# Predlog projekta VIII - Osnove i granice funkcionalnosti programa za prepoznavanje muzickih numera

Potreba za prepoznavanjem numera i njihovih slicnosti postoji od razonodnih upotreba do komercijalnih. Programi razvijani u ove svrhe kao sto su Shazam, Soundhound i Adobe Audition postoje vec dugi niz godina. Ovi softveri kao osnovnu funkciju imaju kreiranje digitalnog otiska pesme i poredjenje dobijenih otisaka sa vec postojecim pesmama iz svoje baze. Ovaj nacin upotrebe se pretezno koristi u razonodne svrhe radi pronalazenja pesme u izlasku. Ali druga namena za koju se mogu koristiti neki od gorenavedenih softvera je prepoznavanje plagijata.

Ovaj rad se fokusira na teorijski princip funkcionisanja aplikacije *Shazam*, kao i testiranja granica softvera i mogucnosti njegovog zavaravanja. Koriscenjem literature koja je dostupna na internetu vezana za osnove njegovog funkcionisanja:

1. Blog Ane Harris na sajtu *Medium*

[*https://medium.com/@anaharris/how-does-shazam-work-d38f74e41359*](https://medium.com/@anaharris/how-does-shazam-work-d38f74e41359)

1. Clanak Jovana Jovanovica na stranici toptal

<https://www.toptal.com/algorithms/shazam-it-music-processing-fingerprinting-and-recognition>

1. Naucni rad Avery Li-Chun Wang *An Industrial-Strength Audio Search Algorithm*

Osnovni princip rada programa *Shazam* bazira se na kreiranjem digiralnog otiska muzicke numere. Otisak se formira pomocu podataka o vrhovima koje dobijamo sa spektrograma datog snimka. Pomocu odredjenog broja proja vrhova, njihovih frekvenci i vremenskih udaljenosti relativno jednih u odnosu na druge dobijamo dobru osnovu za podatke potrebne za formiranje digitalnog otiska. Odatle koriscenjem *hash* mapa svodimo dobijene podatke na znatno manje prostora sto nam omogucuje brze poredjenje sa bazom podataka.

Eksploatacijom poznavanja frekvencijskih opsega u kojima aplikacija pretezno formira digitalne otiske cemo testirati njene granice prepoznavanja numera koje su filtrirane ili imaju unetu distorziju na odredjenim frekvencama kao i prisustvo buke i suma u generalnom opsegu od 20Hz do 20kHz. Nakon ovoga ce isti snimci biti reprodukovati na slusalicama osobama koje su upoznate sa datim numerama radi poredjenja mogucnosti prepoznavanja softvera i ljudskog uha.

Ovaj rad je dalje moguce nastaviti kreiranjem softvera koji koriscenjem ovakvog digitalnog otiskivanja prepoznaje plagijate i njihov stepen slicnosti.

**Opšte napomene:**

Svaki projekat treba da sadrži:

* Kratak opis problema koji se razmatra u okviru projekta kao i definicije svih pojmova i veličina koje se koriste
* Kratak pregled metodologije na koju se oslanjamo u realizaciji projekta (navodi iz literature koju ste koristili)
* Opis baze muzičkih uzoraka na kojoj se vrši eksperiment
* Opis toka istraživanja koje ste sproveli
* Vrednovanje dobijenih rezultata i diskusija slabosti i robusnosti metode (da se nabroje probelmi sa kojima ste sretali u realizaciji)