

UNIVERZITET SINGIDUNUM
BEOGRAD
DEPARTMAN ZA POSLEDIPLOMSKE STUDIJE
I MEĐUNARODNU SARADNJU

DOKTORSKA DISERTACIJA
JEDNA KLASA SISTEMA ZA GENERISANJE KRIPTOLOŠKIH
KLJUČEVA NA OSNOVU BIOMETRIJSKIH PODATAKA

MENTOR

prof. dr Milan Milosavljević

KANDIDAT

Saša Adamović, master

BEOGRAD, 2013.

Sažetak

Procenjuje se da celokupna oblast zaštite informacija u narednom periodu treba da obezbedi radikalan kvalitativan skok i pomeranje osnovne paradigme računarske sigurnosti ka paradigmi informaciono teorijske sigurnosti. Ovaj pomeraj bi omogućio sintezu čitave klase zaštitnih mehanizama, čija kompromitacija više ne bi zavisila od računarske snage napadača. Interesantno je zapaziti da preduslov za ovaj poduhvat leži u razvoju novih metoda digitalne obrade signale koja na fizičkom sloju savremenih računarsko komunikacionih mreža treba da obezbede ekstrakciju dovoljne količine čisto slučajnih nizova pogodnih za primenu u novim protokolima uspostavljanja i razmene kriptoloških ključeva maksimalne dužine i neodređenosti. U domenu pouzdane i za praksu prihvatljive autentifikacije, biometrija postaje značajan izvor kriptoloških parametara. Digitalna obrada izvornih biometrijskih signala je od presudnog značaja kako za uspostavljanje tako i za ekstrakciju autentifikacionih kodova maksimalne entropije. U domenu upravljanja autorskim pravima digitalnog sadržaja, digitalna obrada signala nosioca i utiskivanje digitalnih pečata, vidljivih i nevidljivih, takođe predstavlja nezaobilazan gradivni element. Sve ovo ukazuje na novo mesto ove discipline u okviru opšte teorije i prakse sinteze sistema za zaštitu informacija. U predloženoj tezi se identifikuje, analizira i razvija klasa podsistema digitalne obrade signala koje ispunjavaju nove zahteve proistekle iz ovih novih trendova u zaštiti informacija u savremenim računarsko komunikacionim sistemima. Prvenstveno se radi na razvoju jedne klase sistema za generisanje kriptoloških ključeva na osnovu biometrije irisa.

Abstract

It is estimated that the total area of information security in the future should provide a qualitative leap radical and fundamental paradigm shift to a new paradigm of theoretical computer security information security. This would allow the synthesis of the whole class of protection mechanisms, whose compromise would no longer be dependent on the computing power of an attacker. It is interesting to note that the prerequisite for this thesis is the development of new methods of digital signal processing that the physical layer of modern communication networks should provide sufficient information. Purely random extraction can be used in the establishment of new protocols and encryption key exchange with a maximum length and imprecision. In the field of reliable and acceptable practice for authentication, biometrics becomes an important source parameters cryptology. Source of digital signal processing biometric is essential to the establishment and to extract a maximum entropy codes authentication. In the field of copyright management of digital content on digital media and digital embossing seals, visible and invisible, are also an essential element. All this leads to a new place in this discipline in the theory and practice of general synthesis systems to protect information. In the proposed thesis identifies, analyzes and develops a class of subsystems digital signal processing to meet the new demands arising from these new developments in the protection of information in computer systems and modern communication. Working primarily on the development of a class of systems of encryption keys based on iris biometrics.

Zahvalnost

Zahvaljujem se mentoru prof. dr Milanu Milosavljeviću na neizmernoj podršci dobijenoj iz oblasti teorije informacija, kriptologije i filozofije, čije je duboko razumevanje nauke i ovog područja omogućilo da se rad pokrene, pravilno usmeri i završi na vreme.

Zahvaljujem se neizmerno svojoj porodici, ocu Željku, majci Slavici i bratu Milanu koji su topline doma podržavali moj rad na ovom velikom projektu.

Zahvaljujem se Univerzitetu Singidunum na čelu sa rektorom prof. dr Milovanom Stanišićem koji je prihvatio ovaj rad i obezbedio uslove za njegovu realizaciju.

Indeks slika

Slika 1. Izvor biometrijske kriptografije	20
Slika 2. Šema za generisanje biometrijskog ključa	29
Slika 3. Primena biometrijskih sistema	34
Slika 4. Piramida znanja - PIZP	36
Slika 5. Konveksnost funkcije $\ln_x$ i tangenta u tački $x=1$	42
Slika 6. Dvonivovski verovatnosni sistem mogućnosti	43
Slika 7. Entropija slučajne promenljive sa dva ishoda	45
Slika 8. Dve slučajne promenljive	53
Slika 9. Uslovna entropija	54
Slika 10. Uzajamna informacija	55
Slika 11. Šema za generisanje iris inf. izvora	61
Slika 12. I- Slika oka , II- detekcija kapaka i trepavica, III-označavanje iris regiona, IV- segmentiran iris	62
Slika 13. Daugmanov Rubber sheet model	63
Slika 14. Realni i imaginarni deo impulsivnog odziva	65
Slika 15. Faza kodovanja	67
Slika 16. Model komunikacionog kanala	69
Slika 17. Primena Hamingvog blokovskog zaštitnog koda	71
Slika 18. Primer konvolucionog kodera	72
Slika 19. Dijagram stanja konvolucionog kodera	73
Slika 20. RS kodna reč	75
Slika 21. Šema konkatentnog koda	77
Slika 22. Jedan primer blokovskog interlivinga	78
Slika 23. Šema blokovskog interlivinga	79
Slika 24. Šema konvolucionog interlivera	80
Slika 25. Generička šema predloženog rešenja	81
Slika 26. Proces generisanja iris koda	82
Slika 27. Primeri slika irisa iz “CASIA-Iris V4” baze podataka	84
Slika 28. Računanje lokalne entropije nad slikom normalizovanog irisa	86
Slika 29. Histogram prosečne vrednosti intenziteta piksela za sliku irisa	87
Slika 30. 3D informacioni model lokalne entropije (3x3)	88
Slika 31. 3D informacioni model lokalne entropije (9x9)	89
Slika 32. Vrednosti lokalne entropije po segmentima	90
Slika 33. Lokalna entropija za prvi segment [3x3]	90
Slika 34. Lokalna entropija za prvi segment [9x9]	90
Slika 35. Lokalna entropija za drugi segment [3x3]	91
Slika 36. Lokalna entropija za drugi segment [9x9]	91
Slika 37. Iris kod interpretiran kao crno-bela slika	93
Slika 38. Vrednosti lokalne entropije za različite vrednosti parametra λ	97
Slika 39. Primer za aproksimaciju II reda	98
Slika 40. Entropija prvog segmenta irisa po modelima aproksimacije	100

Slika 41. Uzajamna informacija po redovima u iris kodu za prvi segment (različiti irisi)....	103
Slika 42. Uzajamna informacija po redovima u iris kodu za prvi segment (isti irisi).....	104
Slika 43. Grafička prezentacija prostora biometrijskog ključa	105
Slika 44. Analiza ekvivalentnog kanala	107
Slika 45. Binarna struktura iris koda	109
Slika 46. Biometrijski komunikacioni kanal sa šumom	111
Slika 47. Broj grešaka između različitih irisa.....	113
Slika 48. Broj grešaka između istih irisa	114
Slika 49. Kapacitet greške za RS(127,20).....	115
Slika 50. Maksimalni kapacitet RS koda.....	116
Slika 51. Jedna šema primene sistema	121
Slika 52. LG Iris 4000 skener irisa	124
Slika 53. Bimodalni biometrijski sistem	124

Indeks tabela

Tabela 1. Analiza vejevlet filtra korišćenih u fazi kodovanja	68
Tabela 2. Statistički pokazatelji za dobijeni histogram.....	87
Tabela 3. Prosečne vrednosti lokalne entropije za četiri veličine prozora.....	91
Tabela 4. Statistički pokazatelji po segmentima	92
Tabela 5. Prostiranje lokalne entropije u iris kodu za parametar filtra ($\lambda_n = 12$).....	93
Tabela 6. Prostiranje lokalne entropije u iris kodu za parametar filtra ($\lambda_n = 18$).....	94
Tabela 7. Prostiranje lokalne entropije u iris kodu za parametar filtra ($\lambda_n = 24$)	95
Tabela 8. Modelovanje iris informacionog izvora primenom Šenonovih modela aproksimacije	99
Tabela 9. Prosečna uzajamna informacija između različitih irisa.....	102
Tabela 10. Prosečna uzajamna informacija između istih irisa	103
Tabela 11. DOF – broj stepeni slobode iris koda.....	106
Tabela 12. Uticaj slučajnog interlivera na uzajamne informacije i entropiju	110
Tabela 13. Teorijski kapacitet RS koda	114
Tabela 14. Stvarni kapacitet RS koda	115
Tabela 15. Performanse sistema kada je $n = 127$	116
Tabela 16. Broj pogrešnih bitova u ključu u slučaju iste i različite permutacije	117
Tabela 17. Poređenje sa postojećim rezultatima u ovoj oblasti	119

Indeks listinga programskih kodova

Listing 1. Šenonov model aproksimacije II reda.....	126
Listing 2. Šenonov model aproksimacije III reda	126
Listing 3. Šenonov model aproksimacije IV reda	127
Listing 4. Šenonov model aproksimacije V reda	127
Listing 5. Uzajamna informacija između dva irisa	128
Listing 6. Generisanje ključa veličine 200 bitova	129
Listing 7. Generisanje ključa veličine 300 bitova	130
Listing 8. Generisanje ključa veličine 400 bitova	131

Sadržaj

1. Uvod.....	10
1.1. Motivacija i opšta razmatranja	10
1.2. Predmet istraživanja	11
1.3. Doprinosi disertacije	12
1.4. Struktura disertacije.....	14
2. Pregled u oblasti istraživanja	16
2.1. Nedostaci biometrijskih sistema.....	17
2.2. Biometrijska kriptografija	19
2.3. Prednosti i značaj biometrijske kriptografije	20
2.3. Razvoj oblasti biometrijske kriptografije.....	21
2.3.1. Zaštita privatnosti biometrijskih podataka	22
2.3.2. Generisanje ključeva iz biometrijskih podataka	26
2.3.3. Generisanje ključeva na osnovu biometrijskih podataka	28
2.4. Primene tehnologije zasnovane biometrijskoj kriptografiji.....	34
3. Teorijske osnove istraživanja.....	36
3.1. Pojam neodređenosti i informacije.....	39
3.1.1. Entropija	44
3.1.2. Informacija i njena mera.....	49
3.1.3. Grafička prezentacija informaciono-teorijskih mera	52
3.1.4. Šenonova matematička teorija komunikacije	55
3.2. Iris kao biometrijski informacioni izvor.....	61
3.2.1. Segmentacija.....	61
3.2.2. Noramalizacija.....	62
3.2.3. Kodovanje.....	64
3.2.4. Jedinstvenost iris biometrije	67
3.3. Zaštitni kodovi.....	69
3.2.1. Klase zaštitnih kodova.....	71
3.2.2. Rid-Solomon kodovi.....	73
3.2.3. Tehnike zaštitnog kodovanja	75
3.2.4. Interliveri	78
4. Novi sistem za generisanje ključeva na osnovu biometrije irisa	81

4.1. Generička šema predloženog rešenja	81
4.2. Postavka eksperimentalnog okruženja	83
4.3. Informaciono-teorijska analiza Iris biometrijskog izvora	85
4.3.1. Informaciona analiza tekstone slike irisa	85
4.3.2. Informaciona analiza iris koda	92
4.3.3. Šenonovi modeli aproksimacije za modelovanje iris inf. izvora	98
4.3.4. Analiza uzajamne informacije između istih i različitih irisa	101
4.3.5. Kompliksnost iris koda - DOF	106
4.4. Izbor adekvatnog zaštitnog koda	106
4.4.1. Analiza ekvivalentnog kanala	107
4.4.2. Uloga zaštitnog koda i interlivera	108
4.4.3. Motiv za uvođenje zaštitnog koda i interlivera	111
4.5. Evaluacija sistema	112
4.5.1. Prikaz eksperimentalnih rezultata	112
4.5.2. Diskusija, kritike i poređenje sa postojećim rezultatima	118
4.5.3. Oblast primene predloženog rešenja	120
5. Zaključak	122
5.1. Sumarni ciljevi istraživanja	122
5.2. Ostvareni rezultati i doprinosi	122
5.3. Predlog daljeg rada	123
Dodatak A	126
Skraćenice	132
Bibliografija	133
Objavljeni radovi u naučnim časopisima i konferencijama	141

1. Uvod

1.1. Motivacija i opšta razmatranja

Savremene informacione tehnologije i globalizacija dramatično su promenili današnji svet računarskih mreža, koje se odlikuju visokim stepenom integracije različitih elektronskih servisa. Obzirom da je broj Internet servisa i novih korisnika usluga svakodnevno u porastu, povećavaju se i količina i vrednost informacija koje se razmenjuju. Informacije koje se razmenjuju putem mreža i skladište na umreženim memorijskim lokacijama, mogu biti kompromitovane ukoliko nisu adekvatno zaštićene. U svim tipovima mrežne komunikacije poseban izazov predstavljaju mehanizmi za kontrolu pristupa. Njihova uloga je da zaštite podatke i informacije od neovlašćenih pristupa putem dovoljno sigurnih postupaka za proveru identiteta.

Tradicionalni sistemi za proveru identiteta zasnovani su na nečemu što znamo - na primer lozinkama – ili na nečemu što posedujemo - *smart* karticama, tokenima, mobilnim telefonima... Ovi uređaji ne mogu garantovati da se radi o legitimnim korisnicima, jer ne postoji jaka veza između autentifikatora i servisa za autentifikaciju. Problem nastaje u slučaju njihovog kompromitovanja (krađa, gubljenje, kopiranje i sl.).

Problem je rešen uvođenjem još jednog atributa „nečeg što jesmo“. Biometrija ili bio-informacija se nameće kao moguće rešenje za obezbeđenje čvrste veze između autentifikatora i sistema za autentifikaciju. Biometrijski sistemi su bazirani na fizičkim ili bihejvioralnim karakteristikama ljudskih bića kao što je lice, glas, otisak prsta, iris i drugo. Biometrijski podaci imaju potencijalnu prednost da budu jedinstveni identifikatori jedne osobe. U zavisnosti od vrste biometrijskog izvora, mogu da sadrže dovoljnu ili nedovoljnu količinu informacija za određenu primenu. Za estimaciju maksimalne količine informacija neophodno je dobro poznavati osobine biometrijskih podataka okarakterisane prema zadatom tipu biometrijskog izvora, kao i tehnologije za precizno očitavanje i ekstraktovanje informacija. Obzirom na osobine koje poseduju biometrijski podaci biometrija je postala ozbiljan kandidat za zamenu tradicionalnih metoda na polju autentifikacije i autorizacije pojedinaca.

