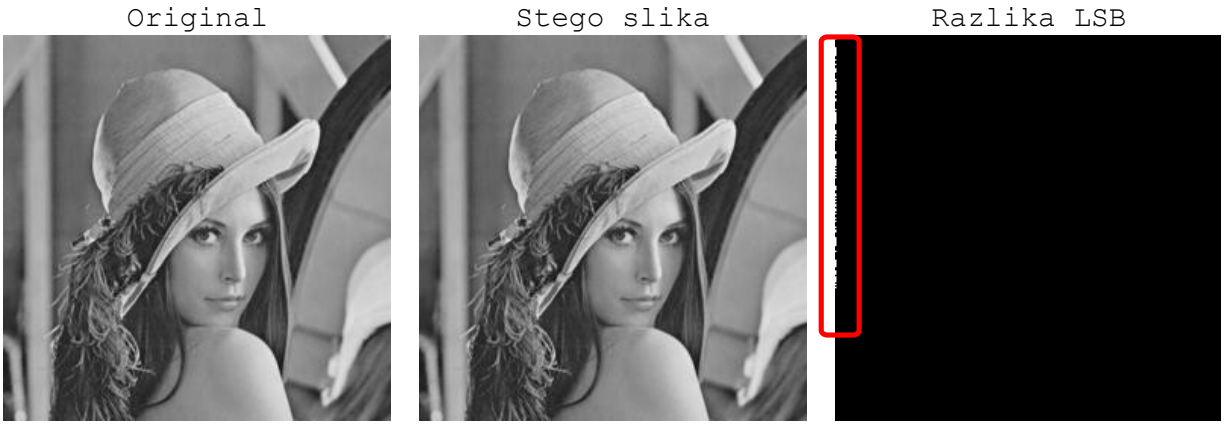
Proces steganografije obično uključuje umetanje tajne poruke unutar nekog prijenosnog medija koji ima ulogu prikrivanja postojanja tajne poruke. Nositelj mora biti takav skup podataka koji je sastavni dio uobičajene svakodnevne komunikacije te kao takav ne privlači posebnu pozornost na sebe (tekst, slika, audio ili video zapis).[3] U svrhu dodatne zaštite, moguća je i uporaba steganografskog ključa kojim se tajna poruka kriptira prije umetanja. Steganografski medij se stoga može prikazati u sljedećem obliku:



Slika 2.3 Slanje tajne poruke [3]

Poruka se u digitalni medij može sakriti sekvencijskim načinom (*eng. sequential path*) ili raspršiti slučajnim odabirom (*eng. random path*). Skrivanje tajne poruke "`Ja boravim u Zagrebu!`" prikazano je na *Slika 2.4* i *Slika 2.5*. Poruka nije kriptirana i sadrži 21 znak. Digitalan medij je slika dimenzija 256 × 256 piksela kojom se prenosi poruka. Kako bi se poruka mogla moći sakriti, na slici je potrebno izmijeniti 21 × 8 = 168 bitova originalnog sadržaja.

Sekvencijski način skrivanja poruke skriva poruku tako da se bitovi najmanje važnosti digitalnog medija zamijene s bitovima tajne poruke odozgo prema dolje. Takvo skrivanje poruke (bez enkripcije) vrlo je lako otkriti.



Slika 2.4 Skrivanje tajne poruke sekvencijskim načinom [2] [8]

Drugi način skrivanja poruke je raspršiti ju slučajnim odabirom po cijelom digitalnom mediju. Slučajan odabir se zapravo računa prema simetričnom kriptografskom algoritmu toka RC4 uporabom steganografskog ključa kojega moraju i pošiljatelj i primatelj poznavati kako bi se moglo odrediti koji pikseli nose informaciju. RC4 je vrlo jednostavan algoritam koji ulazne podatke enkriptira bit po bit tako da provede XOR operacije između toka izvornog teksta i toka pseudo-slučajnih brojeva generiranih u obliku S-box tablice na temelju predanog steganografskog ključa. Za primjer na *Slika 2.5* steganografski ključ je riječ "`password`".

Slika 2.5 Skrivanje tajne poruke slučajnim odabirom [2] [8]

2.4 Digitalni vodeni pečat

Digitalno označavanje (*eng. Digital Watermarking*) temelji se na umetanju podatka, vodenog pečata (*eng. Watermark*), u izvorni dokument sa svrhom njegove ponovne detekcije. Dokument koji se označava može biti bilo koja vrsta informacije: slika, video, zvuk ili tekst. Pečat može sadržavati bilo koju informaciju, kao na primjer logo firme, identifikaciju kupca, prodavača ili nešto drugo.[3] Najčešće primjene digitalnog vodenog pečata je očuvanje autorskih prava, identifikacija vlasnika, financijske transakcije i kontrola kopiranja zaštićenog sadržaja.

