UNIVERZITET U NIŠU
ELEKTRONSKI FAKULTET
Katedra za računarstvo

Audio steganografija: Diskretno sekvencijalno širenje spektra

-Digitalna forenzika-

Mentor: prof. dr Bratislav Predić

Student: Milica Todorović, 1256

Niš, 2021.

Sadržaj

1 Uvod

Današnje digitalno doba omogućava jednostavan pristup velikoj količini podataka u različitim formatima, kao što su audio i video snimci, slike, tekst, što čini podatke podložnim različitim napadima. Podaci se mogu ilegalno kopirati, kršenjem autorskih prava, menjati ili im se može pristupati bez odobrenja vlasnika. Iz ovih razloga, pojavljuje se potreba za skivanjem tajnih informacija u okviru različitih tipova digitalnih medijuma. Glavni zadatak steganografije je umetanje i skrivanje tajnih podataka u okviru različitih tipova medijuma. Steganografija omogućava sigurnu i tajnu komunikaciju tako da se prisustvo dodatnih informacija ne može detektovati od strane neautorizovanog korisnika.

Audio steganografija, odnosno korišćenje audio snimaka kao prenosnog medijuma, predstavlja interesantan podskup steganografskih tehnika, s obzirom na visok nivo redundantnosti i visoku brzinu prenosa zvučnih signala. Ipak, skrivanje informacija u okviru audio signala komunikacije u realnom vremenu ne predstavlja jednostavan zadatak.

U ovom radu biće predstavljena primena tehnike direktnog sekvencijalnog širenja spektra u audio steganografiji, koja predstavlja jednu od tehnika širenja spektra. Tehnika širenja spektra je inicijalno razvijena u svrhe sigurne komunikacije za potrebe vojske i često se koristi za slanje tajnih informacija kroz radio talase. Ovom tehnikom se poruka prenosi kao talas nalik šumu. Metoda širenja spektra se izuzetno razvila na ovom polju zbog svoje robustnosti i otpornosti na napade šumom. Navedene prednosti tehnike direktnog sekvencijalnog širenja spektra predstavljaju motivaciju za njen odabir kao centralne teme ovog rada.

U drugom poglavlju ovog rada biće dat pregled digitalne reprezentacije zvuka kao i kratak pregled formata za čuvanje digitalnih audio fajlova, sa akcentom na WAV formatu. Treće poglavlje baviće se opisom steganografskog sistema i tipovima steganografije zavisno od korišćenog digitalnog medijuma. Tema četvrtog poglavlja biće audio steganografija, odnosno tehnike koje se koriste u ovom vidu steganografije, kao i metrike koje se mogu koristiti za evaluaciju datih metoda. Fokus petog poglavlja je na tehnikama širenja spektra, pri čemu je posebna pažnja posvećena primeni metode direktnog sekvencijalnog širenja spektra u audio steganografiji. Finalno, u šestom poglavlju biće predstavljani implementacioni detalji prethodno opisane tehnike, kao i rezultati dobijeni njenom primenom.

2 Digitalna reprezentacija zvuka

Zvuk se može definisati kao: (a) Oscilacija u pritisku, pomeranje čestica, brzina čestica itd., odnosno propagacija u medijumu sa unutrašnjim silama ili superpozicija takve propagirane oscijacije. (b) Auditorna senzacija proizvedena oscijalicijama opisanim pod (a). Fizički, zvuk predstavlja vibraciju koja se propagira kao akustični talas kroz transmisioni medium kao što je gas, tečnost ili čvrst materijal, kao longitudalni talas i takođe transverzalni talas kroz čvrste materijale. Radi jednostavnosti, zvučni talas, se može posmatrati kao sinusni talas opisan svojom amplitudom i frekvencijom. [1]

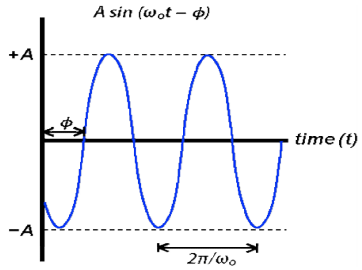
2.1 Sinusni talasi

Sinusni talas predstavlja matematičku krivu glatke periodične oscilacije. Sinusni talas prestavlja kontinualni talas. Grafik ovog talasa (funkcije) može se predstaviti sinusnom funkcijom (Slika 1). U osnovnom obliku, dati talas može se predstaviti jednačinom u funkciji od vremena:

$$y(t) = A\sin(2\pi ft + \varphi) = A\sin(\omega t + \varphi) \quad ,$$

pri čemu je **A** – amplituda (najviša devijacija funkcije od nule), **f** – frekvencija (broj oscilacija u svakoj sekundi vremena), **ω** – ugaona frekvencija (brzina promene argumenta funkcije u radijanima/sekund) i **φ** – faza (specificira gde u ciklusu je oscilacija u trenutku t=0).

Značaj sinusnog talasa proizilazi iz činjenice da talas zadržava svoj oblik nakon dodavanja još jednog sinusnog talasa iste frekvencije a arbitrarne faze i amplitude. Ovakva osobina čini ovu vrstu talasa, akustično unikatnim.



Slika 1: Sinusni talas

2.2 Digitalni audio signali

Digitalni audio signal predstavlja reprezentaciju zvuka snimljenu ili konvertovanu u digitalni format.

Analogni zvučni signal se konvertuje u digitalni signal korišćenjem analognog-u-digitalni konvertera (A/D). Proces konverzije u datoj tački u vremenu, kada je amplituda talasa „uhvaćena“ predstavlja proces semplovanja. Proces semplovanja ponavlja se onoliko puta koliko je potrebno da bi se dobila frekvencija semplovanja potrebna da bi se dati zvuk čuo. Na osnovu Nikvistove (*eng. Nyquist*) teoreme, frekvencija semplovanja mora biti barem dva puta veća od frekvencije zvuka koji je potrebno čuti. Na primer, neka je najviša frekvencija standardnog audio snimka 22050Hz, tada minimalna frekvencija semplovanja mora biti 44100Hz, što će značiti da postoji 44100 semplova po sekundi.[2]

A/D konvertuje dati analogni signal u digitalni, koristeći neku metodu kodiranja, pri čemu se standardno koristi impulsna kodna modulacija (*eng. Pulse-Code Modulation - PCM*). Impulsna kodna modulacija je metoda koja se koristi za digitalnu reprezentaciju analognih signala. U PCM strimu, amplituda analognog signala se sempluje regularno na uniformnim intervalima, a zatim se svaki sempl kvantizuje na najbližu vrednost u okviru datog raspona digitalnih podeoka. Osnovne karakteristike PCM strima su: učestalost semplovanja (*eng. sampling rate*) i bitska dubina/rezolucija (*eng. bit depth/resolution*), odnosno broj mogućih digitalnih vrednosti kojima se može predstavati svaki sempl.