Kasnije je nastala ideja o generisanju kriptoloških ključeva na osnovu biometrijskih izvora. Hardverska i softverska rešenja za informacionu bezbednost, koja učestvuju na direktan ili indirektan način u skladištenju, obradi i prenosu podataka, moraju da koriste adekvatne kriptografske mehanizme koji obezbeđuju računarsku ili praktičnu sigurnost koja podrazumeva sigurnost u odnosu na protivnika koji ima specifične ograničene vremenske i računarske resurse. Snaga ovih kriptografskih mehanizama je određena snagom kriptološkog ključa. Danas mnoge zemlje razmatraju ili već koriste biometrijske podatke za podizanje sigurnosti u cilju zaštite pojedinca od krađe identiteta ili lažnog predstavljanja.

Aktuelnost i značaj ove doktorske disertacije, kao i potrebe za istraživanjem u ovoj oblasti, potvrđene su od Ministarstva za nauku i tehnološki razvoj Republike Srbije, sa izuzetno

visokom ocenom projekta TR-32054 pod nazivom, “*Digitalna obrada signala u sintezi sistema za zaštitu informacija*“, odobrenog u kategoriji projekata tehnološkog razvoja za period 2011-2014 na kom autor ove disertacije učestvuje kao istraživač, a mentor kao rukovodilac projekta.

1.2. Predmet istraživanja

Tehnološki razvoj Interneta i globalnih komunikacionih mreža stavio je u prvi plan pitanje sigurnosti i integritet prenesenih i memorisanih informacija. U istu kategoriju spada i pitanje pouzdanih metoda autentifikacije učesnika globalne mreže. Mehanizmi zaštite koji se baziraju na principu računarske složenosti potrebne za njihovo probijanje, sve više otvaraju prostor ka informaciono-teorijskim kriterijumima sigurnosti koji jasno definišu potrebne i dovoljne uslove za takozvanu bezuslovnu sigurnost. Sistemi sa ovim svojstvom su invarijantni na računarsku moć protivnika. Teorijski rezultati u poslednjih desetak godina ostvaruju solidnu osnovu za praktičnu realizaciju pojedinih segmenata bezbednosnih servisa, npr. tajnosti koji bi u dobroj meri posedovali ovo poželjno svojstvo. Predloženi rad u tezi u celini teži ka sticanju iskustva u projektovanju i realizaciji pojedinih gradivnih blokova ove nove klase sistema zaštite.

Polazna tačka istraživanja su teorijske osnove i Šenonov koncept savršenih šifarskih sistema. Teorijsko informaciona analiza biometrije, kao svojevrsnog informacionog izvora, davala bi konkretna rešenja za sintezu kriptografskog sistema sa teorijski garantovanim performansama. Performanse biometrijskih kripto sistema zasnivaju se na nominovanoj sigurnosti koja zahteva pažljiv i kompletan opis operacionog scenarija u kome se egzaktno proučava:

- Od kakvih napada kriptografski sistem treba da obezbedi zaštitu?
- Šta je poznato napadaču?
- Šta je poznato korisnicima sistema?
- Koje su fizičke pretpostavke koje bliže određuju šta napadač i korisnici mogu da urade?

Problem istraživanja predstavlja kompleksan način obrade biometrijskih podataka. Biometrijski podaci sadrže šum koji znatno povećava varijabilnost podataka i tako utiče na informacioni kvalitet jednog biometrijskog uzorka. Problemi ovakve vrste mogu da se prevaziđu upotrebom odgovarajućih tehnika za obradu signala i kodova za ispravljanje grešaka. Atraktivne osobine jedinstvenosti stavljaju biometriju na prvo mesto, ali sa druge strane pojavljuju se i brojne slabosti. Za razliku od lozinki, jednom izgubljena ili kompromitovana biometrija ne može ponovo da se koristi, niti zameni sa nekom drugom biometrijskom informacijom. Na osnovu ovih činjenica može se zaključiti da je biometrija ograničen resurs. Biometrijski sistemi moraju da odgovore na problem skladištenja i sigurne komunikacije između aplikacija u kojima se biometrija koristi što ujedno i čini osnovni problem istraživanja. Imajući u vidu navedene slabosti i zahtevane uslove korišćenja,

potrebno je analizirati različite scenarije generisanja kriptoloških ključeva iz biometrijskih uzoraka.

Generalna hipoteza - Postoje različiti biometrijski izvori koji sadrže jedinstvene informacije o ličnim karakteristikama jedne osobe. Iris kao biometrijski podatak sadrži dovoljnu količinu informacija za razvoj jedne klase sistema za generisanje kriptoloških ključeva, dovoljnih dužina za potrebe današnjih kriptografskih mehanizama, sa visokim stepenom nominovane sigurnosti.

Generalna hipoteza se operacionalizuje preko sledećih posebnih hipoteza:

- 1) Biometrijski podaci su jedinstveni i predstavljaju dobru polaznu osnovu za sintezu optimizovanog sistema za ekstrakciju kriptoloških ključeva visokog kvaliteta;
- 2) Uvođenjem informacionih mera, entropije, lokalne entropije i uzajamne informacije, moguće je identifikovati segmente u biometriji irisa koji su najpogodniji za proces generisanja kriptoloških ključeva;
- 3) Optimizacijom parametara odgovarajuće transformacione funkcije, očekuje se veća entropija i manje uzajamne informacije u transformacionom domenu, čime su postavljeni okviri za sintezu sistema za ekstrakciju čisto slučajnih nizova iz biometrije irisa;
- 4) Biometrija iz istih biometrijskih izvora se ne reprezentuje uvek na isti način zbog prisustva šuma u biometrijskim podacima, te je neophodno projektovati tolerantnu šemu za nastale greške;
- 5) Biometrijski izvor i biometrijske podatke moguće je posmatrati i okarakterisati kao komunikacioni kanal sa šumom, što uvodi u upotrebu široku paletu tehnika za obradu signala iz oblasti digitalnih komunikacija;

Istraživanje ima dva cilja: naučni, društveni i tehnički.

Naučni doprinos rada ogleda se u novoj naučno-relevantnoj informaciji koja se može okarakterisati kao rezultat potrebe i želje za definisanjem nove klase sistema za generisanje kriptografskih ključeva za primenu u savremenim bezuslovno sigurnim kriptografskim mehanizmima i predloga za dalji razvoj. Društveni cilj je primena rezultata istraživanja u svrhu podizanja bezbednosti pojedinaca u informacionom društvu preko prevencije od krađe identiteta zasnovanom na biometrijskim autentifikacionim tehnikama i da odgovori na tehnološke izazove sa kojima se susrećemo danas i sa kojima ćemo se susretati u budućnosti na polju informacione bezbednosti. Tehnički cilj obuhvata omogućavanje razvoja uređaja sa boljim performansama, kao i sa manjim greškama i većom društvenom prihvatljivošću.

1.3. Doprinosi disertacije

Rad je utemeljen na dosadašnjim dostignućima poznatih naučnih stvaralaca, kao i na sopstvenom radu na ovoj tezi.

Osnovni doprinos ovog rada na polju kriptografske bezbednosti čini sledeće:

1. Sprovođenje informacione analize nad biometrijom irisa u cilju sinteze optimizovanog sistema za ekstrakciju kriptoloških ključeva visokog kvaliteta;
2. Optimizacija parametara odgovarajuće transformacione funkcije u cilju dobijanja što veće entropije i što niže uzajamne informacije u transformacionom domenu;
3. Estimacija lokalne entropije i uzajamne informacije, identifikovanje segmenata irisa koji su najpogodniji za ove potrebe;
4. Postavljeni su okviri za sintezu sistema za ekstrakciju čisto slučajnih nizova iz iris biometrije, bez narušavanja njenih osvedočenih autentifikacionih svojstava;
5. Dizajnirana je nova generička šema sistema za generisanje kriptoloških ključeva. Naša šema kombinuje kod za ispravljanje grešaka i tehniku interlivinga;
6. Postignuta je izuzetno velika dužina ključa od 400 bitova, sa stopom lažnog prihvatanja od 0% i stopom lažnog odbacivanja od 4,75%. Generička šema čvrsto se oslanja na dobijene okvire nakon detaljno sprovedene teorijsko-informacione analize. Raniji pokušaji u ovoj oblasti su imali prilično neprihvatljive rezultate za sisteme ovakve namene ili su generisali relativno kratke ključeve koji su bili neupotrebljivi za aktuelne kriptografske algoritme;
7. Rad je jedinstven zbog svoje specifičnosti koja se ogleda u gore već navedenoj teorijsko-informacionoj analizi, preko koje smo usvojili zajedničku jedinicu mere (Šenonov bit) koja je osnova za sintezu svakog kriptosistema, što nije bio slučaj do sada;
8. Na osnovu mesta problema (između biometrije i kriptografije) definisali smo novu oblast pod nazivom „biometrijska kriptografija“ koja se bavi velikim problemom kao što je uspostavljanje veze između biometrije i kriptografije na takav način da je moguće zadovoljiti tačnost koja je zahtevana od strane kriptografije i pored varijabilnosti koja je prisutna kod biometrijskih podataka;
9. U tezi će biti izučen fenomen objedinjavanja biometrije i kriptologije, kao posledica zahteva iz obe discipline. U kriptologiji postoji trajni zahtev za generisanjem kvalitetnih kriptoloških ključeva koji se lako distribuiraju i memorišu, dok šira praktična upotreba biometrije povlači za sobom zahteve za bezbednim memorisanjem i opozivom biometrijskih podataka. Oba zahteva se ispunjavaju u sistemima za generisanje kriptoloških ključeva na osnovu biometrijskih podataka;
10. Predviđen je razvoj i implementacija nove klase algoritama za formiranje biometrijskih obrazaca zasnovanih na irisu imajući u vidu pre svega primenu u generisanju odgovarajućih kriptoloških ključeva.

U ovoj tezi je predstavljeno praktično realizovano sopstveno rešenje za generisanje kriptoloških ključeva na osnovu biometrijskih podataka sa sopstvenim razvijenim algoritmima. Algoritmi koji su razvijeni u okviru ove doktorske disertacije obezbediće:

- testiranje matematičkih modela za obradu biometrijskih podataka;
- opštu informacionu, kao i specijalnu analizu iris biometrije;

- evaluaciju performansi novog predloženog rešenja;
- analizu mogućih napada na sistem i stepen ranjivosti sistema;
- razvoj softvera u Matlab-u sa potpunom implementacijom projektovanog rešenja za generisanje kriptoloških ključeva na osnovu biometrijskih podataka.

1.4. Struktura disertacije

Proces naučnog istraživanja je sproveden u nekoliko koraka:

- 1) Uvod;
- 2) Pregled u oblasti istraživanja;
- 3) Teorijske osnove istraživanja;
- 4) Novi sistem za generisanje ključeva na osnovu biometrije irisa;
- 5) Zaključak sa sumarnim rezultatima i predlog za budući rad.

U nastavku teze u njenom drugom delu, razmatrano je opšte stanje u oblasti istraživanja. Prvo su uočeni nedostaci tradicionalnih biometrijskih sistema. Analizirani su koraci koji su doveli do razvoja jedne potpuno nove oblasti pod nazivom „biometrijska kriptografija“. Uvidom u mnoge naučne radove, navedene su prepoznate prednosti biometrijske kriptografije, kao i njen značaj na polju zaštite privatnosti biometrijskih podataka. Razvoj ove oblasti daje potpuno novu upotrebnu vrednost biometrijskim podacima. Osim postignutih rezultata na polju zaštite privatnosti, nastaje potpuno nova ideja o mogućnosti ekstrakcije kriptoloških ključeva na osnovu biometrijskih podataka. Razmatraju se novi principi i pristupi za potrebe dizajniranja biometrijskih kripto sistema za razliku od onih koji postoje kod tradicionalnih biometrijskih sistema u okviru servisa za autentifikaciju, pre svega u pogledu odvojene optimizacije pojedinih podsistema. Na ovaj način je omogućeno nezavisno podešavanje procesa za generisanje kriptološkog ključa prema principu maksimalne entropije ključa, ali bez narušavanja performansi podsistema za autentifikaciju. Na kraju su date moguće primene ove tehnologije koja se zasniva u potpunosti na sistemima za biometrijsku kriptografiju.

U trećem delu rada obrazložene su teorijske osnove istraživanja. Uveden je i objašnjen pojam neodređenosti i informacije. Date su osnove za sve Šenonove informacione mere, kao i grafička reprezentacija teorijskih-informaciono mera. U nastavku ovog dela analiziran je iris kao biometrijski informacioni izvor. Objašnjene su sve faze generisanja iris koda, polazeći od faze segmentacije preko normalizacije do faze kodovanja. Date su osnovne procene jedinstvenosti biometrije irisa na osnovu statističkih pokazatelja. Na kraju ovog dela rada razmatrane su osnovne tehnike iz teorije komunikacija. Pod ovim tehnikama podrazumevamo različite klase zaštitnih kodova (kodovi za ispravljanje grešaka) i interlivere.

U četvrtom delu rada otpočeli smo sa sintezom sistema za generisanje ključeva na osnovu biometrije irisa. Napravljena je generička šema novopredloženog rešenja. Primarno

originalno rešenje koje će biti obrađeno, odnosi se na razvoj sopstvenog algoritma za generisanje kriptoloških ključeva na osnovu biometrije irisa. Prikazan je izbor eksperimentalnog okruženja zajedno sa testnim uzorcima iz standardizovanih baza podataka. Prvi korak je predstavljen sa rigoroznom teorijskom-informaciono analizom iris kao biometrijskog izvora. Računata je entropija teksture slike irisa posle faze normalizacije, a zatim i entropija iris koda. Primenjeni su Šenonovi modeli aproksimacije za modelovanje irisa kao informacionog izvora. Određene su uzajamne informacije između istih i različitih irisa, kao i kompleksnost iris koda sa aspekta jedinstvenosti. U ovom koraku pošli smo u potragu za ekvivalentnim komunikacionim kanalom. Napravljen je izbor adekvatnog zaštitnog koda i interlivera. Objašnjen je motiv za uvođenje zaštitnih kodova i interlivera na osnovu pronalaska ekvivalentnog komunikacionog kanala. U poslednjem koraku urađena je evaluacija kompletnog sistema. Prikazani su svi eksperimentalni rezultati. Vođena je kritička diskusija na osnovu poređenja sa postojećim rezultatima od sličnih sistema. Na kraju je data oblast primene predloženog rešenja.

