Primena Šamirovog tajnog deljenja na audio fajlove

Mihajlo Milojević SV57/2023

1. Opis projekta

U ovom projektu, koristiću princip **Šamirovog tajnog deljenja**[1] (Shamir's Secret Sharing) kako bih omogućio bezbednu dekompoziciju i rekonstrukciju audio fajlova. Šamirovo tajno deljenje (SSS) je kriptografski algoritam koji omogućava podelu tajne (npr. ključa, fajla ili informacije) na *n* delova (senki) tako da je za rekonstrukciju tajne potrebno najmanje *k* od tih delova. Svaka pojedinačna senka ne otkriva informacije o originalnoj tajni. Umesto klasične primene na slike, ovaj projekat će se fokusirati na audio fajlove specifično .wav format. Glavni cilj je podela audio fajla na „senke" (eng. shadows) koje pojedinačno ne otkrivaju informacije o originalnom fajlu, dok će se uz minimalan potreban broj senki moći rekonstruisati originalni audio.

2. Skup podataka

Za potrebe testiranja koristiću nekoliko audio fajlova u ***.wav***[2] formatu raznih dužina koji se mogu preuzeti sa javno dostupnih baza podataka ili generisati korišćenjem sintetizovanih tonova (npr. sinusni, pravougaoni, trougaoni talasi). Biće izabrani i audio fajlovi sa ljudskim govorom, kao i pesme kako bih testirao sposobnost algoritma da pravilno dekomponuje i rekonstruiše ove fajlove.

3. Ciljevi projekta

- Kreirati sistem za bezbedno dekomponovanje audio fajla na senke koristeći Šamirov algoritam.
- Implementirati opciju rekonstrukcije audio fajla iz zadatog broja senki (*k* od *n*) i osigurati da, u slučaju manjeg broja dostupnih senki, rekonstrukcija bude nemoguća ili proizvede neprepoznatljiv zvuk.
- Omogućiti učitavanje i čuvanje audio fajlova i senki na disk radi daljeg poređenja.
- Vizuelizovati i uporediti sličnost između originalnog i rekonstruisanog zvuka.

4. Metodologija

1. Pretprocesiranje

- Učitavanje audio fajla i konverzija u niz uzoraka (diskretizacija u vremenskoj domeni).

2. Dekompozicija (Enkripcija)

- Definisanje parametara ***n*** (broj senki) i ***k*** (minimalan broj senki za rekonstrukciju).
- Primena Šamirovog algoritma kako bih generisao *n* senki
- Čuvanje senki na disk u obliku fajlova.

3. Rekonstrukcija (Dekripcija)

- Učitavanje najmanje ***k*** senki.
- Rekonstrukcija originalnog zvuka primenom Lagranžove interpolacije.

4. Provera ispravnosti

- Analiza sigurnosti senki - provera da li manji broj od *k* senki otkriva informacije o originalnom zvuku.
- Analiza razlike između originalnog i rekonstruisanog zvuka koristeći vizuelizaciju spektrograma i metrika poput **SNR**[3] (Signal-to-Noise Ratio) i **MSE**[4] (Mean Squared Error).

5. Očekivani rezultati

- Potpuno funkcionalna implementacija dekompozicije i rekonstrukcije audio fajlova koristeći Šamirovo tajno deljenje.
- Jasan vizuelni i numerički prikaz razlika između originalnog i rekonstruisanog zvuka.
- Prikaz neprepoznatljivog zvuka prilikom pokušaja rekonstrukcije sa manje od k senki.

6. Tehnologija

Projekat će biti izrađen u programskom jeziku **python**[5]. Od biblioteka će biti korišćen **numpy**[6] za čuvanje i obradu vrednosti signala, matplotlib za vizelizaciju audio signala, **scipy**[7] za računanje metrike i Lagranžovu interpolaciju, ugrađeni **wave**[8] paket za rad sa .wav fajlovima.

Moguće je dodati biblioteke koje nisu navedene u ovom predlogu ukoliko bude bilo potrebe za njihovu upotrebu.

7. Literatura

[1] https://en.wikipedia.org/wiki/Shamir%27s_secret_sharing

[2] https://en.wikipedia.org/wiki/WAV

[3] https://en.wikipedia.org/wiki/Signal-to-noise_ratio

[4] https://en.wikipedia.org/wiki/Mean_squared_error#Interpretation

[5] https://www.python.org/

[6] https://numpy.org/

[7] https://scipy.org/

[8] https://docs.python.org/3/library/wave.html

[9] https://www.researchgate.net/publication/274248847_Audio_Secret_Management_Scheme_Using_Shamir's_Secret_Sharing