2.3 Formati digitalnih audio zapisa

Zvučni zapisi se, po broju kanala, mogu podeliti na jednokanalne (mono) i višekanalne (stereo) zapise. Stereo zapisi uglavnom imaju dva kanala (levi i desni), što bi značilo da zapis sadrži dva zvučna talasa, levi i desni.

Formati digitalnih audio zapisa mogu se grubo podeliti na:

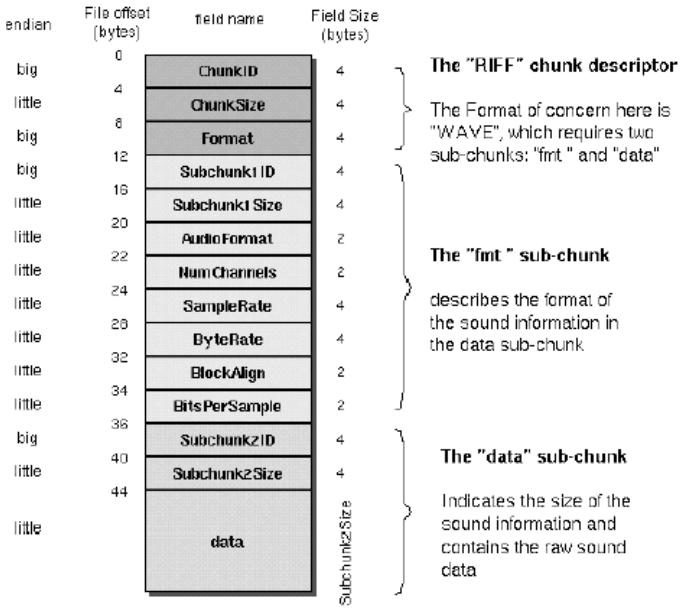
- Nekompresovane audio formate: WAV, AU, itd.
- Formate sa kompresijom bez gubitka (*eng. lossless compression*): FLAC
- Formate sa kompresijom sa gubitkom (*eng. lossy compression*): MP3

Ovaj rad će razmatrati nekompresovane audio formate, konkretno WAV format, te će njegov opis biti dat u nastavku.

2.3.1 WAV/WAVE audio format

WAV/WAVE (*eng. Waveform Audio File Format*) je audio fajl format standard razvijen od strane IBM-a i Majkrosoft-a, za čuvanje audio bitstrimova na računaru. Ovaj format čuva audio snimak u nekompresovanom, sirovom (*eng. raw*), stanju. Česta verzija ovog formata je Microsoft WAV 16-obitski PCM format. Broj bitova označava bitsku dubinu pri PCM kodiranju, odnosno, u ovom konkretnom slučaju, podrazumeva da se svaka amplituda enkodira 16-obitnim označenim integerom. Može se zaključiti da je maksimalna vrednost amplitude, podržana ovim formatom, 32768, dok je minimalna -32768.

WAV fajl format predstavlja podgrupu Majkrosoft-ove RIFF specifikacije za čuvanje multimedijalnih fajlova. Struktura RIFF fajl-a sastoji se iz zaglavlja (*eng. header*) i sekvence delova podataka (*eng. data chunks*). WAV fajl je zapravo RIFF fajl sa „WAV“ odeljkom. Ovaj odeljak podeljen je na dva pododeljka: „fmt“ odeljak, koji specifira format podataka i neke od metadata informacija (broj kanala, učestalost semplovanja itd.) i „data“ odeljka, koji specificira sempl podataka. [2]



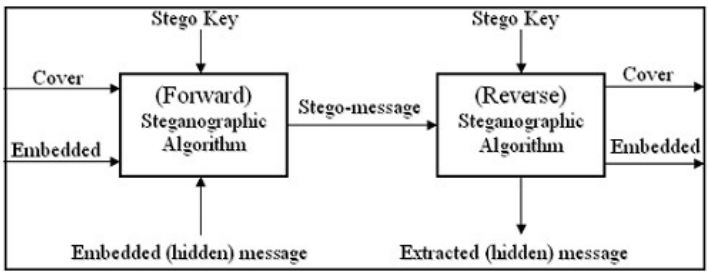
Slika 2: Izgled WAVE fajl formata

3 Steganografija

Steganografija je nauka ili umetnost sakrivanja poruka u okviru drugih izvora informacija kao što su tekst/dokumenti, audio snimci, video zapisi, slike itd, tako da ona nije vidljiva od strane neautorizovanih korisnika. Steganografija, zajedno sa kriptografijom, predstavlja sigurnosnu metodu koja omogućava poverljivost podataka. Ipak, između steganografije i kriptografije postoji jasno definisana razlika. Esencijalno, kriptografija šifrira tajnu poruku tako da se ona ne može pročitati od strane nepoželjnih korisnika, dok steganografija sakriva tajnu poruku u okviru kompjuterskog fajla tako da se ne može videti od strane nepozeljnih korisnika. Gledano sa strane, enkriptovana poruka automatski implicira tajnu konverzaciju. Sa druge strane, sakrivena poruka ne privlači pažnju, te se ne bi moglo posumnjati na postojanje tajne konverzacije. Iz tog razloga, steganografija se često smatra nevidljivim vidom transmisije osetljivih podataka kroz javne kanale na način da niko, sem originalnog pošiljaoca i zadatog prijemnika, ne zna za postojanje date tajne konverzacije.

3.1 Steganografski sistem

Steganografski sistem sastoji se iz tri komponente: prenosni objekat (Slika 3: Cover) – objekat koji sakriva poruku, tajna poruka (Slika 3: Embedded message) i stego-objekat (Slika 3: Stego-message) – prenosni objekat sa umetnutom sakrivenom porukom. Pre slanja, poruka se enkriptuje korišćenjem određenog stego ključa (Slika 3: Stego Key), tajna poruka se dekriptuje sa strani prijemnika, uz pomoć istog ključa. U vreme transmisije poruka je zaštićena, nakon dekripcije poruka postaje nezaštićena i može se kopirati i distribuirati. Tajna ili enkriptivovana poruka može biti tekstualni fajt, šifrovan tekst, slika itd.



Slika 3: Grafički prikaz opšteg steganografskog sistema

3.2 Steganografija u digitalnim medijumima

Mogu se izdvojiti različite tehnike steganografije zavisno od tipa prenosnog objekta koje se moraju ispoštovati kako bi se ostvarila zaštita tajnih podataka.

3.2.1 Steganografija teksta

Tehnike steganografije teksta uključuju broj tab-ova, razmaka, velikih slova, slično korišćenju Morezovog koda za skrivanje podataka.

3.2.2 Steganografija slika

Ovaj tip steganografije podrazuma skrivanje poruka u okviru slika. Data tehnika koristi intenzitete piksela za skrivanje informacija, pri čemu se često koriste slike bitskih dubina 8 i 24. Veća količina informacija može se sakriti u okviru slike veće veličine. Međutim, velike slike mogu da zahtevaju neki vid kompresije kako bi se izbegla detekcija postojanja skrivene poruke.