U poslednjem, petom delu rada dat je zaključak. Sumirani su svi ciljevi koji su postavljeni na početku istraživanja. Dati su svi ostvareni rezultati i doprinosi u radu. Razmatrane su moguće oblasti u kojima se rešenje može primeniti. Kao osnovne oblasti primene identifikovani su skoro svi kripto sistemi, prvenstveno oni sistemi koji su od izuzetne važnosti za informacionu bezbednost jedne države.

2. Pregled u oblasti istraživanja

Konvencionalna kriptografija za šifrovanje koristi kriptološke ključeve. Ključ predstavlja binarnu sekvencu tačno određene dužine, obično 128, 192, 256 bitova ili više. Ovi ključevi mogu biti simetrični, javni ili privatni, i predstavljaju suštinski deo svakog kriptografskog sistema. Na primer, infrastruktura sa javnim ključevima - PKI. Ključevi koji se koriste u javnoj kriptografiji su velike dužine, obično 1024 ili 2048 bitova. Ljudi ne mogu da pamte ovako duge ključeve, kao što je to slučaj sa tradicionalnim lozinkama. Ovakvi ključevi često su zaštićeni lozinkama ili PIN-om koji se pamti. Upravljanje lozinkama je najslabija tačka svakog kripto sistema iz razloga što lozinka može biti ukradena, izgubljena ili pogođena primenom metoda za potpunu pretragu (eng. *Brute force*).

Biometrijski sistem za prepoznavanje uvek proizvodi jedan odgovor „DA“ ili „NE“, a to je samo jedan bit informacije. Dakle, očigledna je uloga biometrije u konvencionalnim krypto sistemima, a to je uloga upravljanja lozinkama. U procesu prepoznavanja, nakon faze verifikacije ukoliko je odgovor „DA“, sistem otključava lozinku ili ključ. Ključ mora da se čuva na sigurnom mestu ili uređaju od poverenja.

Interakcija između biometrije i kriptografije predstavlja dve potencijalne komplementarne tehnologije. Postoje različiti biometrijski izvori koji sadrže jedinstvene informacije u vidu ličnih karakteristika jedne osobe. Na ovoj pretpostavci može biti izgrađen napredni sistem za autentifikaciju pojedinca sa visokim stepenom sigurnosti. Na primer, čvrstom kombinacijom biometrije i kriptografije moguće je potencijalno ostvariti jaku vezu između digitalnog potpisa i osobe koja ga je sačinila sa visokim stepenom sigurnosti.

Međutim, postoje neizostavni tehnološki izazovi na koje je potrebno odgovoriti, a to su tačnost, pouzdanost, sigurnost podataka, prihvatljivost, cena, interoperabilnost, kao i izazovi vezani za obezbeđenje efikasne zaštite privatnosti. Sve ovo predstavlja uobičajenu ranjivost odnosno nedostatak biometrijskih sistema sa aspekta društvene prihvatljivosti. Biometrijske tehnologije stavljene u kontekst kriptografskih mehanizama, predstavljaju grupu povezanih tehnologija preko kojih je moguće na potpuno siguran način generisati kriptološki ključ na osnovu biometrijskih podataka.

Autentifikacija ili provera legitimnosti korisnika je sastavni deo opšte bezbednosti svih informacionih sistema. Tradicionalno, autorizacija korisnika znači uvođenje korisničkog imena i lozinke, a to je tehnika koja je u upotrebi već decenijama. Napravljene su mnoge inkrementalne promene na ovom polju, a izdvajamo samo osnovne. Lozinke se više ne šalju u otvorenoj formi preko mreže i postavljen je zahtev se generisanje jačih lozinki, ali se i dalje, osnovni pristup, u suštini, nije promenio. Slabosti ovog pristupa su dobro poznate i zbog njih su prisutni problemi na svakodnevnom nivou.

Za većinu poslovnih okruženja potencijalne pretnje uključuju:

- napad na privatnost, u kojem pojedinci neovlašćeno pristupaju tuđim podacima;
- napad na poslovne tajne, način na koji pojedinci kompromituju neku organizaciju, tako što joj nanose ogromne finansijske gubitke;
- napad sa ciljem zloupotrebe, u kojima pojedinac neovlašćeno raspolaže sa tuđim podacima, tj. lažno se predstavlja (krađa identiteta);

Sa druge strane, kriptografski mehanizmi koji obezbeđuju servis poverljivosti, integriteta i neporecivosti zahtevaju kriptološke ključeve većih dužina zbog kontinualnog tehnološkog razvoja. Drugi problemi nastaju usled odabira sigurnog mesta za čuvanje ključeva i načina za distribuciju. Ovakve ključeve nije moguće zapamtiti i neophodno je koristiti različite hardverske tokene za skladištenje kriptoloških ključeva. Takođe, slabosti ovog pristupa za skladištenje ključeva su dobro poznate, ukoliko hardverski medijum bude kompromitovan na neki od ovih načina.

Biometrija je ponudila nove i bolje pristupe za autentifikaciju korisnika, kao i tehnike za obezbeđenje mnogih servisa savremene informacione bezbednosti. Obzirom da su tradicionalne lozinke najslabije karike u sistemu bezbednosti, biometrija predstavlja novu potencijalnu tehnologiju koja treba da bude brzo usvojena u cilju rešavanja problema na svim nivoima bezbednosti.

Autentifikacija na bazi biometrije obezbeđuje pogodnu i sigurnu proveru identiteta korisnika kako u lokalnom mrežnom okruženju, tako i preko Veba. Ukoliko je zahtevan viši nivo bezbednosti, na mreži je moguće koristiti biometriju samostalno ili u kombinaciji sa drugim akreditivima. Biometrijske sisteme je značajno teže prevariti nego dosadašnje tradicionalne sisteme na bazi lozinki. Imajući u vidu raznolikost biometrijski izvora, moguće je izvršiti njihovu podelu prema kvalitetu, u odnosu na količinu informacija koje sadrže. U zavisnosti od količine informacija, moguće je generisati biometrijske ključeve za kriptografske potrebe. Osim primene za generisanje ključeva, biometrija nudi rešenja za čuvanje i skladištenje kriptoloških ključeva, kao i njihovu distribuciju.

2.1. Nedostaci biometrijskih sistema

Razvojem biometrijskih sistema i njihovom prvom tradicionalnom primenom, prepoznati su i prvi nedostaci. Nedostaci mogu biti različito okarakterisani i možemo da ih podelimo u dve osnovne kategorije. Prvu kategoriju čine svi nedostaci koji se odnose na lošu zaštitu privatnosti biometrijskih podataka, a koji su u direktnoj vezi sa osnovnim osobinama biometrijskih podataka (jedinственost, nezamenljivost, sadrže šum). Drugu kategoriju čine nedostaci koji su vezani za bezbednost sistema, kod kojih se biometrijski podaci koriste u svrhu autentifikacije.

Tradicionalni model primene biometrije irisa je praktično neprihvatljiv kada je u pitanju zaštita privatnosti i nije u skladu sa drugim aspektima bezbednosti. Iako biometrijski obrasci mogu biti šifrovani, to nije dovoljno dobro rešenje za bezbednost biometrijskih podataka. Jer,

ukoliko se dogodi da baza podataka bude kompromitovana, biometrijski podaci će biti izgubljeni zauvek. Sa druge strane, biometrijski obrasci se ne nalaze uvek u šifrovanoj formi, zbog nemogućnosti obavljanja faze prepoznavanja u kojoj se vrši provera stepena poklapanja dva biometrijska obrasca. U momentu provere poklapanja, biometrijski podaci moraju da budu prisutni u dešifrovanoj formi.

Ukoliko je napad u ovoj fazi uspešno realizovan, napadač će posedovati odgovarajuću originalnu biometriju. Ponovljenim napadom, pristup sistemu je zagarantovan. Nastala situacija može da postane još teža, ukoliko je vlasnik izgubljene biometrije koristio biometriju za pristup u više različitih i nepovezanih servisa. Ovo predstavlja ozbiljan bezbednosni problem, a takođe može biti problem koji nosi određene pravne konsekvence.

Između biometrijskih podataka postoji velika varijabilnost, nezavisno od tipa biometrijskog izvora. Varijabilnost biometrijskih podataka nastaje skoro u svim fazama obrade biometrijskih podataka. Od momenta očitavanja, pa sve do faze za generisanje biometrijskih obrazaca. Biometrijski podaci poseduju šum zbog kojeg nije moguće dva puta očitati potpuno identičan biometrijski podatak, sa istog biometrijskog izvora. Na primer, tehnika za prepoznavanje identiteta korisnika na osnovu biometrije lica sklona je visokom stepenu varijabilnosti zbog različitih spoljnih uticaja - osvetljenja i ugla pomeranja korisnika. Ovo je jedan od razloga zbog koga smo zamoljeni da ne koristimo sliku sa osmehom u ličnim dokumentima. Slično ovome, brojni faktori utiču na neefikasno izdvajanje konzistentne informacije i iz uzoraka biometrije prsta. Na osnovu analize osobina za različite tipove biometrija, utvrđeno je da iris oka sadrži najveću količinu informacija i najmanju varijabilnost, a sa druge strane odlikuje ga i visok stepen pouzdanosti i konzistentnosti.

Kao posledica varijabilnosti, sveže generisani biometrijski obrasci mogu biti u potpunoj suprotnosti sa uskladištenim referentnim biometrijskim obrascima. Kod biometrijskih sistema nije moguće postići tačnost poklapanja dve biometrije od 100%. Kada biometrijski sistem nije u stanju da potvrdi visok stepen sličnosti u fazi prepoznavanja, prinuđen je da odbaci legitimnog korisnika. Ovakav događaj naziva se „lažno odbacivanje“ (eng. *False Rejected Rate*) od strane sistema u fazi prepoznavanja. Ukoliko se dogodi ovakav slučaj, korisnik je prinuđen da ponovi postupak iz početka.

Biometrijski sistemi za prepoznavanje mogu biti dizajnirani tako da se stopa lažnog odbacivanja smanji i da prepoznavanje bude uspešno, što će omogućiti pravilno funkcionisanje sistema i obezbediti njegovu društvenu prihvatljivost. Standardne metode za poboljšanje rada sistema za prepoznavanje su dodatno kontrolisanje uslova u kojima se očitavaju izvori biometrijskih podataka, kao i optimizacija matematičkih algoritama. Jedan od načina da se to uradi, jeste da se smanji prag ili granica za prihvatanje - FRR. Postoje problemi sa ovim pristupom jer direktno utičemo na stopu „lažnog prihvatanja“ (eng. *False Acceptance Rate*), a taj događaj bi značio da je sistem napravio grešku kod identifikacije i prihvatio pogrešnu osobu na bazi tuđeg referentnog biometrijskog obrasca.

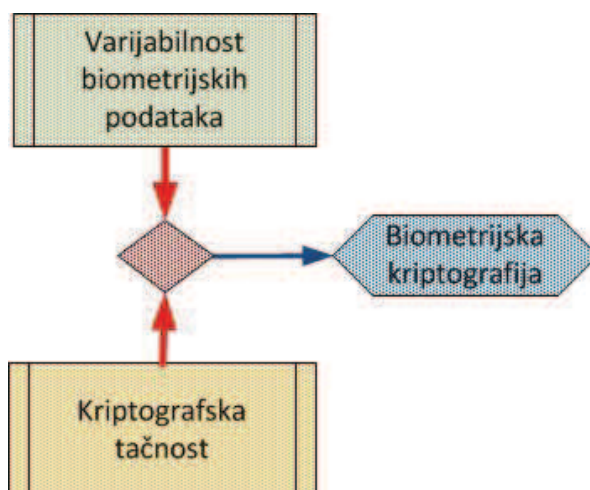
Imajući u vidu da je moguće smanjenje jednog parametra na račun drugog i obrnuto, neophodno je napraviti kompromis između njih, tako da sistem bude prilagođen okruženju. Neka okruženja dozvoljavaju visok FRR, dok druga mogu tolerisati visok FAR.

Vrednosti parametara FAR i FRR koje su navedene od strane prodavca sistema za prepoznavanje, na osnovu biometrijskih podataka, često su nepouzidane. Iz tog razloga, neophodno je konsultovati nezavisne izvore informacija, kao što su javna takmičenja na polju biometrijskih istraživanja u organizaciji američkog Nacionalnog instituta za standarde (NIST). Za većinu biometrijskih sistema, vrednost FRR kreće od 0.1 % do 20 %, što znači da će legitimni korisnik biti odbijen najmanje jednom od 1000 puta. FAR kreće od jednog prihvaćenog u 100 (aplikacije niskog stepena sigurnosti) i jednog prihvaćenog u 10.000.000 (aplikacije visokog stepena sigurnosti). Drugi izazov kod biometrijskih sistema je brzina prepoznavanja. Sistem mora da bude sposoban da napravi preciznu odluku u realnom vremenu. Sa druge strane, najbitnija sposobnost biometrijskih sistema koju bi trebalo da poseduju je otpornost na učestale napade.

2.2. Biometrijska kriptografija

Originalni koncept „biometrijske kriptografije“ je primenjen nad otiscima prstiju 1994. god. Pionir u ovoj oblasti bio je dr. Džordž Tomko (*dr. George Tomko*), osnivač (*Mytec technologies*) Toronto, Kanada. Od tada mnoge istraživačke grupe su učestvovala u razvoju ove oblasti kao i drugih srodnih tehnologija. Pored naziva biometrijska kriptografija (eng. *Biometric Encryption*), korišćeni su i drugi nazivi za ovu tehnologiju, kao što su: biometrijski kripto sistemi (eng. *biometric cryptosystem*), privatni obrasci (eng. *private template*), biometrijsko zaključavanje (eng. *biometric locking*), sigurna skica (eng. *secure sketch*), biometrijsko vezivanje ključeva (eng. *fuzzy commitment scheme, fuzzy vault*), biometrijsko generisanje ključeva (eng. *biometric key generation*), virtuelni PIN (eng. *virtual PIN*), biometrijski potpis (eng. *biometric signature*) i biometrijsko hešovanje (eng. *bioHashing*).