Vodeni pečat se najčešće prije umetanja u digitalni medij pretvori u čistu crno-bijelu sliku (bez nijansi sivih boja). Na taj način se dobije niz nula i jedinica koji se poput teksta lako umetnu u digitalni medij. Takvom pretvorbom se čuva originalna dimenzija pečata. *Slika 2.6* prikazuje umetanje crno-bijelog pečata istih dimenzija kao i polazni digitalni medij (256 × 256 piksela)

Slika 2.6 Crno-bijeli digitalni vodeni pečat [8]

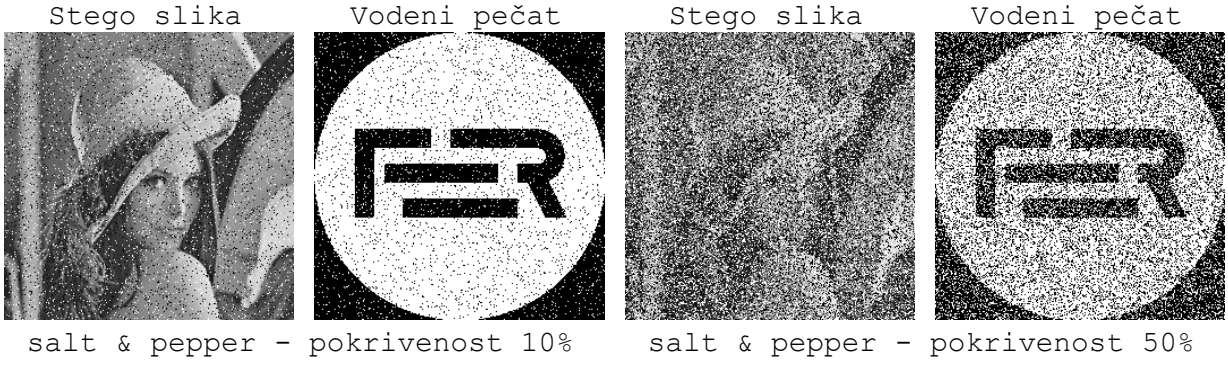
Ukoliko se želi koristiti digitalni vodeni pečat sa svim nijansama sive boje, tada se on mora umanjiti za otprilike 35% veličine digitalnog medija. *Slika 2.7* prikazuje umetanje umanjenog pečata s nijansama sive boje čije su dimenzije 90 × 90 piksela. Digitalni medij s dimenzijom 256 × 256 piksela može iskoristiti skoro sve svoje piksele za pohranu digitalnog vodenog pečata.

Slika 2.7 Umanjeni digitalni vodeni pečat [8]

2.5 Napadi na supstituciju LSB

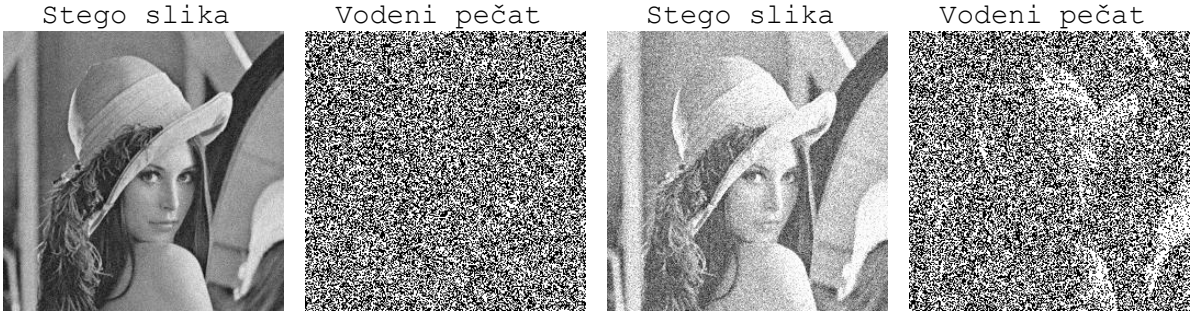
Cilj napada na supstituciju LSB je nemogućnost ponovne detekcije skrivenog teksta ili vodenog žiga iz zadane stego slike. Sukladno intenzitetu distorzije na stego slici, detekcija teksta ili vodenog žiga provodi se više ili manje uspješno. Najpoznatiji napadi na supstituciju LSB su umetanje aditivnog šuma, rotacija i JPEG kompresija stego slike.

Slika 2.8 prikazuje napad salt & pepper šumom koji nasumično postavlja samo neke piksele stego slike u crni ili bijeli piksel.