3.2.3 Steganografija mrežnih protokola

Korišćenje mrežnih protokola (TCP, UDP, IP itd.) kao prenosnog objekta, pri čemu se za prenos koriste dati protokoli spada u oblast steganografije mrežnih protokola. U OSI modelu moguće je izvršiti steganografiju korišćenjem bitova hedera TCP/IP polja koja nisu iskorišćena.

3.2.4 Steganografija audio snimaka (Audio steganografija)

Audio steganografija podrazumeva korišćenje audio snimaka kao objekata za skrivanje podataka. Audio snimci predstavljaju veoma važan medijum zbog sve veće popularnosti internet telefonije, prenosa glasa internetom (*eng. Voice over Internet Protocol*). Koriste se različiti formati audio snimaka, među kojima su najpopularniji: WAVE, MIDI i AU.

3.2.5 Steganografija video snimaka (Video steganografija)

Ova tehnika omogućava skrivanje informacija u okviru digitalnog video formata. Video snimak, odnosno kombinacija većeg broja slika i zvuka, koristi se kao prenosni objekat za date tajne podatke. Može se koristiti diskretna kosinusna transformacija za umetanje podataka u okviru svake slike datog videa, tako da date izmene nisu vidljive od strane ljudskog oka. Različiti video formati se mogu koristiti u ovoj tehnici, pri čemu su najpopularniji: MP4, MPEG, AVI, H.264 itd.

4 Audio Steganografija

Audio steganografija predstavlja vid steganografije koji omogućava skrivanje poruka u okviru digitalnih audio zapisa. Audio steganografija podrazumeva skrivanje poruka u okviru WAV, AU, i u nekim slučajevima MP3 fajlova. Kao i kod drugih tipa medija, moguće je modifikovati zvučne snimke tako da oni sadrže skrivene poruke, pri čemu se ne vrši narušavanje originalnog signala.

U audio steganografiji, poruke se skrivaju eksploatisanjem nedostataka ljudskog auditornog sistema (*eng. Human Auditory System – HAS*). Izazov skrivanja podataka u okviru zvučnih zapisa je upravo ljudski auditorni sistem čiji je opseg dozvoljenih frekvencija veoma veliki (20Hz - 20.000 Hz), što ga čini izutetno osetljivim na dodatne nasumične šumove. Ugrađivanje poruka u ovaj tip medijuma, ispostavlja se kao komplikovaniji zadatak, u odnosu napr. slike, zahvaljujući dinamičkom suverenitu ljudskog auditornog sistema u odnosu na ljudski vizuelni sistem. Količina podataka koju je moguće ugraditi u video frejmove je veća od količine koja se može sakriti transparentno u okviru audio sekvenci jer audio signal ima manje dimenzija od videa. Ljudski auditorni sistem, ipak, ima neke nedostatke („rupe“). Dva atributa ljudskog auditornog sistema se dominantno koriste u steganografskim algoritmima su temporalno maskiranje i maskiranje frekvencije. Koncept perceptivnih rupa ljudskog auditornog sistema se koristi i u određenim tipovima kompresija (MP3). Rupe se koriste kako bi smanjile broj bitova potrebnih za enkodiranje audio signala bez ometanja percepcije datog kodiranog audio signala, u algoritmima kompresije. U okviru algoritama za skrivanje podataka, karakteristike maskiranja se koriste za ugrađivanje dodatnih bitova u postojeće bit strimove, opet, bez dodavanja šuma datom audio signalu, koji bi se mogao opaziti.[3]

4.1 Tipovi audio steganografije

4.1.1 Umetanje

Tehnika umetanja postavlja podatke koje je potrebno sakriti u beznačajan deo prenosnog fajla, koji je uglavnom ignorisan od strane operativnih sistema i softverskih aplikacija. Na primer, većina fajlova sadrže EOF (*eng. end-of-file*) marker koji označava da nema više podataka koji se mogu pročitati iz datog fajl izvora. Slično, izvršni fajlovi se

uglavnom završavaju EOF markerom na kraju binarnih instrukcija. Tehnike steganografije metodom umetanja, koriste date EOF sekcije i umeću skrivene podatke nakon EOF markera što nema nikakve posledice na prenosni fajl i često se zanemaruje od strane izvršnog okruženja. Ova tehnika povećava veličinu prenosnog fajla, za količinu skrivenih podataka, te skrivanje prevelike količine informacija može izazvati sumnje u autentičnost prenosnog fajla. [4]

4.1.2 Zamena

Ova tehnika zamenjuje „nebitne“ bitove prenosnog fajla sa bitovima poruke koju treba sakriti. Nevažnim bitovima se smatraju oni bitovi koji se mogu izmeniti bez smanjivanja kvaliteta prenosnog fajla ili narušavanja njegovog integriteta. Na primer, u audio snimcima, svaki sempl zvuka sačinjen je od niza bitova. Ukoliko se najmanje značajan bit date sekvence modifikuje, data promena ima minimalni uticaj na percepciju finalnog zvuka, te ljudsko uho ne može da napravi razliku između originalne i izmenjenje verzije. Ova tehnika se oslanja se na nedostatak ljudskog auditornog sistema koji ne može da uoči razliku dva zvuka koja se minimalno razlikuju. [4]

4.1.3 Generativna

Za razliku od prethodnih tehnika, ova tehnika ne zahteva prethodno postojanje prenosnog fajla. Generativna tehnika čita podatke koje treba sakriti i od njih kreira nove podatke. Ovo je dinamička metoda koja kreira prenosni fajl na osnovu informacije koja se nalazi u poruci za skrivanje. Na primer, jedna generativna tehnika uzima karaktere skrivene poruke i pretvara ih u odgovarajuće audio freknvecije, koje se finalno mogu koristiti za generisanje audio fajla. Ova tehnika generiše prenosni fajl samo u svrhe sakrivanja poruke, pri čemu je rezultat originalni fajl, samim tim je otporan na uporedne tehnike za otkrivanje sakrivenih poruka. [4]

4.2 Tehnike audio steganografije

4.2.1 Najmanje značajan bit

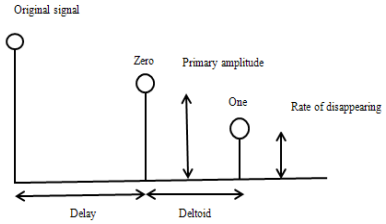
Metoda najmanje značajnog bita (*eng. Least significant bit - LSB*), je jedna od najranijih tehnika korišćenja u skrivanju informacija. Generalno, bazira se na ugrađivanju svakog bita poruke u najniži bit prenosnog audio fajla na željeni način. Tako da, audio sekvenca semplovanu sa 16kHz dozvoljava sakrivanje 16kbps podataka. LSB tehnika daje veliki kapacitet umetanja podataka i relativno je jednostavna za implementaciju ili kombinovanje sa drugim tehnikama. Međutim, sigurnost ove tehnike je dosta niska, te su

snimci podložni čak i jednostavnim napadima. Podaci stego-objekta bi, najverovatnije, bili uništeni amplifikacijom, dodavanjem šuma, filtriranjem i kompresijom sa gubitkom. Pored toga, podaci su sakriveni na deterministički način, te se podaci mogu otkriti jednostavnom eliminacijom cele ravni najmanje značajnih bitova.