Biometrijska kriptografija može da se definiše, u osnovi, kao proces koji koristi PIN, lozinku ili kriptološki ključ u kombinaciji sa biometrijskim podacima, na takav način da originalni biometrijski podaci, niti ključ ne mogu biti otkriveni ili regenerisani na osnovu referentnog podatka koji je predviđen za skladištenje. Ovo bi značilo da je moguće sa posebno dizajniranim algoritmima ponoviti ili regenerisati originalni ključ sa novim biometrijskim obrascem bez skladištenja prvog generisanog biometrijskog obrasca u početnoj fazi generisanja biometrijskog ključa. Dva biometrijska obrasca ne moraju da budu identična. Važno je da oba obrasca potiču od istog biometrijskog izvora da bi bilo moguće regenerisati ključ. Korišćenjem ove tehnike moguće je napraviti određen kompromis između biometrijske varijabilnosti i zahtevane kriptografske preciznosti (**Slika 1**).



Slika 1. Izvor biometrijske kriptografije

Biometrijska kriptografija i srodne tehnologije su stvorile interesovanje glavnih akademskih istraživačkih centara specijalizovanih za biometriju poput državnog Univerziteta Mičigen, Univerziteta u Virdžiniji, Univerziteta Karnegi Melon u Kembridžu (Velika Britanija) i Univerziteta u Bolonji (Italija). Među sadašnjim liderima u industriji nalaze se IBM (T.J. Watson Research Center), (RSA Laboratories), (Lucent Technologies), (Sandia National Laboratories) i (Philips Research).

2.3. Prednosti i značaj biometrijske kriptografije

Primenom tehnologije za biometrijsku kriptografiju postignut je ogroman potencijal za poboljšanje privatnosti i informacione bezbednosti. Ključne prednosti, kao i značaj ove tehnologije su navedene u nastavku.

- a) Kod biometrijskih sistema strogo je zabranjeno čuvanje biometrijskih obrazaca u bazama podataka. Ovo je regulisano zakonima za zaštitu privatnosti kod upotrebe biometrijskih sistema. Zahtevana je minimizacija podataka, preko koje se smanjuje količina ličnih podataka. Ovo je postignuto generisanjem nezavisnih referentnih podataka koji ne odaju originalne biometrijske podatke. Bezbednosni problemi više su prisutni na strani baza podataka za skladištenje biometrije, zbog straha od zloupotrebe (krađa, presretanje, zamene i drugo).
- b) Biometrijska kriptografija omogućava korisnicima višestruku upotrebu biometrijskih obrazaca. Ukoliko je jedan nalog kompromitovan, značajno su smanjene šanse za kompromitovanje drugih naloga za pristup drugim servisima. Sa druge strane, obezbeđena je mogućnost ponovnog generisanja na osnovu originalne biometrije, kao i povlačenje referentnih podataka iz upotrebe. Tradicionalni biometrijski sistemi ne poseduju navedene mogućnosti.
- c) Povećana je bezbednost kod servisa autentifikacije, tako što je ostvarena čvrsta veza između autentifikatora (korisnika) i korisničkog naloga za pristup

određenom servisu. Obezbeđena je zaštita lozinki i ključeva kreiranjem referentnih podataka. U momentu predstavljanja lozinke ili ključa, zahtevana je originalna biometrija i referentni podatak. Napadač nema mogućnost da dobije ključ na osnovu referentnog podatka, kao ni informaciju o biometriji korisnika servisa.

- d) Biometrijska kriptografija povećava bezbednost ličnih podataka i komunikacija. Korisnici sistema za biometrijsku kriptografiju mogu da iskoriste pogodnost i jednostavnost date tehnologije za kriptografsku zaštitu ličnih podataka. Obzirom da je priroda izvora ključa lična i na strani pojedinca, ova činjenica predstavlja moćan alat za široku primenu kriptografske zaštite.
- e) Biometrijskom kriptografijom se stiče javno poverenje koje je neophodno za uspeh svakog biometrijskog sistema. Politika upravljanja podacima i procedure mogu da idu daleko onoliko koliko se neguje poverenje u javnosti. Biometrijski sistem mora da obezbedi elemente privatnosti, poverenja i bezbednosti. Ovo su ključni elementi koji utiču na njegovu prihvatljivost. To znači da biometrijski podaci moraju da budu stavljeni pod isključivu kontrolom pojedinca, kao i da je obezbeđen visok nivo zaštite privatnosti i podataka.

2.3. Razvoj oblasti biometrijske kriptografije

Praktično sve vrste biometrija (otisak prsta, iris, lice, dinamika kucanja, glas, svojeručni potpis i druge) su testirane preko tehnike za generisanje ključa, na osnovu biometrije (zaključavanje ključa biometrijskim obrascem) ili tehnikom koja obezbeđuje generisanje ključa iz biometrijskih podataka (primena ekstraktorskih funkcija). Najbolji rezultat postigli su autori [1], takoreći najperspektivniji rezultat postignut je sa iris biimetrijom sa FRR parametrom od 0,47% i FAR parametrom od 0% za generisanje kriptološkog ključa dužine 140 bitova. Ove greške su neznatno veće kod konvencionalnih biometrijskih sistema za prepoznavanje. Upotreba otiska prsta je takođe moguća u pogledu tačnosti za biometrijsku kriptografiju sa FRR parametrom većim od 10%, ali za razliku od irisa postoji primetna degradacija tačnosti biometrijskog sistema za otiske prstiju.

Treća biometrija, posle irisa i otiska prsta, jeste lice. Poslednjih godina napravljena su primetna poboljšanja kod biometrijskih sistema za prepoznavanje, na osnovu lica sa aspekta performansi sistema. Očekivani FRR parametar je 3,5%, parametar FAR je 0. Ovim sistemom moguće je generisati ključ dužine 56 bitova, što su pokazali autori [2]. Postignut je dovoljno dobar rezultat koji omogućava izbacivanje tradicionalnih lozinki iz upotrebe.

Nije poznato da li drugi biometrijski izvori poseduju dovoljnu količinu informacija za generisanje ključa minimalne dužine od 128 bitova. U ovom slučaju, entropija je mera koja se koristi za postavljanje teorijskih okvira za dužinu kriptoloških ključeva. Entropiju možemo definisati kao neredundantnu informaciju koja je sadržana u biometrijskim izvorima. Ova mera je od velike važnosti jer su na njoj zasnovani kriptografski principi koji garantuju nivo bezbednosti kriptografskih mehanizama. Teorijski sada možemo da kažemo da dužina ključa

ne sme da prelazi ukupnu količinu informacija koju poseduje biometrijski izvor, a predstavljeno je preko entropije, gde se za jedinicu informacije koristi 1 bit.

Radovi koji su objavljivani od 2001. godine daju opštu teorijsku osnovu za tehnologije biometrijske kriptografije, sa kriptografske tačke gledišta. U njima se dokazuje da sistem može biti tako dizajniran da bude otporan na kriptoanalitičke metode, čak i na metod potpune pretrage. U ovom slučaju napadač mora da isproba sve moguće kombinacije, kako bi uspeo da otkrije ključ ili biometrijski obrazac, čime bi sistem bio bezbedan u odnosu na dostupnu tehnologiju. U ovom dokazu se pretpostavlja, isto kao kod konvencionalne kriptografije, da je algoritam javan, a da napadač kod sebe može posedovati biometrijski obrazac koji nije odgovarajući sa nijednim referentnim podatkom u bazi podataka.

Međutim, napadač bi mogao da pokuša da izvede sofisticiranije napade, tako što bi iskoristio nasledne slabosti. Nasledne slabosti možemo tumačiti kao uzajamnu informaciju ili zavisnost koja postoji između dva biometrijska podataka ili između dva referentna biometrijska podatka koja u sebi sadrže zaključan ključ i originalni biometrijski obrazac. Istraživanje na polju informacione analize je u velikoj meri zanemareno. Ako bi ovakav napad bio uspešan, efikasna sigurnost sistema sa 140 bitova recimo kod autora [1], bila bi smanjena na 70 ili 50 bitova. U cilju određivanja efikasnije sigurnosti sistema, smatramo da je neophodno sprovesti detaljnu informacionu analizu nad podacima koji učestvuju u šemi za tehnologije biometrijske kriptografije.

2.3.1. Zaštita privatnosti biometrijskih podataka

Navedeni bezbednosni problemi ne bi postojali, ukoliko bi se više vodilo računa o zaštiti privatnosti pojedinaca. Zaštita privatnosti podrazumeva da se stvarna biometrija ne čuva kao takva u bazama podataka. Generalno je poželjno, po mogućnosti, izbeći bilo kakvo čuvanje biometrijskih podataka u bazama podataka. U kontekstu sa prethodnim, biometrijski obrazac nekog pojedinca bi bio sačuvan (memorisan) kao rezultat neke neinvertibilne funkcije. Na ovaj način originalna biometrija u formi slike ili digitalnog biometrijskog obrasca ne bi bila sačuvana u bazi podataka ili na drugim medijima za čuvanje podataka. Opisanom tehnikom o kojoj je reč u nastavku rada, obezbeđena je mogućnost za generisanje više biometrijskih obrazaca na osnovu jednog originalnog obrasca, a da pri tom nisu međusobno korelisani (sadrže uzajamnu informaciju). Generisani obrasci mogu biti sačuvani na istim ili različitim mestima za skladištenje, a da pripadaju samo jednom pojedincu.

Primenom šeme za zaštitu privatnosti, obezbeđuje se mogućnost za opoziv izdate biometrije "eng. *Cancelable Iris Biometric*". Na ovaj način, originalna biometrija neće biti sačuvana na klijentskim računarima, u bazama podataka ili na drugim mestima slične namene. Posle neophodne obrade biometrijskih podataka (proces segmentacije, normalizacije i kodovanja), originalni biometrijski obrazac biće uništen nakon primene transformacione neinvertibilne funkcije. Rezultat ovakve funkcije transformacije, predstavlja novu verziju biometrijskog obrasca (referentni biometrijski podatak) koji je potpuno bezbedan za korišćenje u ovim sistemima, jer ne postoji mogućnost za otkrivanje originalnog biometrijskog obrasca. Naziv ove šeme za primenu biometrije, sada je potpuno opravdan jer vrlo lako je moguće generisati

novi, potpuno drugačiji obrazac promenom samo inicijalnih parametara date funkcije transformacije. Na ovaj način, stekle su se mogućnosti za generisanje na hiljade različitih referentnih biometrijskih obrazaca na osnovu jednog biometrijskog izvora od istog pojedinca. U nastavku ovog poglavlja, dat je pregled na osnovu mnogobrojnih radova u kojima su predložene slične tehnike za zaštitu privatnosti biometrijskih podataka.

Autori [3] su analizirali tehnike za zaštitu privatnosti biometrije irisa i tehnike za povećanje stepena tačnosti algoritama za prepoznavanje ili poređenje dva obrasca irisa. U radu su korišćene jednosmerne funkcije (eng. *One-way function*) za zaštitu originalnih biometrijskih podataka i drugih autorizacionih informacija. Autori ovog rada polaze od pretpostavke da je za razvoj sigurnih aplikacija i sistema neophodno predvideti u dizajnu i implementirati siguran servis za korisničku autentifikaciju. Prvo je razmatrana „*Of-line*“, a kasnije „*On-line*“ korisnička autentifikacija zasnovana na biometrijskim sistemima, u kojima se kao algoritam za proveru prepoznavanja identiteta koristi Hamingovo rastojanje (eng. *Hamming distance*). Prvi model koji je predstavljen od autora, [3 str. 3] koristi se za povećanje tačnosti algoritama za identifikaciju i autorizaciju u sigurnim aplikacijama, tako što primenjuje tehniku čvrstog vezivanja (eng. *Binding data*) biometrijskog obrasca sa autorizacionim podacima i tako novodobijeni podatak smešta na *smart* kartice. Druga tehnika koju su predložili autori [3 str. 7] je specijalno dizajnirana i koristi se za zaštitu privatnosti biometrijskih podataka. Biometrijski podaci su enkapsulirani u autorizacionim informacijama i ne zahtevaju sigurne hardverske *tokene*. Kompletno istraživanje je sprovedeno u cilju olakšanja praktične primene biometrijskih sistema u servisima sigurnih aplikacija. U cilju obezbeđenja zaštite privatnosti biometrijskih podataka, predloženo je sledeće rešenje [3 str. 7, 8]. Administrator sistema korisniku servisa generiše javni i privatni ključ koji će koristiti za autentifikaciju. Javni ključ je smešten u biometrijskom čitaču. Na *smart* kartici se nalaze informacije, kao što su ime osobe, osnovne informacije o izdavaču ATM-a, kao i digitalni potpis čvrsto vezanih podataka sa slučajno generisanom binarnom sekvencom. Za proces autentifikacije neophodno je da osoba priloži *smart* kartica, nakon toga biometrijski obrazac se generiše nezavisno od korisnika. U sledećem koraku se konstrukcijom kodova za ispravljanje grešaka (eng. *Error-Correcting-Code*) dobija ispravna biometrija i u poslednjem koraku se vrši verifikacija. Digitalni potpis je verifikovan javnim ključem administratora sistema. Autori u radu na ovaj način rešavaju pitanje privatnosti.

Autori [4] predlažu novu šemu, poznatu pod nazivom „*Fuzzy commitment scheme*“. Predložena šema povezuje dve tehnike, tehniku za prikrivanje (eng. *Concealing*) i tehniku za ostvarivanje čvrste veze između podataka (eng. *Binding*). U poređenju sa konvencionalnom kriptografijom za otključavanje ovog nastalog tipa podatka, neophodan je jedinstven identifikator, slično kao kriptološki ključ koji se koristi za dešifrovanje. Karakteristike predložene šeme su vrlo primenjive kod aplikacija kao što su biometrijski servisi za autentifikaciju, čiji podaci zadrže šum sa karakterističnim osobinama. Predložena šema [4 str. 4, 5] je tolerantna na nastale greške. Sposobna da zaštiti biometrijske podatke na sličan način kao i konvencionalna kriptografija, ili kao jednosmerne funkcije koje se koriste za zaštitu alfanumeričkih lozinki. O navedenoj šemi više će biti reči u delu rada pod brojem 2.3.3.