Slika 2.8 Napad salt & pepper šumom [8]

Slika 2.9 prikazuje napad Gaussovim šumom koji svaki piksel slike promijeni za određenu nijansu sive čiji intenzitet ovisi o zadanim parametrima μ i σ.



Slika 2.9 Napad Gaussovim šumom [8]

Slika 2.10 prikazuje napad rotacijom stego slike oko svojeg središta za zadani kut u stupnjevima.

Slika 2.10 Napad rotacijom [8]

Slika 2.11 prikazuje je napad JPEG kompresijom u kojem se treba specificirati željena kvaliteta kompresije Q.

Slika 2.11 Napad JPEG kompresijom [7] [8]

3. Zaključak

Steganografija predstavlja dobru alternativu kriptografiji jer raspolaže vrlo efikasnim i snažnim tehnikama koje ljudima omogućavaju zaštićenu i skrivenu komunikaciju. Kombinirana s kriptografijom, predstavlja dodatni sigurnosni sloj u zaštiti informacija. Steganografska tehnologija vrlo je jednostavna za upotrebu, a izrazito se teško detektira.

Kao i svaka steganografska tehnika, supstitucija LSB ima svoje prednosti i nedostatke. Supstitucija LSB je najčešće korištena metoda zbog svoje brzine i jednostavnosti, te zato jer pruža prijenos velikog kapaciteta skrivenog sadržaja. Iako je supstitucija LSB vrlo efikasna, neki napadi kao što su JPEG kompresija i napad Gaussovim šumom mogu u potpunosti uništiti skrivenu informaciju, dok napadi rotacijom i napad salt & pepper šumom ne predstavljaju problem.

S obzirom na veliku količinu multimedijskog sadržaja na mrežama, nemoguće je analizirati i detektirati svaki steganografski sadržaj što omogućava zlouporabu steganografskih tehnika, naročito u terorističkim aktivnostima i dijeljenju ilegalnih materijala. Radi toga je razvijena znanstvena disciplina - stegoanaliza, koja se bavi detekcijom skrivenog sadržaja unutar medija. Zbog navedenog problema, napravljeni su alati StegSpy [5] i DIIT [6] (*eng. Digital Invisible Ink Toolkit*) s ciljem detekcije steganografske tehnike i skrivene tajne poruke.

4. Sažetak

Supstitucija LSB u steganografiji

U ovome radu je nakon pojašnjenja pohrane slike u računalu opisan rad supstitucije LSB, načini pohrane tajnog teksta ili vodenog žiga u digitalan medij i prikazani su napadi na supstituciju LSB. Objašnjena je važnost steganografije koja osim za tajnu komunikaciju služi i za zaštitu od ilegalnog korištenja autorskih djela te njihovo umnožavanje. Korišteno programsko okruženje je Matlab R2018a.

Ključne riječi: steganografija, supstitucija LSB, vodeni pečat, zaštita, Matlab

LSB substitution in steganography

This paper describes the operation of LSB substitution after clarifying image storage in computer memory, it also describes ways to store secret text or watermark in digital media and few attacks on LSB substitution are shown as well. The importance of steganography, which, in addition to secret communication, also serves to protect against the illegal use of copyrighted works, is explained. The software used is Matlab R2018a.

Key words: steganography, LSB substitution, watermark, security, Matlab

5. Literatura

[1] Fangjun Huang; Bin Li; Yun Qing Shi; Jiwu Huang; Guorong Xuan: " Image Steganalysis ", Springer-Verlag, Berlin, Heidelberg, 2010

[2] Dr. Kaitai Liang: " Multimedia Security and Digital Forensics ", University of Surrey, 2019

[3] CARNet, s Interneta, https://www.cert.hr/wp-content/uploads/2006/06/CCERT-PUBDOC-2006-04-154.pdf, 29. ožujka 2020

[4] Sanja Zeljković: Steganografija, s Interneta, http://e.math.hr/old/stegano/index.html, 2. travnja 2020

[5] Bill Englehardt: StegSpy, s Interneta, http://www.spy-hunter.com/stegspydownload.htm, 16. travnja 2020

[6] K. Hempstalk, University of Waikato: Digital Invisible Ink Toolkit, s Interneta, http://diit.sourceforge.net/, 16. travnja 2020

[7] Birendra Bikram Singh: Jpeg Compression, s Interneta, https://uk.mathworks.com/matlabcentral/fileexchange/38518-jpeg-compression, 16. travnja 2020

[8] Tao Chen: 8-bit 256 x 256 Grayscale Lena Image, s Interneta, https://www.researchgate.net/figure/8-bit-256-x-256-Grayscale-Lena-Image_fig1_3935609, 16. travnja 2020