Otpornost ove tehnike na distorziju i dodavanje šuma, može se postići povećavanjem dubine sloja umetanja sa 4. na 6. i 8., bez remećenja perceptualne transparentnosti stego fajla. [5]

4.2.2 Skrivanje u ehu

Metoda skrivanja u ehu ugrađuje skrivenu poruku u audio signal dodavajući kratak eho diskretnom signalu. Originalnom audio signalu se dodaju rezonanca, pri čemu stego signal zadržava iste perceptivne i statističke karakteristike. Skrivanje podataka u eho signalu zahteva manipulaciju sledećim parametrima: inicijalna amplituda, pomeraj (*eng. delay*) i brzina raspadanja takva da se dodatni eho ne primećuje. Eho predstavlja repliku originalnog zvuka, dodatu neko vreme nakon originalnog zvuka. Amplituda eha se mora smanjiti kako bi isti bio neprimetan. Pomeraj od 1ms između eho i originalnog signala stvara neprimetan efekat. Međutim, samo jedan bit tajne poruke bi se mogao enkodirati ukoliko bi se generisao samo jedan eho iz originalnog signala. Te se, pre procesa enkodiranja originalni signal deli na manje delove. Nakon procesa enkodiranja, delovi se spajaju kako bi se proizveo finalni signal. Ograničenje ove tehnike je nizak kapacitet skrivanja, jer bi bilo zahtevno dodavati eho za svaki bit koji je potrebno sakriti. [5]

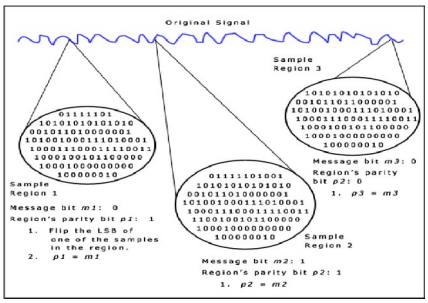


Slika 4: Dijagram tehnike skrivanja u ehu

4.2.3 Kodiranje pomoću bita parnosti

Ova tehnika predstavlja jednu od najrobustnijih tehnika audio steganografije. Umesto deljenja signala na individualne semplove, signal se deli na odvojene grupe semplova. Individualni semplovi se grupišu i izračunava se bit parnosti svake grupe ne bi li se umetnuli bitovi podataka jedan po jedan. Vrši se provera bita parnosti grupe, ukoliko se bit parnosti i bit podataka slažu, ne radi se ništa. U suprotnom, modifikuju se najmanje značajni bitovi bilo kog individualnog sempla date grupe kako bi bit parnosti postao jednak odgovarajućem bitu

poruke. Može se primetiti, da postoji više od jedne mogućnosti za enkodiranje tajnog bita, te signal može biti izmenjen na manje očigledan način. [5]



Slika 5: Dijagram kodiranja pomoću bita parnosti

4.2.4 Skrivanje u intervalima tišine

Data tehnika eksploatiše intervale tišine u signalima govora na jednostavan i efektivan način. Inicijalno je potrebno odrediti intervale tišine datog audio snimka i njihove respektivne dužine (broj semplova u datom intervalu). Vrednosti koje predstavljaju dužinu intervala tišine su smanjene nekom vrednošću takvom da $0 < vrednost < 2^n$, pri čemu je n broj bitova potrebnih za reprezentaciju elementa podataka koji je potrebno sakriti. Prenosni audio fajl sada ima novu dužinu intervala tišine. Povraćaj poruke vrši se pomoću $izmenjena\_dužina \% 2^n$. Ova tehnika ima dobru perceptivnu transparentnost ali je osetljiva na kompresiju. Modifikacije dužina intervala tišine dovode do nekorektne ekstrakcije podataka. [4]

4.2.5 Diskretna talasna transformacija

Diskretna talasna transformacija (*eng. Discrete Wavelet Transform – DWT*) podrazumeva sakrivanje poruke u najmanje značajnim bitovima talasnih koeficijenata originalnog signala. Često, podaci se skrivaju u celobrojnim talasnim koeficijentima, izbegavajući sekcije tišine audio fajla, ne bi li se smanjila verovatnoća detekcije izmenjenog fajla. Prenosni audio fajl razbija se u talasne koeficijente, svaki signal se skalira u odnosu na maksimalnu vrednost i broj bitova po semplu. Algoritam određuje broj bitova koji se mogu, na siguran način, sakriti u svaki sempl. Nakon toga, stego ključ se ugrađuje u detalje signala sa najmanjom frekvencijom, što čini ključ otpornijim na distorziju. Nedostatak ove metode leži u tome što se podaci mogu izgubiti prilikom povraćaja poruke, s obzirom da ova tehnika nije u potpunosti precizna.[4]

4.2.6 Kodiranje faze

Ova tehnika eksploatiše neosetljivost ljudskog auditornog sistema na relativne faze distinktnih spektralnih komponenti. Metoda se bazira na zameni odabranih komponenti faze originalnog spektra audio snimka sa tajnim podacima. Kodiranje faze ima veću toleranciju na distorziju signala u poređenju sa ostalim predstavljenim tehnikama.

Data tehnika može se objasniti sledećim koracima:

1. Originalni signal deli se u manje sekcije čine dužine su jednake veličini poruke koju treba enkodirati
2. Diskretna Furijeova transformacija primenjuje se na svaki segment ne bi li se kreirala matrica faza i Furijeovih transformacionih magnituda
3. Računa se razlika faza susednih elemenata
4. Promene u fazi između uzastopnih segmenata su jednostavne za detekciju, te se tajna poruka umeće u vektor faze prvog segmenta signala na sledeći način:

$$phase_{new} = \begin{Bmatrix} \frac{\pi}{2}, & message\,bit = 0 \\ -\frac{\pi}{2}, & message\,bit = 1 \end{Bmatrix}$$

5. Nova matrica faza se kreira korišćenjem razlike nove faze prvog segmenta sekcije i originalne faze. [5]

4.2.7 Širenje spektra

Ova tehnika koristi koncept širenja spektra (*eng. Spread Spectrum*), inicijalno korišćenog u komunikaciji podataka da osigura povoljan povraćaj signala poslatog kroz kanal u kome je prisutna određena doza šuma. Ovom metodom se skrivene informacije raširuju kroz širok opseg frekvencija. Odnos signala i šuma u svakom opsegu frekvencija mora biti dovoljno mali da se ne može detektovati prisustvo dodatnih podataka. Čak i ukoliko se delovi podataka uklone iz nekolicine opsega, količina informacija pristutna u ostalim opsezima je dovoljna za povraćaj podataka. Ispostavlja se da je veoma teško ukloniti podatke bez potpunog oštećenja prenosnog fajla. Ovaj pristup je veoma robustan i siguran način prenosa podataka, ali može da unese nasumični šum koji sprečava problem gubitka podataka. [6]