Autori [5] su analizirali prednosti biometrije sa aspekta novopredloženog rešenja kod sistema za kontrolu pristupa, u cilju obezbeđenja naprednih sistema za autentifikaciju. Identifikovane su slabosti kod tradicionalnih biometrijskih sistema i predstavljeno je novo rešenje kojim bi se eliminisale neke od ovih slabosti. Za ilustraciju, autori u radu koriste otisak prsta kao biometrijski podatak. Analiza je obuhvatila i druge biometrijske metode. Osnovni problem tradicionalnih biometrijskih sistema je nemogućnost zamene originalne biometrije sa novom, potreba koja se javlja ukoliko dođe do kompromitovanja biometrijskih podataka. U cilju ublažavanja ovog problema, autori [5 str. 629] predlažu šemu pod nazivom (eng. *Cancelable biometrics*). Šema se sastoji od namerne, ali ponovljive distorzije (izobličenja) biometrijskog signala (podatka) uz pomoć izabrane transformacije. Na ovaj način, biometrijski signal je izobličen u svim sferama primene, za upis u sistem i za svaku proveru identiteta. Sa ovim pristupom, svaka nova instanca može koristiti različitu transformaciju čineći kros-podudaranje nemogućim. Osim toga, ako jedna instanca biometrijskog obrasca bude kompromitovana, vrlo lako se može kreirati nova sa transformacionom funkcijom, slično kao kada imamo potpuno novog pojedinca. Za proces distorzije, izabrana je jedna neinvertibilna funkcija. Na ovaj način, čak i kada je napadaču poznat dizajn algoritma transformacione funkcije i transformisani biometrijski podatak, napadač nije u mogućnosti da rekonstruiše originalni biometrijski podatak. Metod distorzije, kao transformacione funkcije je moguće primeniti u domenu bilo kog signala. Biometrijski signal se može direktno transformisati posle akvizicije ili kao i obično, posle procesa generisanja iris obrasca. Moguće je uticati na povećanje dužine biometrijskog obrasca, kao i na bitovsku snagu sistema (eng. *Entropy*). U idealnom slučaju neinvertibilna funkcija treba da bude takva, da čak i sa poznavanjem originalne biometrije nije moguće povratiti jednu ili više biometrijskih instanci. Primeri su dati od strane autora na [5 str. 629-630]. U prvom primeru za biometriju lica, (eng. *Face biometric*) korišćena je distorzija nad originalnom slikom, nakon čega se distorzovana slika nije podudarala sa originalnom. U drugom primeru za biometriju prsta, korišćen je jedan od metoda za skremblovanje signala, čime je postignut identičan rezultat kao i u prethodnom primeru. Metod koji je korišćen za distorziju signala, opisan je u radu [6].

Pored navedenih neinvertibilnih funkcija, u transformacionom domenu mogu biti primenjene i šifarske funkcije (eng. *Encryption function*) za zaštitu privatnosti biometrijskih podataka. Tehnika koju predlažu autori u ovom radu drugačija je od šifarskih funkcija jer ukoliko budu kompromitovani neki parametri sistema, kao što je kriptološki ključ, biometrija je izgubljena zauvek, dok ova mogućnost ne postoji kod neinvertibilnih funkcija. Razlika je u tome što transformisani signal zadržava iste osobine koje su bitne za servis autentifikacije, dok te osobine nisu sadržane u šifrovanim biometrijskim podacima. Parametar distorzije u transformacionom domenu, moguće je kontrolisati od strane servisa ili još bolje od strane pojedinca koji se identifikuje. U cilju podizanja najvišeg stepena privatnosti, predloženo je čuvanje parametra transformacije (eng. *State*) na *smart* karticama koje su dodatno zaštićene sa PIN kodom. Ukoliko se desi da kartica bude izgubljena ili bude ukradena, taj događaj neće značajno uticati na povredu privatnosti pojedinca i celokupnu bezbednost sistema. Na ovaj način je ublažena zabrinutost za kršenje privatnosti, upotrebom sistema za autentifikaciju na bazi biometrijskih podataka. Detaljan opis predloženog rešenja dat je u radu [5 str. 632].

Autori [7] predlažu novi metod za zaštitu privatnosti kod biometrije irisa. Za ekstrakciju iris koda ili generisanje biometrijskog obrasca, koriste funkciju transformacije 1D log-Gabor filtr, slično implementaciji [8]. U prvoj fazi, koriste razmotanu sliku irisa kojoj dodaju eksternu, slučajno generisanu informaciju za proces distorzije signala. Transformaciju koju koriste je neinvertibilna i opoziva. Detaljan opis predloženog rešenja dat je u radu [7 str. 2].

Autori [9] u cilju prevazilaženja problema neopozivosti biometrijskih podataka tj. zaštite privatnosti biometrijskih podataka, predložili su šemu sa dva faktora zaštite koristeći biometriju irisa i lozinku. U ovoj šemi, svakom korisniku se određuje zaseban ključ za mešanje iris koda koji predstavlja jedinstvenu permutaciju. Drugi faktor ove šeme predstavlja uvođenje kodova za detekciju i ispravljanje grešaka (eng. *Error-Correcting-Code*) u cilju smanjenja varijabilnosti između biometrijskih podataka. Šema mešanja kod različitih irisa u fazi poređenja, povećava Hamingovo rastojanje ostavljajući originalno rastojanje nepromenjeno, dok ECC kodovi smanjuju Hamingovo rastojanje u fazi poređenja istih irisa u većoj meri, nego u fazi poređenja različitih irisa. Primenom šeme mešanja postignuti su bolji rezultati koji se manifestuju kroz Hamingovo rastojanje u poređenju istih i različitih irisa. Ključ koji se koristi za mešanje, zaštićen je lozinkom koja ovaj sistem čini biometrijski opozivim. Na ovaj način, biometrijski podaci su u zaštićenoj formi koja obezbeđuje privatnost biometrijskih podataka. Detaljan opis predloženog rešenja dat je u radu [9 pp. 123, 124].

Autori [10] predstavili su novu šemu za zaštitu biometrije irisa. Slično kao i u drugim radovima, predložena je neinvertibilna transformacija koja bi se primenila nad iris obrascem, čak i ako je transformacija poznata napadaču. Predloženi metod sadrži tri faze. U prvoj fazi generisan je iris obrazac, u drugoj fazi ekstraktovani su konzistentni bitovi irisa i u trećem koraku primenjen je algoritam koji pruža zaštitu privatnosti i opoziv iris biometrije. Detaljan opis predloženog rešenja, dat je u radu [10 str. 883].

Na osnovu uvida u referentne radove koji se bave pronalaženjem rešenja koje bi obezbedilo zaštitu privatnosti biometrijskih podataka (eng. *Cancelable Biometric*), možemo sumirati određene ideje i izvesti sledeće zaključke:

- 1) Većina autora predlaže jednosmerne funkcije ili transformacije. Funkcije treba da poseduju osobinu neinvertibilnosti preko koje je obezbeđena mogućnost opoziva biometrijskih podataka, kao i ponovno generisanje biometrijskih podataka u svrhu autentifikacije. Na ovaj način su zamenjeni tradicionalni biometrijski sistemi;
- 2) Pojedine transformacione funkcije zadržavaju u podacima osnovne karakteristike biometrijskih podataka koje pozitivno utiču na Hamingovo rastojanje, metod koji koristimo za prepoznavanje ili poređenje biometrijskih podataka;
- 3) Prepoznate su određene slabosti kod transformacija sa lozinkama koje predstavljaju jedinstvenu "*permutaciju*". Problem je poistovećen sa kompromitovanjem kriptološkog ključa u kriptografiji;
- 4) Pored primene transformacija, neki autori poistovećuju biometrijske podatke sa komunikacionim kanalima uz prisustvo šuma. Ovakva ideja uvodi primenu kodova za

ispravljanje grešaka (eng. *Error Correcting Code*), a metod je preuzet iz oblasti komunikacija;

- 5) Tehnika zaštitnog kodovanja ECC je omogućila praktičnu realizaciju ideje o dizajniranju sistema za autentifikaciju, kao i kod tehnika za stvaranje čvrste veze između nezavisnih podataka (eng. *Data binding*), sa ciljem maskiranja ili skrivanja biometrijskih podataka;
- 6) Sva istraživanja su sprovedena u cilju obezbeđenja privatnosti ili zaštite biometrijskih podataka kod primene u sistemima za autentifikaciju, a sa druge strane dovela su do stvaranja potpuno nove ideje o mogućnosti generisanja i distribucije biometrijskih ključeva na bazi biometrijskih podataka, o čemu će se kasnije više diskutovati i što predstavlja glavni predmet istraživanja u ovoj disertaciji;

2.3.2. Generisanje ključeva iz biometrijskih podataka

U ovom delu rada napravićemo pregled prvih tehnika za generisanje kriptoloških ključeva iz biometrijskih podataka. Tehnika obezbeđuje zaštitu privatnosti biometrijskih podataka, opoziv, kao i mogućnost ponovnog generisanja javnih biometrijskih podataka za primenu kod sistema za autentifikaciju, kriptografsku zaštitu podataka, digitalno potpisivanje i druge kriptografske aplikacije. Sigurnost kod ovih aplikacija obezbeđuju bešumni, slučajni, uniformni nizovi u ulozi kriptoloških ključeva.

Kriptografija se tradicionalno oslanja na ravnomerno raspoređene i precizno ponovljive slučajne nizove u ulozi tajne ili kriptološkog ključa. Realnost međutim otežava generisanje, čuvanje i pouzdano preuzimanje takvih slučajnih nizova. Uglavnom ti nizovi nisu sa uniformnom ili ravnomernom raspodelom, a još teže su precizno ponovljivi, kada se to zahteva. Na primer, nizovi koji proizlaze iz obrađene zenice oka ili irisa, nisu sa uniformnom distribucijom, niti se precizno očitavaju svaki put kada se iznova generiše obrazac irisa.

Da bi ilustrovali upotrebu slučajnih nizova za autentifikaciju, na jednostavnom primeru ćemo razmotriti upotrebu lozinke za svrhu autentifikacije. Korisnik želi sa svojom lozinkom L da pristupi svom korisničkom nalogu. Servis koji autentifikuje korisnika, poseduje informaciju $y = f(L)$. Kada korisnik priloži lozinku L , servis računa $f(L)=y'$. Ukoliko su vrednosti y i y' jednake, proces autentifikacije biće uspešan. U ovom slučaju pretpostavićemo da je ovo siguran način za verifikaciju korisnika. Sigurnost sistema je zasnovana na neinvertibilnosti funkcije $f(L)$. Ovakve funkcije nazivaju se jednosmerne funkcije ili jednosmerne funkcije (eng. *one-way functions*).

Nažalost, ovo rešenje poseduje nekoliko problema kada se koristi lozinka L u realnom životu. Prvo, definicija jednosmerne funkcije podrazumeva da je L generisano na slučajan način sa uniformnom distribucijom. U našem slučaju, biometrijski podaci u ulozi lozinke su daleko od uniformne distribucije. Drugo, tradicionalna lozinka je precizno ponovljiva svaki put kada se koristi za autentifikaciju, dok to nije slučaj sa biometrijskim podacima. Sa druge strane, tradicionalna upotreba lozinke obezbeđuje relativno nizak nivo sigurnosti, jer je čovek u

stanju da zapamti relativno kratke lozinke. Visok nivo sigurnosti moguće je jedino obezbediti sa upotrebom biometrijskih lozinki. Ove lozinke mogu biti generisane na osnovu biometrijskih podataka kao što je biometrija prsta, irisa, glasa, geometrija dlana ili još bolje kombinacijom više biometrijskih izvora podataka. Biometrijski podaci sadrže neuporedivo veću količinu informacija nego tradicionalne lozinke koje su predviđene za pamćenje. Osnovni problem biometrijskih podataka je varijabilnost koja postoji u signalima, nakon očitavanja biometrijskih podataka iz dva identična biometrijska izvora. Drugim rečima, potrebna je mogućnost da se toleriše ograničen broj grešaka, zadržavajući sigurnost na visokom nivou, za razliku od nivoa sigurnosti koji je postignut upotrebom kratkih lozinki.

Autentifikacija lozinkama, o kojoj smo već diskutovali, je jedan jednostavan primer kriptografske aplikacije gde se postavljaju pitanja uniformnosti i tolerancije na grešku kod biometrijskih podataka. Ostali primeri primene uključuju mnoge kriptografske aplikacije, algoritme za šifrovanje i dešifrovanje, digitalni potpis ili identifikaciju, gde kriptološki ključ koji se koristi potiče ili proizilazi iz neuniformnih signala sa šumom.

Važan i opšti problem je proces konvertovanja neuniformnih biometrijskih podataka sa šumom u čisto slučajne uniformne nizove sa maksimalnom entropijom. Entropiju koju navodimo, odnosi se na entropiju binarnih informacionih izvora koju definišemo kao meru neodređenosti nekog sistema ili kao prosečnu količinu informacije koja je sadržana u jednom bitu.

Opšta ideja je zasnovana na dva koraka. Ekstraktujuća funkcija konvertuje biometrijske podatke u slučajne nizove sa uniformnom distribucijom koje je moguće koristiti u različitim kriptografskim aplikacijama za zaštitu podataka. U poslednjem koraku funkcija generiše javnu informaciju (eng. *Secure sketch*) koja se koristi za rekonstrukciju potpuno istih privatnih (tajnih) nizova u prisustvu biometrijskih obrazaca, dovoljno blizu aproksimacije originalnih biometrijskih podataka.

Autori [11] predlažu funkciju pod nazivom „*Fuzzy extractors*”. Funkcija ima osnovni zadatak da pretvara neuniformni slučajni podatak minimalne entropije u slučajni podatak sa uniformnom distribucijom i visokom entropijom.

Autori [12] predlažu šemu koja obezbeđuje generisanje više ključeva za primenu u različitim aplikacijama. Modifikovana je postojeća „*Fuzzy extractor*” šema, tako što je korišćena fiksna i jedinstvena permutacija nad obrascem irisa u fazi generisanja ključa. Kompromitovanjem jednog ključa koji je generisan iz biometrije jednog pojedinca, neće dovesti u pitanje ostale ključeve koji su izvedeni iz biometrije od istog pojedinca. U revidiranom dizajnu predložene šeme, ostaje pretpostavka da originalni biometrijski podaci ostaju tajna ili da neće biti kompromitovani.

Sledeći teorijski rad autora [11], takođe se bavi ekstraktovanjem ključa iz biometrijskih podataka sa šumom. Originalna šema je predložena od strane autora u radu [11 str. 3]. U datom predlogu, autori primenjuju kodove za ispravljanje grešaka za ulazni signal, a zatim i heš funkcije. Dokazano je da je količina curenja informacija o ulaznim podacima

zanemarljiva u odnosu na analizu izlaznih podataka. Ova vrsta pristupa može biti korisna samo ako se biometrijski podaci sa šumom čuvaju u tajnosti. Međutim, problem biometrijskih podataka je što ih ljudi ostavljaju svuda oko sebe ili mogu biti ukradeni. U ovom slučaju, biometrijski podatak u rukama napadača otkrio bi mnogo o entropiji ključa.