4.3 Metrike evaluacije

Evaluacija tehnike steganografije može se vršiti različitim vrstama poređenja. Neke od metrika koje se mogu koristiti pri evaluaciji su:

- **Učestalost bitske greške:** Tajnu poruku bi, idealno bilo, izdvojiti bez greške, međutim u realnom komunikacionom kanalu, može nastupiti greška. Greška se može meriti učestalošću bitske greške (*eng. bit error rate – BER*), koja predstavlja odnos broja grešaka sa totalnim brojem bitova u datom prenosnom objektu.
- **Srednja kvadratne greška:** Srednja kvadratna greška (*eng. mean square error – MSE*) dobija se upoređivanjem originalnog i stego objekta. Distorzija između dva objekta može se izračunati kao: $MSE = \frac{1}{n}\sum_{i=0}^{n-1}(O(i)-S(i))^2$, pri čemu su **O** i **S** originalni i stego signal respektivno.
- **Odnos maksimuma signala i šuma:** Cilj steganografije je sakrivanje tajne poruke u okviru prenosnog objekta tako da se očuva kvalitet prenosnog objekta. Odnos maksimuma signala i šuma (*eng. peak signal to noise ratio - PSNR*) se uglavnom koristi kao mera kvaliteta rekonstrukcije kod kompresija sa gubitkom, pri čemu je veća vrednost PSNR indikator višeg kvaliteta, odnosno manje distorzije. PSNR se može izračunati kao: $PSNR = 10\log_{10}(\frac{max(O)^2}{MSE})$.

5 Tehnika širenja spektra

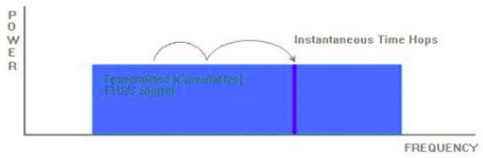
Tehnike širenja spektra se veoma često koriste u tipovima prenosa podataka, kao što su CDMA mobilne komunikacije. H. Lamar i G. Antheil su 1941 patentirali datu tehniku kao metodu za tajnu komunkaciju korišćenu u vojne svrhe. Ove tehnike podrazumevaju da se energija generisana u određenim opsezima ciljano širi u okviru frekventivnog domena, rezultujući signalom sa širim opsegom frekvencija.

5.1 Tipovi tehnika širenja spektra

Tehnike širenja spektra podrazumevaju sledeće tehnologije: skokovito frekventivno širenje spektra (*eng. Frequency Hopping Spread Spectrum, FHSS*), direktno sekvencijalno širenje spektra (*eng. Direct Sequence Spread Spectrum, DSSS*), skokovito vremensko širenje spektra (*eng. Time Hopping Spread Spectrum, THSS*).

5.1.1 Skokovito frekventivno širenje spektra (FHSS)

Ova metoda prenosi signal ubrzanim skakanjem prenosnika po različitim kanalima frekvencija, korišćenjem pseudoslučajne sekvence poznate i prenosniku i prijemniku. Redosledom skakanja s jedne na drugu frekvenciju upravlja generator pseudoslučajnih brojeva. Dokle god sve stanice koriste istu klicu pseudoslučajnog niza brojeva koje proizvodi generator i dokle god su vremenski sinhronizovane, one će istovremeno prelaziti s jedne na drugu frekvenciju. Glavna mana ove tehnije je mali propusni opseg. [7]



Slika 6: Grafički prikaz FHSS

5.1.2 Direktno sekvencijalno širenje spektra (DSSS)

DSSS deli podatke koji se trebaju poslati na manje delove a zatim se svaki deo dodeljuje određenom kanalu frekvencije kroz ceo spektar. Prenosnik koristi modulacionu

tehniku faznim pomerajem (koristeći generator pseudoslučajnih brojeva) kako bi modulirao svaki deo podataka sa sekvencom veće brzine prenosa podataka (*eng. bit rate*). [7]

5.1.3 Skokovito vremensko širenje spektra (THSS)

THSS prenosi kratke nalete informacija impulsima pseudoslučajne dužine trajanja ili prenošenjem na slučajnim pozicijama. Ova tehnika se može implementirati na dva načina: interval naleta informacija određuje se pseudoslučajnim generatorom ili informacije idu u isto vreme u svakom bitskom periodu, a pseudoslučajni generator menja svoju dužinu trajanja. [7]

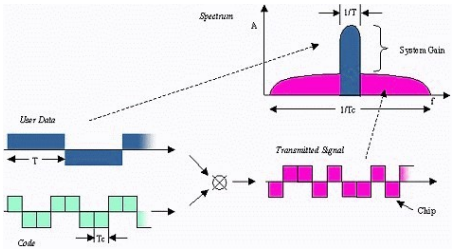
5.2 Audio steganografija korišćenjem DSSS

U ovom odeljku biće dat detaljan pregled tehnike direktnog sekvencijalnog širenja spektra, s obzirom da je to tehnika audio steganografije implementirana uz ovaj rad.

Tehnika direktnog sekvencijalnog širenja spektra, kao i druge tehnike širenja spektra, zasniva se na korišćenju generatora preudoslučajnih brojeva, odnosno u ovom konkretnom slučaju, sekvence pseudo šuma (*eng. Pseudo Noise Sequence, PN sequence*). Ova tehnika modulacije može se opisati sledećim koracima:

1. Pripremiti tajnu poruku koju je potrebno raširiti
2. Pripremiti PN sekvencu koja će se iskoristiti za modulaciju date poruke
3. Izvršiti modulaciju poruke sa PN sekvencom, pri čemu frekvencija sekvence mora biti veća (signal mora biti brži) od frekvencije signala tajne poruke
4. Umetnuti modulirani signal u prenosni audio fajl
5. Prijemnik za dekodiranje poruke mora imati istu PN sekvencu
6. Modulirati dobijeni signal sa istom PN sekvencom
7. Rezultujući signal je tajna poruka

Posebno je potrebno obratiti pažnju na ovu sekvencu, konkretno na njenu frekvenciju. Frekvencija PN sekvence, takođe nazvana stopa čipa (*eng. chip rate*), utiče na količinu podataka koja se može ugraditi u prenosni objekat. Ilustracija se može videti na sledećem primeru, pretpostavimo da je potrebno poslati poruku dužine 8 bitova, pri čemu je frekvencija semplovanja prenosnog audio snimka 44.1kHz. Tada je potrebno poslati podatke 44100 semplovima za svaki bit. To bi bila posledica PN sekvence frekvencije 44.1kHz. Sekvenca te frekvencije ima isto značenje kao umetanje 1 bita poruke po sekundi. Ova količina podataka se smatra malom, s obzirom da bi u audio snimku od 4 minuta (240s) se moglo umetnuti 30 bajtova podataka. Frekvenciju PN sekvence moguće je smanjiti pojačavanjem signala skrivene poruke. Pojačavanje signala može se izvršiti korišćenjem faktora jačine, koji će pojačati signal sa određenim odnosom. [2]



Slika 7: Grafički prikaz tehnike direktnog sekvencijalnog širenja spektra

5.2.1 Enkodiranje tajne poruke

Ova implementacija podrazumeva prenosni mono audio fajl WAV formata, bitske dubine 16. Tajna poruka data je kao bajt sekvenca, u ovom slučaju tekstualni fajl.

Tajnu poruku, datu kao sekvencu bajtova, potrebno je prvo konvertovati u sekvencu bitova, a zatim kodirati amplitudu signala kao: amplituda 1, ukoliko je bit 1, amplituda -1, ukoliko je bit 0. U nastavku poglavlja biće dato objašnjenje zašto je reprezentacija {-1, 1} povoljnija od {0, 1}. Amplitude signala tajne poruke mogu se predstaviti kao:

$$A=\{a_i|a_i\in\{-1,1\}\} \quad .$$

Dalje, učitavanjem prenosnog WAV fajla dobija se amplituda datog audio signala. Amplituda je predstavljena 16bitnim označenim integerom u opsegu $[2^{15}-1,-2^{15}+1]$, te je potrebno amplitudu datog signala takođe skalirati na opseg [-1, 1].

Sledeći korak je kreiranje dugačke PN sekvence čije vrednosti mogu biti -1 ili 1. Neka PN sekvenca ima frekvenciju (stopu čipa) **cr**, dok signal tajne informacije ima ukupno **n** signala (dužina sekvence bitova tajne poruke), tada generisana PN sekvenca mora biti dužine **cr x n**. PN sekvenca je tada:

$$PN=\{pn_i|pn_i\in\{-1,1\}\} \quad .$$

Dalje, vrši se modulacija signala poruke sa PN sekvencom. Prvo je potrebno generisati signal **B**, koji će biti distribuirani signal poruke **A cr** puta duži od originalnog signala **A**. Širenje informacija iz A, u signal B može se definisati kao:

$$B=\{b_i|b_i=a_j, j\cdot cr\le i\le(j+i)\cdot cr\} \quad .$$

Sada je moguće modulirati signale **B** i **P**, pri čemu se takođe vrši i množenje faktorom jačine **α**. Data poruka se sada može umetnuti u prenosni audio signal. Neka je **v** originalni prenosni signal, tada je **v`** prenosni signal koji sadrži datu tajnu poruku:

$$v'_i=v_i+\alpha\cdot b_i\cdot p_i \quad .$$

Ovim procesom će se generisati šum koji će biti dodat originalnom prenosnom audio fajlu. Ukoliko je faktor jačine previše veliki, šum će biti previše veliki i može dovesti do

oštećenja prenosnog objekta, te je neophodan pažljiv odabir frekvencije PN sekvence i faktora jačine. [2]

5.2.2 Dekodiranje tajne poruke

Efekat prethodno generisane PN sekvence je dodavanje slučajno kreiranog šuma, te je povraćaj sakrivene poruke zahteva korišćenje iste PN sekvence.

Dekodiranje se zasniva na množenju signala PN sekvence sa signalom dobijenog prenosnog objekta, pri čemu važi:

$$\sum_{i=j\cdot cr}^{(j+1)\cdot cr-1} p_i v'_i = \sum_{i=j\cdot cr}^{(j+1)\cdot cr-1} p_i v_i + \sum_{i=j\cdot cr}^{(j+1)\cdot cr-1} \alpha b_i p_i^2 \quad .$$

Može se zapaziti da $\sum_{i=j\cdot cr}^{(j+1)\cdot cr-1} p_i v_i \sim 0$ za veliki broj semplova (visoku frekvenciju slučajne sekvence). Ovo je posledica slučajne vrednosti PN sekvence koja prouzrokuju približavanje sume signala 0 ili određenom pragu.

Drugi sabirak takođe ima interesatnu karakteristiku. Naime, kako vrednosti PN sekvence mogu biti samo -1 ili 1, $p_i^2 = 1$, te dati izraz postaje $\sum_{i=j\cdot cr}^{(j+1)\cdot cr-1} \alpha b_i$. Kako je $\mathbf{b_i}$ takođe definisano vrednostima -1 ili 1, može se zaključiti da ukoliko je vrednost izraza veća od 0, preuzeta informacije je 1, a ukoliko je vrednost izraza manja od 0, informacija je 0. Ova karakteristika je motivacija iza odabira domena prethodno definisanih signala $\mathbf{P}$ i $\mathbf{B}$.

Prethodno objašnjenje pokazuje da vrednost αb_i mora biti veća od određenog praga ne bi li se jednoznačno povratila tajna informacija. Ponovo se zaključuje da je potrebno pažljivo odabrati faktor jačine α. [2]

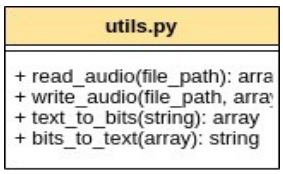
6 Implementacija i rezultati

Uz ovaj rad data je implementacija audio steganografije korišćenjem metode direktnog sekvencijalnog širenja spektra (DSSS) opisane u prethodnom poglavlju. Implementacija je izvršena u *Python* programskom jeziku, pri čemu je za učitavanje i snimanje audio fajlova korišćena biblioteka *SoundFile*.

Steganografija kao prenosni objekat koristi 16-obitni jednokanalni (mono) audio snimak WAV formata, dok je tajna poruka data u formatu tekstualnog *.txt* fajla.

U ovom poglavlju biće opisani značajniji implementacioni detalji kao i rezultati dobijeni implementacijom date tehnike.

6.1 Pomoćne metode

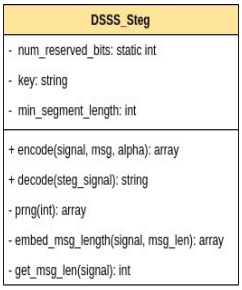


Slika 8: Dijagram utils.py modula

Implementacija date metode steganografije zahteva rad sa audio snimcima odnosno određenu dozu obrade podataka reprezentovane nizovima bitova. Korišćene pomoćne metode nalaze se u *utils.py* modulu, pri čemu su najznačajnije predstavljene na (Slika 8).

Prvenstveno je potrebno učitati audio snimak, a kasnije i sačuvati modifikovani stego audio signal, što je, kao što je već napomenuto, izvršeno uz pomoć *SoundFile* biblioteke. Učitavanje fajla vraća više dimenzionalni niz, pri čemu je broj dimenzija određen brojem kanala zvučnog zapisa, dužine broja semplova (frejmova). S obzirom da ova implementacija podrazumeva jednokanalne zapise, oni su reprezentovani jednodimenzionalnim nizovima. Vrednosti pojedinačnih semplova date su kao racionalni brojevi u opsegu [-1.0, 1.0). Rate semplovanja je takođe povratna vrednost metode *read_audio*. Ostale pomoćne metode odnose se na konverziju bitske reprezentacije u tekstualnu ili brojčanu i obrnuto.