Autori [13] identifikovali su potencijalne ranjivosti kod postojećih biometrijskih sistema na bazi *Fuzzy* ekstraktora. Predstavili su šemu za zaštitu glavnog biometrijskog obrasca, kao i načine kako da se obezbedi zaštita biometrijskih obrazaca u odnosu na glavne karakteristike *Fuzzy* ekstraktor šeme za dati iris obrazac. Detaljno su opisali proces za generisanje i proces verifikacije na osnovu predložene *Fuzzy* ekstraktor šeme u radu. Autori su u zaključku izneli osnovne doprinose u radu, sa aspekta različitosti (kros-korelisanosti), opoziva, bezbednosti i performansi predloženog sistema.

Definicije koje smo naveli do sada su formalne, kao i efikasne tehnike za pretvaranje biometrijskih podataka u kvalitetnu informaciju koja bi obezbedila pouzdanu i sigurnu autentifikaciju na osnovu biometrijskih podataka, kao i primenu kod bilo koje kriptografske aplikacije. Opisana tehnika nije ograničena samo na primenu nad biometrijskim podacima, već može biti upotrebljena i za druge podatke koji su namenjeni za generisanje kriptoloških ključeva. Na osnovu bezbednosnih analiza koje su sprovedene u više preglednih radova, potencijalno je moguća dokazana sigurnost za predloženu šemu, samo ukoliko postoji pretpostavka da originalni biometrijski podaci ostaju tajna ili da neće biti ukradeni ili kompromitovani. U svim drugim slučajevima, napadač bi mogao da ima neke informacije o ključu, a to dalje ukazuje na postojanu mogućnost curenja informacija iz javnih podataka koji su predloženi ovom šemom.

2.3.3. Generisanje ključeva na osnovu biometrijskih podataka

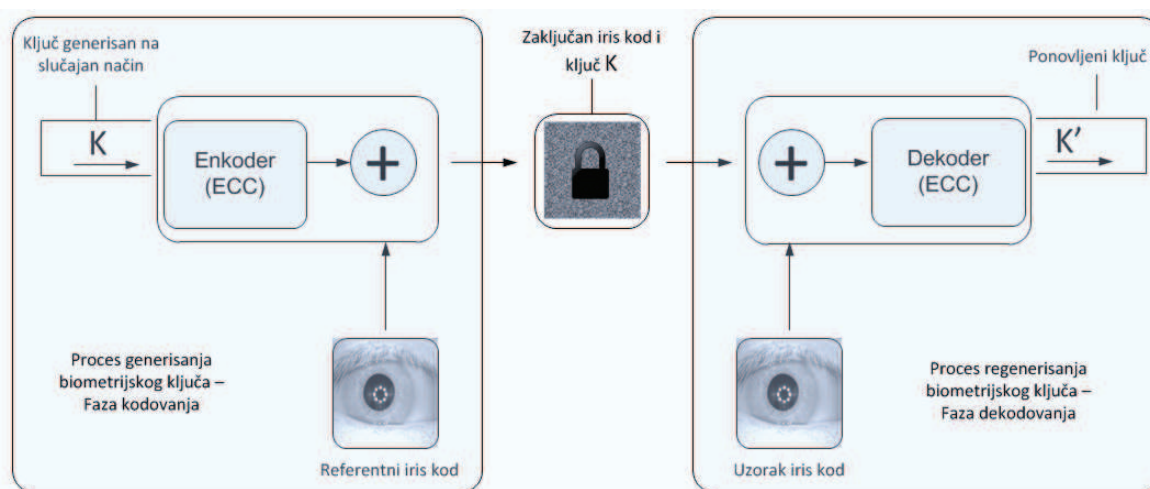
U ovom delu rada napravljen je pregled tehnika za generisanje kriptoloških ključeva na osnovu biometrijskih podataka. Za razliku od prethodne tehnike koja je navedena u delu 2.3.2., kriptološki ključevi se generišu iz biometrijskih podataka. Ova nova tehnika takođe obezbeđuje zaštitu privatnosti biometrijskih podataka, opoziv, kao i mogućnost ponovnog generisanja kriptoloških ključeva za autentifikaciju, kriptografsku zaštitu podataka, digitalno potpisivanje i druge kriptografske aplikacije. Neki autori radova su pokušali da unutar biometrijskih podataka ugrade jedinstvene i ponovljive binarne nizove koji dalje imaju ulogu kriptološkog ključa. Slede autori istaknutih naučnih radova koji su detaljno izučavali ovu tehniku: [14], [15], [16], [17], [18] i [1]. Ova tehnika omogućava ponovno regenerisanje ključa koji je u fazi upisa ugrađen u referentni podatak. Ostvareni doprinos preko ove tehnike je od izuzetnog značaja, iz razloga što je kriptografija konvencijalna nauka i zahteva istu tačnost koja postoji kod kriptoloških mehanizama.

Biometrijsko šifrovanje je adekvatan naziv ovog metoda koji predstavlja proces u kome se čvrsto vezuje kriptološki ključ za biometrijski obrazac (podatak). Suštinski, ključ je šifrovan sa biometrijskim podatkom primenom logičke operacije za sabiranje (eng. *XOR*), a rezultat je biometrijski kodovani ključ koji predstavlja javnu informaciju ili pomoćni podatak. Formom pomoćnog podatka obezbeđena je zaštita privatnosti, kao i sigurnost ugrađenih kriptoloških

ključeva. Pomoćni podaci mogu da budu sačuvani na nekom memorijskom prostoru, bez bezbednosnih rizika.

Kriptološki ključ moguće je ponoviti ili regenerisati ponovljenom operacijom (XOR), na osnovu pomoćnog podatka sa odgovarajućim predloženim biometrijskim obrascem koji je generisan u fazi verifikacije (eng. *Decommitted*).

Proces šifrovanja i dešifrovanja je prirodno konfuzan, iz razloga što je biometrijski podatak svaki put različit, za razliku od konvencijalnih metoda u kriptografiji. Imajući u vidu varijabilnosti u biometrijskim podacima, veliki je tehnološki izazov da ključ bude svaki put precizno ponovljen ili identičan. **Slika 2.**, predstavlja generičku šemu visokog nivoa za generisanje biometrijskih ključeva.



Slika 2. Šema za generisanje biometrijskog ključa

Posle procesa verifikacije ili oporavka ključa, ključ se može koristiti kod bilo kog tipa kriptografskih aplikacija, gde se ključ koristi kao lozinka ili može da služi za generisanje simetričnih ili asimetričnih ključeva. Bitno je istaći da biometrijsko šifrovanje nije samo po sebi kriptografski algoritam. Uloga biometrijskog šifrovanja je da zameni ili da osigura bezbednost kod postojećih šema za autentifikaciju, koje su bazirane na tradicionalnim lozinkama.

Biometrijsko šifrovanje ne treba mešati sa sistemima za šifrovanje biometrijskih podataka ili sa zaštitom biometrijskih obrazaca kriptografskim metodama za potrebe skladištenja, ili sa načinima za skladištenje kriptoloških ključeva na *smart* karticama ili drugim memorijskim *tokenima*, preko kojih je ključ oslobođen nakon uspešne verifikacije.

Međutim, metod za biometrijsko šifrovanje nalazi se u familijarnom odnosu sa metodom poznatim pod nazivom (eng. *Cancelable Biometrics*). Ovaj metod predstavlja prvi metod koji obezbeđuje opoziv biometrijskog obrasca, odnosno izdavanje novog ukoliko je postojeći kompromitovan ili je potreban novi za upotrebu u drugim aplikacijama. Priroda biometrijskih sistema nije takva da originalni biometrijski podaci mogu biti promenjeni sa drugim

biometrijskim podacima. Na primer, ukoliko neko snimi biometriju našeg prsta, lica ili irisa, ne postoji mogućnost generisanja nove biometrije.

Postoje dva osnovna pristupa za generisanje kriptoloških ključeva. U prvom slučaju, imamo potpuno nezavisno generisan ključ koji se čvrsto vezuje sa biometrijskim obrascem (eng. *Data binding*), a u drugom slučaju, kada ekstraktujemo ključ iz biometrijskog obrasca (eng. *Fuzzy extractor*). Kod oba pristupa su prisutni pomoćni podaci koji moraju biti negde skladišteni i od njih ovaj sistem zavisi.

Pomoćni podatak omogućava regenerisanje kriptološkog ključa u fazi verifikacije. Faza verifikacije podrazumeva primenu logičke operacije XOR između tog generisanog biometrijskog obrasca i pomoćnog podatka. Veličina prostora ključeva na pomoćnom podatku zavisi od stepena varijabilnosti biometrijskih obrazaca koji na ovaj način utiču na kapacitet kodova za ispravljanje grešaka i na ukupnu količinu neuzajmne informacije poređenjem različitih biometrijskih obrazaca.

U režimu vezivanja biometrijskog obrasca sa ključem, ključ je slučajno i nezavisno generisan tako da ni korisnik niti bilo ko drugi zna nešto o ključu. Nezavisnost ključa od biometrijskog podatka omogućava izmenu ukoliko je potrebno. Nakon generisanja biometrijskog podatka, algoritam na siguran način vezuje biometriju sa ključem i tako generiše biometrijski šifrovan ključ. Dobijeni pomoćni podaci obezbeđuju zaštitu privatnosti i mogu biti sačuvani u bazi podataka ili lokalno na *smart* karticama, računarima, *tokenu* ili mobilnom telefonu. Na kraju kompletne procedure (kada je generisan pomoćni podatak), ključ i biometrijski obrazac su uništeni na siguran način.

Osnovni metod je tako dizajniran da bude prihvatljiv za određenu količinu šuma kod ulaznih biometrijskih podataka. Ukoliko neko na nelegitiman način pokuša da regeneriše ključ, tako što koristi različit biometrijski obrazac od originalnog, algoritam neće biti u mogućnosti da ponovi isti ključ. Mnoge šeme za biometrijsko šifrovanje čuvaju heš vrednosti ključeva na mestima gde se dati ključevi skladište.

Nakon uspešne verifikacije, algoritam za dobijeni ključ računa heš vrednost na osnovu koje se proverava identičnost dobijenog ključa u isto vreme. U praksi, aplikacije koriste heš vrednosti za derivaciju ključeva. U ovom slučaju to nije moguće jer se originalna heš vrednost čuva na *tokenu*. Sa ovakvom arhitekturom obezbeđenja, napadač ne može da dobije originalni ključ izvan domena metoda za biometrijsko šifrovanje. U praktičnoj primeni biometrijska slika ne treba da bude poslata preko komunikacionog kanala do servera, već je verifikaciju bitno realizovati lokalno.

Autori [4] predlažu novu šemu poznatu pod nazivom „*Fuzzy commitment scheme*”. Ova šema predstavlja metod za prikrivanje (eng. *Concealing*) biometrijskih podataka i šemu za uvezivanje podataka (eng. *Binding*). U konvencionalnoj šemi za otključavanje ovog tipa nastalog podatka neophodan je jedinstven identifikator, sličan kao što je kriptološki ključ kod algoritma za dešifrovanje. Karakteristike predložene šeme su vrlo primenjive kod aplikacija kao što su biometrijski sistemi za autentifikaciju, kod kojih biometrijski podaci sadrže šum.

Šema je tolerantna na nastale greške, sposobna da zaštiti biometrijske podatke slično kao kriptografski mehanizmi zaštite, ili kao jednosmerne funkcije koje se koriste za zaštitu alfanumeričkih lozinki.

Autori [19] predlažu sistem baziran na otisku prsta. Oni su prvi koji su komercijalizovali ovu tehnologiju u proizvod (www.biocrypt.com). Fazna informacija se ekstrahuje sa slike otiska prsta, koristeći Furijeovu transformaciju i kodove za ispravljanje grešaka, da bi smanjili varijabilnost u biometrijskom signalu. Umesto da generišu ključ iz biometrijskih podataka, oni uvode metod biometrijskog zaključavanja. Zapravo, unapred određeni slučajni ključ se zaključava sa dobijenom faznom informacijom (biometrijski obrazac). Dobijeni rezultat moguće je jedino otključati sa originalnom biometrijom. U ovom radu prikazana je obećavajuća ideja o biometrijskom zaključavanju jer je ključ sada moguće definisati nasumično i nezavisno od sistema. Performanse predloženog sistema nisu date.

Autori [16] u svom radu koriste svojeručni potpis. Definisali su 43 karakteristike koje mogu da ekstrahuju iz takve dinamične informacije, kao što su brzina, pritisak, nadmorska visina i azimut (pravac određen uglom u horizontalnoj ravni). Na osnovu svih pribavljenih vrednosti navedenih parametara, generiše se jedan konkatentni niz iz koga se funkcijom ekstrahuje informacija i rezultat je jedan binarni niz. Na ovaj način generisan je ključ sa entropijom od 40 bitova sa stopom lažnog odbacivanja FRR od 28% i stopom lažnog prihvatanja FAR negde oko 1,2%.

Autori [17] kombinuju slične tehnike koje su gore navedene, projektuju sistem baziran na biometriji lica. Usvajaju metod biometrijskog zaključavanja predloženog od autora [19]. Uspeli su da regenerišu ključ sa entropijom od 80 bitova sa stopom lažnog odbacivanja FRR od 0,93%. Dobijene vrednosti parametara FRR predstavljaju najpribližnije očekivane vrednosti za praktičnu primenu sistema. Sistem je testiran na slikama lica. Slike su dobijene preko kontinualnog video signala, a ne na osnovu slika iz baze podataka. Postoje određene sumnje kod evaluacije ovog rada, iz razloga što u kontinualnom video signalu postoji nešto manja varijacija, nego što je u slikama iz baze podataka.

Autori [20] predlažu šemu pod nazivom virtuelni PIN koja je praktično identična šemi u radu [4]. Korišćen je Gabor filtr za ekstrakciju informacionih sadržaja. Autori predlažu korišćenje (eng. *Low Density Parity Check*) zaštitnih kodova za ispravljanje grešaka. Evaluacija predloženog rešenja je urađena na neutvrđenoj bazi podataka.