6.2 DSSS_Steg klasa



Slika 9: Dijagram DSSS_Steg klase

Metoda direktnog sekvencijalnog širenja spektra enkapsulirana je DSSS_Steg klasom, pri čemu je njena struktura data na (Slika 9). Najznačajnije metode ove klase predstavljaju *encode()* i *decode()* metode, pri čemu su ove metode implementirane u skladu sa algoritmima za enkodiranje i dekodiranje tajne poruke opisanim u prethodnom poglavlju.

Atribut klase *key* predstavlja stego ključ kojim se vrši generisanje jedinstvene PN sekvence u procesima enkodiranja i dekodiranja. *Min_segment_len* atribut predstavlja minimalnu stopu čipa (frekvenciju PN sekvence), potrebnu za valjano skrivanje tajne poruke bez oštećenja originalnog signala. Pored toga, poslednjih *num_reserved_bits* bitova signala rezervisano je za skrivanje dužine tajne poruke, ne bi li se postigla takozvana *slepa steganografija*.

6.2.1 Enkodiranje podataka

Encode() metoda uzima kao parametre originalni audio signal i originalnu tajnu poruku u string reprezentaciji, a kreira stego signal kao kombinaciju ulaznih parametara i generisane PN sekvence. Na osnovu minimalne dužine segmenta PN sekvence i dužine originalne poruke određuje se stvarna dužina segmenta sekvence kao i broj ponavljanja date sekvence. Kao što je prethodno opisano, svaki bit originalne poruke modulira se jednim segmentom PN sekvence, te se može zaključiti da poruke čija je dužina bitske reprezentacije veća od prethodno izračunatog dozvoljenog broja ponavljanja sekvence, ne može biti sakrivena u datom audio signalu, što ima za posledicu prekidanje procesa endkodiranja. Generisanje PN sekvence vrši se *prng()* metodom, čiji će detalji biti dati u nastavku, a zatim se sekvenca proširuje na dužinu koja odgovara proizvodu broja segmenata i dužini pojedinačnog segmenta. Signal poruke se ugrađuje u originalni audio signal na način opisan u prethodnom poglavlju, pri čemu ulazni parametar *alpha* predstavlja aktor jačine. Proces enkodiranja završava se umetanjem dužine originalne poruke na kraj originalnog signala, metodom *embed_msg_len()*.

6.2.2 Generisanje PN sekvence

Prethodno je pomenuto da se generisanje PN sekvene vrši *prng()* metodom. Ova metoda koristi zadati ključ, dat kao string, ne bi li generisala celobrojnu vrednost, koja će onda biti korišćenja kao klica (*eng. seed*) *np.random* metode. Celobrojna vrednosti se dobija kao: $\sum_i ord(key_i) * i$. Nakon postavljanja klice, generiše se nasumična sekvenca zadate dužine čije su vrednosti 1 ili -1. Seed-ovanje istom vrednošću garantuje da će sekvenca biti ista pri korišćenju istih ključeva, pri svakom kreiranju.

6.2.3 Moduliranje signala tajne poruke

Metoda kojom se originalna tajna poruka modulira ne bi li dostigla odgovarajuću dužinu, definisana je u okviru *mixer.py → mix().* Mix() metoda funkcioniše sasvim jednostavno, množeći svaki bit poruke nizom jedinica dužine zadatog segmenta. Nakon toga vrši se promena vrednosti signala poruke iz (0, 1) na (-1, 1), iz razloga prethodno opisanih. Dodatna modifikacija izvršena je primenom Hanovog filtera, iz porodice filtera kosinusne sume, koji se koristi za uglačavanje funkcija, pogotovo u oblasti procesiranja digitalnih signala.

6.2.4 Dužina tajne poruke

Funkcije za umetanje i ekstrakciju dužine poruke, takođe imaju relativno jednostavan algoritame. Naime, umetanje, podrazumeva pretvaranje dužine poruke u bitsku reprezentaciju a zatim enkodiranje 0 i 1 kao minimalne i maksimalne vrednosti, respektivno, krajnjeg dela originalnog signala. Tako enkodirana dužina poruke zamenjuje bitove originalnog signala. Metoda ekstrakcije dužine poruke vrši obrnut proces, uzimajući krajnje bitove stego signala i prevođenje nazad u standardnu binarnu reprezentaciju.

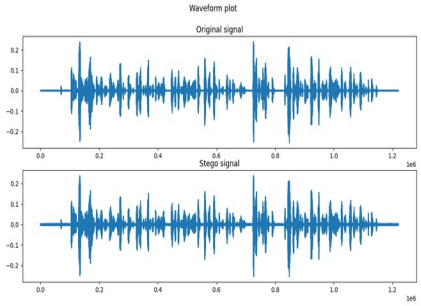
6.2.5 Dekodiranje

Finalno, *decode()* metoda uzima stego signal i iz njega ekstraktuje tajnu poruku, koja je ujedno i povratna vrednost date metode. Prvo je potrebno ekstraktovati dužinu poruke iz dobijenog signala, metodom *get_msg_length()*, a zatim se na, sličan način opisan u prethodnom poglavlju, vrši i ekstrakcija tajne poruke.

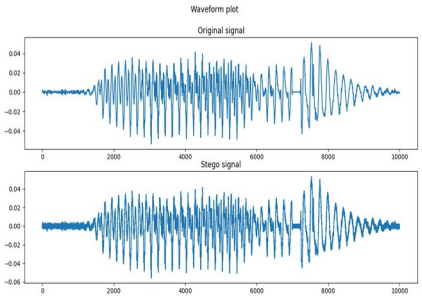
6.3 Rezultati i evaluacija

Finalno, rezultati steganografije mogu se evaluirati standardnim metodama opisanim u 4.3, a datim u okviru *evaluate.py* modula. Takođe, se rezultati mogu i vizuelizovati, u vidu grafika originalnog i stego signala.