Veoma sličan, ali više praktičan pristup je realizovan na biometriji otiska prsta, gde je za ekstrakciju informacionih sadržaja korišćen fazni Gaborov filtr. Autori [21] izdvajaju konzistentne i pouzdane bitove informacija iz biometrije prsta koristeći BCH kodove za ispravljanje nastalih grešaka. Sprovedenom evaluacijom u radu dobijaju sledeće rezultate, vrednost parametra FRR od 5,4% do 9,9% za ključ dužine 49 bitova i 85 bitova. Međutim, parametar FAR od 3,2% do 2,5% je relativno visoka vrednost u odnosu na postignutu dužinu ključa.

Autori [2] su predstavili šemu za biometrijsku kriptografiju na osnovu biometrije lica. Generička šema je zasnovana na šemi čvrstog vezivanja biometrije sa nezavisno generisanim ključem, tehnika koja je predstavljena od strane autora [4]. Sa ciljem poboljšanja tačnosti algoritma, autori ekstraktuju pouzdane i konzistentne bitove informacija iz biometrijskih uzoraka i primenjuju BCH kodove za ispravljanje grešaka. Rezultat koji su dobili relativno je dobar sa FRR parametrom od 3,5% za standardizovanu bazu podataka sa niskim i srednjim nivoom varijabilnosti kod slika i parametrom FAR od 0% u oba slučaja. Algoritam je kasnije implementiran u Philips sistem koji je nazvan "*privIDTM*".

Autori [22] su predstavili bimodalni biometrijski sistem za generisanje kriptološkog ključa na osnovu biometrije glasa i signala elektrokardiograma - EKG. Jedinstveni informacioni sadržaj iz oba biometrijska izvora je ekstraktovan preko transformacione funkcije. Rad je zasnovan na jedinstvenom i kvazi stacionarnom ponašanju EKG signala i govornog signala koji predstavlja bihejvioralnu osobinu pojedinca. Sprovedenom evaluacijom u radu dobijeni su sledeći rezultati: vrednost parametra FAR od 1,27%, i vrednost parametra FRR od 10,62%.

Navedeni primeri predloženih šema generišu ključeve relativno manjih dužina i većina njih ima visoku stopu lažnog odbacivanja FAR, što se negativno odražava na društvenu prihvatljivost nekih rešenja. Na osnovu uvida u postignute rezultate, autori [1] predlažu generičku šemu za generisanje ključeva na osnovu biometrije irisa preko koje su postigli dobre performanse koje obezbeđuju praktičnu primenu njihovog rešenja.

Iz tog razloga, u nastavku ćemo detaljno proći kroz kompletnu šemu algoritma predloženu u njihovom radu. Pre dizajniranja šeme za ispravljanje grešaka, prvo je urađena karakterizacija prisutnih grešaka u iris podatku. Nakon urađene analize, dobili su rezultate na osnovu kojih su izabrali adekvatne zaštitne kodove i optimizovali iste u odnosu na maksimalni dozvoljeni kapacitet za ispravljanje grešaka. Njihov algoritam prvo izdvaja 256 bajtova fazne informacije upotrebom 2D Gabor vejvleta koji je opisan u radu istog autora [23]. Uočeno je da postoje različiti tipovi grešaka u iris podatku. Klasifikacija je napravljena prema dva osnovna tipa grešaka. Prva klasa grešaka, predstavlja pozadinske slučajne greške izazvane od strane CCD senzora kamere ili distorzije irisa koja je prouzrokovana promenom osvetljenja. Kako god, nastale greške nije moguće korigovati postojećim algoritmima za obradu signala. Druga klasa grešaka je u formi paketskih grešaka i nastaje zbog lošeg detektovanja trepavica, kapaka i refleksije. Osnovne ideje za primenu tehnika za kodovanje, predložene su u nekim ranijim radovima koje su korišćene za ispravljanje grešaka, predložene od autora [19] i [24]. Međutim, navedene tehnike ne rade dobro i na iris podatku, jer višestruko skeniranje ne smanjuje količinu greške. Utvrđeno je da na grupi slika od 70 korisnika bez korišćenja maske za pojedinačno skeniranje, stopa greške u proseku iznosi 13,69%, a nakon tri skeniranja stopa greške je smanjena na 10,68% i na 9,36% posle pet skeniranja. Da bi se izborili sa ovim tipom grešaka, autori u radu predlažu upotrebu konkatentnih kodova za ispravljanje grešaka. Za prvu klasu grešaka primenjuju Hadamardove kodove, a zatim na drugu klasu grešaka (paketske greške) primenjuju Rid-Solomon kodove. Kao spoljašnji sloj, koriste se

Hadamardovi kodovi za ispravljanje slučajnih grešaka u binarnom podatku irisa, a za unutrašnji sloj Rid-Solomon kodovi za ispravljanje paketskih grešaka na blokovskom nivou.

U predloženoj šemi sa dva faktora ključ zavisi od kombinacije biometrije i *tokena*. Pretpostavlja se da napadač, ukoliko uspe da dođe u posed *tokena*, može imati potpuno znanje o podacima na njemu. Početni dizajn imao je za cilj da pokaže da ukoliko jedan faktor šeme bude kompromitovan, napadač neće imati mogućnost da regeneriše ključ. Kasnije u radu je pokazana mogućnost da se šema produži sa trećim faktorom, dodavanjem lozinke. Originalna slika šeme je prikazana u radu [1 str. 3]. Šema predstavlja mogućnost premošćivanja nekompatibilnosti koja postoji između varijabilnih biometrijskih podataka i konvencionalne kriptografije.

U prvoj fazi, generiše se nezavisno biometrijski ključ k kao slučajna binarna sekvenca. Ključ k poseduje sve osobine kriptoloških ključeva. Ključ k je tačno zadate dužine. U sledećoj fazi ključ k se koduje konkatentnim kodovima, čija je dužina na kraju faze kodovanja 2048 bitova. Kodovani ključ sa konkatentnim kodovima naziva se „pseudoiris kod“. Zatim kodovani ključ k ili pseudoiris kod se čvrsto vezuje logičkom operacijom XOR za biometriju irisa i tako dobijeni rezultat predstavlja pomoćni podatak koji se skladišti na nekom hardverskom *tokenu* zajedno sa heš vrednošću ključa k .

Nakon kompletnog procesa za upis ili generisanje biometrijskog ključa, neophodno je uništiti ključ na siguran način. Faza dekodovanja je potpuno inverzna fazi kodovanja. Korisnik podmeće originalnu biometriju irisa da bi otključao pseudoiris kod iz pomoćnog podatka, preko ponovljene logičke operacije XOR. Dalje sledi faza dekodovanja sa istim konkatentnim kodovima, samo u suprotnom redosledu od kodovanja. Da bismo potvrdili uspešnost čitave operacije, za dobijeni rezultat k' računa se heš vrednost preko koje se utvrđuje sličnost sa heš vrednošću koja se nalazi na *tokenu*.

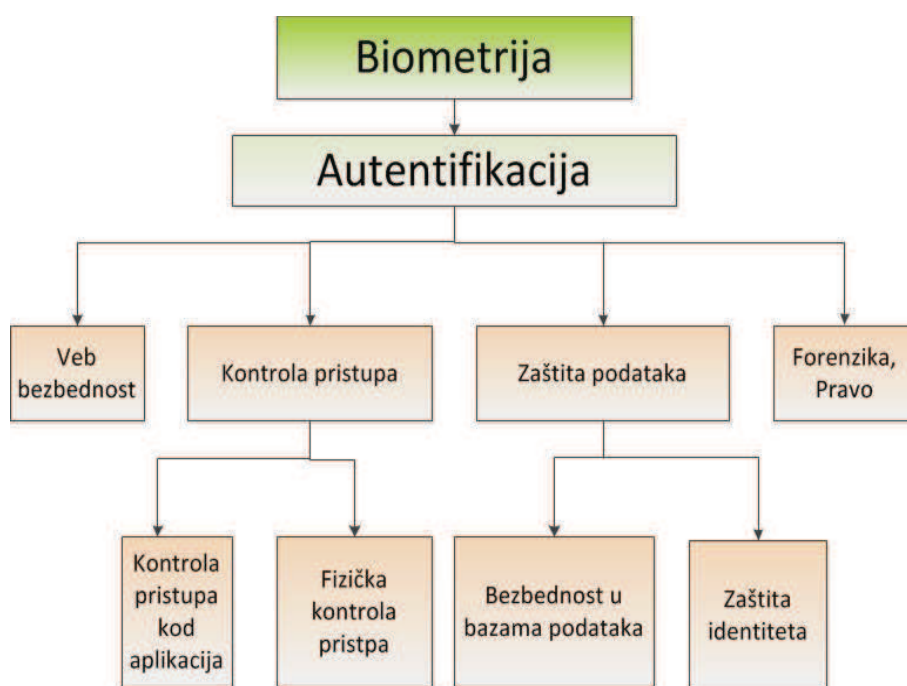
Cilj koji je postignut, zasniva se na detaljnoj analizi obrasca greške u obrascu irisa. Iris kodovi (obrasci) od jednog oka obično se razlikuju od 10 – 20 % bitova, saopšteno od autora [23]. Sa druge strane, razlika između dva irisa kod različitih osoba ili dva irisa iste osobe, iznosi 40 – 60 % bitova. Hadamardov kod je izabran za ispravljanje 25 % pogrešnih bitova [25], koji precizno razdvaja ispravljanje grešaka na biometrijskom uzorku istog i različitog oka.

Autori [26] kombinuju na sličan način tehnike koje su korišćene od strane autora [1]. Za spoljašnji sloj, koriste se Hadamardovi kodovi za ispravljanje slučajnih grešaka u biometrijskom obrascu irisa, a za unutrašnji sloj Rid-Solomon kodovi za ispravljanje paketskih grešaka na blokovskom nivou. Napravljene su određene izmene u šemi. Radi povećanja kapaciteta kodova za ispravljanje grešaka, uvedeni su redundantni bitovi (nule), sa uniformnom distribucijom, čime je dužina iris koda povećana za 40%. Uveden je metod mešanja iris obrasca za permutaciju koja je kontrolisana lozinkom, što je dalje uticalo na povećanje Hamingovog rastojanja između različitih irisa, kao i smanjenje rastojanja između istih irisa. Drugi doprinosi su izraženi preko povećanja entropije iris obrasca. Postignuta je dužina ključa od 186 bitova sa stopom lažnog odbacivanja FRR od 0,76% i stopom lažnog

prihvatanja FAR od 0.096%. Sa aspekta bezbednosti, zaključani iris kod ne otkriva biometrijsku informaciju u slučaju kompromitovanja kriptološkog ključa, *smart* kartice i lozinke. Obezbeđena je mogućnost opoziva i mogućnost ponovnog generisanja biometrijskog ključa.

2.4. Primene tehnologije zasnovane biometrijskoj kriptografiji

U ovoj eri informacionih tehnologija popularnost raste za sisteme koji koriste biometrijsku tehnologiju, zbog njihovog napredovanja, praktičnosti i preciznosti u pouzdanom određivanju identiteta korisnika u odnosu na tradicionalne metode. Biometrijski sistemi mogu da se koriste u svrhe provere verodostojnosti (**Slika 3**), a to uključuje sigurnosne sisteme za kontrolu pristupa kod finansijskih servisa (ATM i banke), vladine institucije, digitalnu forenziku i sprovođenje zakona, granične kontrole pristupa i Web bezbednost koja dalje uključuje sve segmente kriptografske zaštite.



Slika 3. *Primena biometrijskih sistema*

Autori [27] predlažu šemu za zaštitu multimedijalnih sadržaja na osnovu biometrijskih podataka sa primenom kriptografske zaštite. Pomoću biometrijskih podataka, zajedno sa identifikatorom kao što je ključ, moguće je sprečiti neovlašćeno korišćenje zaštićenih sadržaja, kombinacijom simetričnih i asimetričnih ključeva koji su korišćeni u ovu svrhu. Predloženim metodom moguće je dopuniti sisteme za izradu vodenog žiga (eng. *Watermarking*).

Autori [28] predlažu šemu na osnovu biometrije otiska prsta za autentifikaciju, za primenu kod sistema za Veb testiranje i pohađanje Veb kurseva. Krajnji cilj je da rešenje bude

prihvatljivo sa aspekta univerzitetske kulture, a sa druge strane da minimalno remeti tradicionalne procedure i procese.

Autori [29] predlažu šemu za upravljanje digitalnim sadržajima DRM (eng. *Digital rights management*). Šema se zasniva na modelu koji sadrži DRM, i identifikacione metode koje se baziraju na biometriji irisa, kao dodatni sloj bezbednosti za pristup određenim podacima. Šema se odlikuje sa nepromenljivošću i sa zaštitom individualnosti. Šema je dizajnirana tako da razlikuje klijente od nelegitimnih korisnika, tako da samo ovlašćeni klijenti imaju pristup digitalnim sadržajima. Uvođenjem biometrije irisa i razvojem biometrijske kriptografije, tradicionalna autentifikacija bazirana na lozinkama je zamenjena biometrijom irisa.

Autori [30] predlažu šemu za upravljanje kriptološkim ključevima na osnovu biometrije otiska prsta, sposobne za regenerisanje stalno tačnih ključeva na osnovu pravog biometrijskog obrazaca, imajući u vidu količinu šuma koja je sadržana u biometrijskim podacima. Njihova šema je razvijena na tehnici čvrstog vezivanja ključa i biometrijskog obrasca. U datim rezultatima nisu date konkretne dužine ključeva, koje je moguće postići sa predloženom šemom.

Autori [31] predlažu šemu u kojoj koriste iris biometriju za generisanje ključeva u PKI infrastrukturi. Analizirani su detaljno metodi za ekstrakciju informacionih sadržaja iz slike irisa, kao i mehanizmi za generisanje javnih ključeva. Slike su u RGB formatu. Dobijena je veća količina informacija, dovoljna za generisanje ključeva većih dužina za potrebe tradicionalne PKI infrastrukture. Dobijeni rezultati potvrđuju opravdanost predloženog rešenja. U evaluaciji rada nisu date informacione analize koje bi govorile o kvalitetu generisanih ključeva.

Autori [32] izlažu ideju o bio-haotičnoj protočnoj (eng. *Stream cipher*) šifri sa kojom šifruju slike elektronskih medija, koristeći biometrijski ključ i bio-haotičnu funkciju. Bezbednost slika je na ovaj način stavljena na viši nivo i tako je smanjen rizik njihovog kompromitovanja. U radu su prikazane zapravo dve ideje. Razvoj ideje o novoj bio-haotičnoj šifri i ideja o metodu za generisanje biometrijskih nizova u ulozi kriptoloških ključeva za šifrovanje i dešifrovanje.