Tehnika je prvo isprobana na glasovnom audio snimku, snimljenom od strane autora, u 16-obitnom WAV formatu, frekvencije semplovanja 44.1kHz, dužine ~ 27s. Različiti parametri za faktor jačine i frekvenciju PN sekvence su isprobani, a kao metoda evaluacije korišćene su metrike opisane u 4.3, kao i slobodna procena autora o kvalitetu dobijenog stego snimka. Kako frekvencija PN sekvence direktno utiče na moguću dužinu tajne poruke, nastojalo se da frekvencija bude što manja. Za dati snimak se ispostavlja da frekvencija PN sekvence manja od 2^{12}, rezultuje ili nekorektnim dekodiranjem poruke ili prevelikim šumom Korektno dekodiranje se može postići povećavanjem faktora jačine ali se time postiže jasno uočljiv šum, sa druge strane, smanjenjem faktora jačine se čujnost šuma svoji na minimum ali se dobija poruka sa određenom dozom greške (BER ~ 0.2, 0.1). Treba napomenuti da, s obzirom na mali BER koeficijent, poruka, i ako sadrži greške je i dalje razumljiva, te nije potpuno neupotrebljiva u sistemu koji bi posedovao određenu toleranciju na greške, sa druge strane, ovaj rezultat je neprihvatljiv u sistemu koji zahteva potpunu preciznost poruke. Pronađeno je da za dati snimak, optimalne rezultate daje PN sekvenca frekvencije 2^{12}, što podrazumeva maksimalnu dužinu tajne poruke od 30 bajtova, što je relativno kratka poruka. Valjan faktor jačine bi bio ~ 0.002, slično kao i u prethodnom slučaju veći faktor rezultuje većim šumom a manji neispravnom porukom.



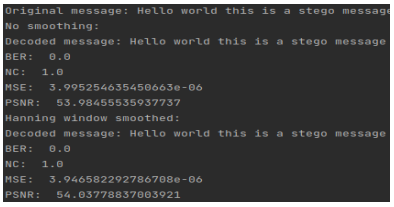
Slika 11: Grafici originalnog i stego signala



Slika 10: Grafici delova originalnog i stego signala

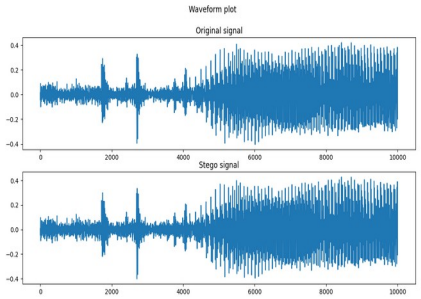
Na (Slika 11, Slika 10) prikazani su grafici originalnog i dobijeno stego signala korišćenjem datih parametara. Može se uočiti vidljivost dodatnog šuma, pogotovo u sekcijama signala u kojima je tišina, ipak dati šum je dovoljno mali da ne unosi prevelike smetnje pri slušanju snimka, već zvuči kao tipičan šum kreiran prilikom snimanja audio zapisa.

Na (Slika 12), dati su i statističke evaluacije date metode, iz kojih se može uočiti njena valjanost. Takođe, može se videti da korišćenje Hanovog filtera poboljšava PSNR i smanjuje MSE.

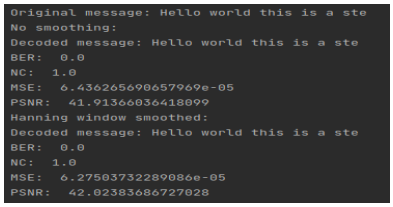


Slika 12: Metrike snimljenog audio fajla

Poslednji primer na kome je tehnika testirana je deo poznatog govora Martin Luther King-a, „I Had a Dream“, semplovanog frekvencijom od 22.05kHz. Kako je frekvencija semplovanja datog snimka duplo manja od prethodnog, zaključuje se da je i raspoloživ prostor za sakrivanje tajne poruke prepolovljen. Ipak, ovaj snimak, zbog relativnog prisustva inicijalnog šuma, dozvoljava veću manipulaciju DSSS parametrima. Ispostavlja se da se faktor jačine šuma može bez problema povećati na ~ 0.8, što dozvoljava smanjenje frekvencije PN sekvence i prouzrokuje da se maksimalna veličina tajne poruke nije prepolovila u odnosu na prethodno opisan slučaj. Na (Slika 13, Slika 14) su prikazani rezultati korišćenjem datih parametara.



Slika 13: Originalni i stego signal: "I had a dream"



Slika 14: Metrike "I had a dream" fajla

7 Zaključak

Skrivanje informacija u okviru zvučnih zapisa, odnosno audio steganografija, prestavlja interesantan i izazovan, i ako manje popularan, tip steganografije. U ovom radu predstavljene su različite tehnike koje se mogu primeniti u oblasti audio steganografije, pri čemu je detaljno obrađena tehnika direktnog sekvecijalnog širenja spektra. Ova tehnika pokazuje se kao veoma robustna i otporna na napade šumom, što su ujedno njene glavne prednosti. Međutim, pokazuje se da je ova metoda sklona dodavanju šuma koji se može opaziti, kao i da je količina podataka koju je moguće sakriti, uz korektnu ekstrakciju, relativno limitirana.

Implementacija date tehnike data uz ovaj rad daje relativno solidne rezultate. Pokazuje se da je pažljivim odabirom parametara moguće umetnuti podatke pri čemu dodatni šum ostaje nečujan. Takođe, ispostavlja se da auditorno „bogatiji“ snimci ostavljaju više prostora za skrivanje informacija, s obzirom da se može dodati veća količina pseudo slučajnog šuma, što je i za očekivati. Data implementacija bi se mogla unaprediti primenom tehnike širenja spektra u frekventivnom domenu, prevođenjem signala korišćenjem Furijeove ili neke druge transformacije, umesto u vremenskom domenu u kome je data. Pored toga, poruka bi se pre umetanja mogla enkriptovati, što bi zajedno sa korišćenjem frekventivnog domena povećalo pouzdanost i sigurnost ove metode. Dodatni prostor za skrivanje informacija, mogao bi se dobiti korišćenjem stereo audio snimaka, s obzirom da bi zvuk bio predstavljen sa dva talasa, pri čemu se bi se identična implementacija tehnike mogla primeniti na oba talasa pojedinačno. Takođe, postoje naznake da se interesantni rezultati mogu dobiti kombinovanjem ove tehnike sa nekom od drugih tehnika predstavljenim u ovom radu.

Na kraju, može se zaključiti da tehnika direktnog sekvencijalnog širenja spektra predstavlja dobru tehniku u oblasti audio steganografije, ali da i dalje postoji prostor za njeno unapređivanje.

Literatura

1: ANSI/ASA S1.1-2013. Available at: https://asastandards.org/asa-standard-term-database.

2: R. Nugraha, "Implementation of Direct Sequence Spread Spectrum Steganography on Audio Data," in *2011 International Conference on Electrical Engineering and Informatics*, Bandung, Indonesia, 2011.

3: K. U. Singh, "A Survey on Audio Steganography Approaches," in *International Journal of Computer Applications*, vol. 95, No. 14, 2014.

4: Y. Bassil, "A Two Intermediates Audio Steganography Technique," in *Journal of Emerging Trends in Computing and Information Sciences (CIS)*, vol. 3, No. 11, 2012.

5: N. Kaur and S. Behal, "Audio Steganography Techniques-A Survey," in *Navneet Kaur Int. Journal of Engineering Research and Applications*, vol. 5, No. 6, 2014.

6: H. Kaur and J. Rani, "A Survey on different techniques of steganography," in *MATEC Web of Conferences*, , 2016.

7: W. Q. Chang et al., "Robust Audio Steganography using Direct-Sequence Spread Spectrum Technology," 2007.