3. Teorijske osnove istraživanja

U teoriji informacija, centralni pojam zauzima pojam informacije kao fenomena i komunikacija kao prenos. Intuitivna predstava pojma informacije je gotovo svima nama bliska, ali naučna zasnovanost i interpretacija informacije utemeljena je tek polovinom prošlog veka, rađanjem nove naučne discipline teorije informacija.

Informacija je postala relevantan pojam za sve nauke koje se bave simboličkom komunikacijom u rasponu od matematike do računarske nauke, ili od logike do lingvistike, odnosno od elektronike do bibliotekarstva, kao i od humanističkih nauka i umetnosti do dokumentalistike, ali i od društvenih nauka do medicine i dr.

Sve je to informaciji dalo interdisciplinarnu dimenziju, jer je svaka nauka pokušala i još pokušava da protumači taj kompleksan pojam. Sve ovo ukazuje da pojam informacija nije lako shvatiti, niti jednostavno protumačiti.

Informacija je reč latinskog porekla „*in formare*“ i izvorno je značila stavljanje u određenu formu, odnosno, davanje oblika nečemu, ali je tokom vremena izgubila prvobitno značenje. Danas nema jedinstvenog pristupa fenomenu informacije, a još manje jednoznačne i opšteprihvaćene definicije informacije.



Slika 4. Piramida znanja - PIZP

(Slika 4), predstavlja piramidu znanja čija je skraćenica PIZM – (podatak, informacija, znanje, mudrost). PIZM je predlog strukturiranja podatka, informacija, znanja i mudrosti u jednu informacionu hijerarhiju gde svaki nivo dodaje određena svojstva, iznad i ispod onog prethodnog.

U ovoj strukturi podatak je najosnovniji nivo, informacija dodaje kontekst, znanje dodaje kako ga upotrebiti i mudrost dodaje kada i zašto ga upotrebiti. PIZM model se temelji na nizu postupaka :

- Podatak dolazi u obliku neobrađenih zapažanja i dimenzija;
- Informacija se oblikuje analiziranjem veza i odnosa između podataka;
- Znanje se oblikuje koristeći informaciju za delovanje;
- Mudrost se oblikuje kroz upotrebu znanja, kroz komunikaciju korisnika i kroz razmišljanja. Opšte prihvaćeno mišljenje je da je podatak manji od informacije, a informacija je manja od znanja, odnosno, da bi se kreirala informacija potreban je podatak i samo kada postoji informacija, znanje može izaći na videlo;

Sadržaj ljudskog uma prema Raselu Ackoffu (*Russell Ackoff*¹) čine: podatak, informacija, znanje, razumevanje (saznajni i analitički proces koji odgovara na pitanje zašto) i mudrost. Prve četiri kategorije odnose se na prošlost i bave se onim što je bilo ili onim što je poznato. Samo se peta kategorija, mudrost, bavi budućim vremenom jer sadrži maštovitost, pronicljivosti i kreativnost. Do mudrosti se ne dolazi brzo i jednostavno. Prvo se mora uspešno proći kroz sve prethodne kategorije.

Pomoću razumevanja se na osnovu postojećih znanja može stvarati novo, odnosno, ljudi koji razumeju mogu preduzimati korisne aktivnosti, zato što mogu sastaviti novo znanje ili barem nove informacije na osnovu već postojećeg znanja. Razlika koja postoji između znanja i razumevanja je kao i između učenja i memorisanja.

U informacionom društvu znanje je rezultat učenja. Povećanje nivoa znanja dovodi do boljeg odlučivanja, a nove informacione tehnologije dovode do ubrzanja i sve većih količina podataka i informacija. Sveprisutnost informacija dovela je do toga da savremeno društvo već označavamo kao informaciono društvo. Razlog za to ne leži samo u sveprisutnosti, već i ključnom značaju informacija u funkcionisanju gotovo svih ljudskih delatnosti.

Danas već imamo informacionu teoriju sistema, informacionu teoriju odlučivanja, informacionu filozofiju, informacionu teoriju svega u okviru teorijske fizike, informacionu medicinu i tako redom.

Ovo je samo jedan primer koji potvrđuje već postojeće stanje stvari po kome je informacija izabrala status resursa, a isti taj status u klasičnoj privrednoj delatnosti imaju sirovina, energija ili ljudski rad. Ako informacija postaje dominantni resurs privredne delatnosti, onda se tim resursom mora pažljivo rukovati.

Stoga prenos, memorisanje i zaštita informacija dolaze na prvo mesto, kao kritične tehnologije kojima se obezbeđuje stabilan razvoj i u krajnjoj liniji opstanak jedne društvene zajednice.

¹ Russell Ackoff , http://en.wikipedia.org/wiki/Russell_L._Ackoff

Upravo teorija informacija daje teorijske i praktične okvire za sintezu sistema koji obezbeđuju efikasnost i raspoloživost informacionih resursa, shvaćenih na gore pomenuti način. Teorija informacija u tehničkom smislu, kako se danas razume i tumači, vodi poreklo od radova Kloda Šenona (Claude Shannon²) objavljenih neposredno nakon Drugog svetskog rata.

Šenon je definisao ovu disciplinu kao teoriju za rešavanje problema prenosa signala preko komunikacionih kanala. Iako je ovakav pristup suviše uzak, pogotovu ako se ima u vidu merenje informacionog (semantičkog) sadržaja poruka koje se memorišu ili prenose, Šenonov pristup predstavlja i dalje centralno jezgro ove teorije.

Teoriju informacija kao akademsku disciplinu je zasnovao Klod Šenon. Šenon je radio u *Bell Laboratories Institute of Technology (1956-1978)* i *Institute for Advanced Study (1940)*. Šenon je tražio odgovor na dva fundamentalna pitanja teorije komunikacija:

- 1) Do koje mere se neki skup podataka može komprimovati? (odgovor: entropija izvora H)
- 2) Šta je krajnja brzina komunikacija za zadati komunikacioni kanal? (odgovor: kapacitet kanala C)

Odnos teorije informacija sa telekomunikacijama nastao je pojavom Šenonovih radova [33] i [34] smatralo se da je nemoguće slati informacije pozitivnom brzinom sa zanemarljivom verovatnoćom greške. Šenon je u svom fundamentalnom radu iz 1948. godine iznenadio celokupnu naučnu zajednicu iz domena telekomunikacija pokazavši da je grešku prenosa moguće smanjiti proizvoljno za sve brzine kapaciteta kanala. Šenon je dalje pokazao da informacioni izvori kao što su muzika ili govor poseduju kompleksnost ispod koje se signali ovih izvora ne mogu dalje komprimovati bez gubitka. Ovu kompleksnost je nazvao entropija i razlikovao je od pojma uvedenog mnogo ranije u oblasti termodinamike. Dakle, ako je entropija izvora manja od kapaciteta kanala, moguće je dizajnirati komunikacionu šemu koja je asimptotski bez greške prenosa.

Pojava turbo kodova približava nas Šenonovoj granici (dakle komunikaciji sa maksimalno mogućom brzinom u blizini kapaciteta kanala, sa nultom greškom prenosa). Front oblast je danas dominantno „mrežno kodovanje“ (eng. *Network coding*), odnosno teorija simultanog komuniciranja više čvorova u jednoj mreži, različitim brzinama u prisustvu šumova i interferencija.

Kolmogorov, Čaitin i Solomonov su razvili teoriju da je kompleksnost jednog niza jednaka dužini najkraćeg računarskog programa koji generiše date podatke. Kompleksnost je dakle jednaka dužini minimalnog opisa (Deskriptivna kompleksnost). Pokazalo se da je ideja Kolmogorovljeva kompleksnost univerzalna i nezavisna od konkretnog računara.

Kolmogorovljeva kompleksnost K je približno jednaka Šenonovoj entropiji H , ukoliko je posmatrana sekvenca za koju se ona računa izvučena slučajno u skladu sa nekom raspodelom

² Claude Shannon - http://en.wikipedia.org/wiki/Claude_Shannon

verovatnoće koja ima entropiju H . Stoga je veza između Šenonove teorije informacija i Kolmogorovljeve kompleksnosti gotovo perfektna.

Odnos teorije informacija sa fizikom predstavlja statistička mehanika, koja je mesto rođenja entropije i drugog zakona termodinamike. Entropija uvek raste. Drugi zakon termodinamike daje teorijsku osnovu za odbacivanje svake tvrdnje o egzistenciji perpetumobila.

Odnos teorije informacija sa matematikom predstavljen je sledećim fundamentalnim veličinama teorije informacija:

- Entropija
- Relativna entropija
- Uzajamna informacija

Sa svoje strane ove veličine dobro opisuju ponašanje dugačkih sekvenci slučajnih veličina i omogućavaju nam procenu verovatnoća retkih događaja (teorija velikih devijacija), kao i ocene greške eksponenta pri testiranju hipoteza.

Odnos teorije informacija sa filozofskim naukama. Solomonov i Čaitin su ubedljivi u dokazu da je univerzalno najbolja procedura predikcije ona koja uzima težinsku sumu predikcija svih programa koji dobro objašnjavaju podatke. Naravno da je ova univerzalna procedura nepraktična, budući da je pronalaženje svih programa koji ne mogu da objasne date podatke vremenski zahtevan postupak.

Odnos teorije informacija sa ekonomijom predstavljen je repetitivnim investicijama u stacionarno berzansko tržište koje rezultuje u eksponencijalnom rastu dobitka. Rast dobitka je dualan entropiji berzanskog tržišta. Postoji jednoznačna veza između teorije optimalnog investiranja u berzansko tržište i teorije informacija.

Odnos teorije informacija sa računarstvom i komunikacijama nastao je pravljenjem velikih računara na osnovu malih komponenti koje nas dovode i do računarskih i do komunikacionih limita. Računarstvo je limitirano komunikacijom, dok su komunikacije računarski limitirane. Stoga su svi rezultati teorije informacija u domenu komunikacija od direktnog uticaja i na računarstvo.

3.1. Pojam neodređenosti i informacije

Neku situaciju ćemo označiti kao neodređenost ukoliko postoji više mogućnosti, pri čemu nije poznato koja će od njih biti realizovana ili izabrana. Zapravo realne situacije potpune određenosti u kojima se ishodi mogu predviđati sa sigurnošću su više izuzetak nego pravilo. Uvedimo na formalniji način pojam neodređenih situacija, koji je početni gradivni blok savremene teorije informacija. Pretpostavimo neku situaciju S u kojoj je moguće m različitih ishoda. Označimo ove mogućnosti sa $e_1, e_2, \dots, e_m$. Ovaj skup mogućnosti ćemo označiti sa $S(e_1, e_2, \dots, e_m)$ i nazvati konačnom šemom izbora.

Po nekom mehanizmu ili na osnovu nekog procesa ili na osnovu nečije volje vrši se izbor jedne od ovih mogućnosti. Neodređenost asocirana zadatoj šemi izbora S , nastaje usled našeg neznanja ili nemogućnosti da znamo koja će od m alternativa biti izabrana. Na koji način možemo izmeriti količinu neodređenosti u datoj šemi izbora S ? Intuitivno predosećamo da što je veća kardinalnost $|S|$ skupa S , odnosno broj elementa ovog skupa, veća je i neodređenost. Zašto onda ne bismo jednostavno uzeli $|S|$ kao meru neodređenosti? Iako je to u principu moguće, jedan drugi pristup će se pokazati delotvornijim. Predstavimo sebi sledeću igru: zamisli se jedna od mogućnosti iz skupa $S(e_1, e_2, \dots, e_m)$, a zatim se na osnovu pitanja koja su tako formulisana da se na njih odgovara samo sa „da“ ili „ne“, pokušava pogoditi zamišljena vrednost.

Mogućnosti uvek možemo označiti redom brojevima $1, 2, 3, \dots, m=|S|$, tako da pitanja mogu biti tipično sledećeg oblika: Da li je zamišljeni broj paran? Da li je zamišljeni broj manji od 20? Da li je zamišljeni broj veći od 12? i tako redom. Odmah zapažamo da ukoliko nam je potrebno više pitanja, neodređenost šeme izbora S je veća. Stoga je osnovano smatrati da je broj pitanja tipa „da“ i „ne“ neophodnih za određivanje nepoznate zamišljene mogućnosti iz skupa S , dobra mera neodređenosti šeme izbora S .

Od važnosti je i kakva pitanja postavljamo. Ako su pitanja tipa da li je zamišljeni broj jednak 1, pa ako nije da li je jednak 2, pa ako nije da li je jednak 3, jasno je da nas to vodi ka postavljanju svih m mogućih pitanja. Ovo sigurno nije optimalan način za dobijanje tačnog odgovora u najmanjem broju koraka. Međutim, ako je prvo pitanje da li je zamišljeni broj manji od $m/2$, u narednom pitanju sužavamo pretragu na polovinu početnih mogućnosti.

Ako i u narednim koracima nastavimo da postavljamo pitanja na ovaj način, pokazaće se da je izabrana strategija vrlo efikasna. U cilju detaljnijeg upoznavanja sa ovim postupkom, pretpostavimo prvo da je ukupan broj početnih mogućnosti stepen broja 2, odnosno da važi $m = 2^n$. Tada možemo podeliti S postavljanjem prvog pitanja (da li je zamišljeni broj veći od 2^{n-1} ?) na dva skupa iste veličine: $\{1, 2, \dots, 2^{n-1}\}$ i $\{2^{n-1}+1, \dots, 2^n\}$.

Svaka od ovih polovina će bit podeljena na dve polovine narednim pitanjem. Ako je odgovor na prvo pitanje – „ne“, tada sledeće pitanje generiše skupove $\{2^{n-1}+1, \dots, 2^n\}$ i $\{2^{n-2}+1, \dots, 2^{n-1}\}$. Ako je odgovor na prvo pitanje – „da“, tada naredno pitanje pravi razliku između skupova $\{2^{n-1}+1, \dots, 2^{n-1}+2^{n-2}\}$ i $\{2^{n-1}+2^{n-2}+1, \dots, 2^n\}$.

Formalizovali smo situacije neodređenosti uvođenjem sistema S , u kome se bira jedna od mogućnosti, ali se ne zna koja. Neodređenost pridružena sistemu mogućnosti se meri najmanjim brojem pitanja koja treba postaviti u cilju nalaženja stvarnog izbora. Na ovaj način smo došli do mere neodređenosti izražene sa $\log|S|$. Ako su upotrebljena binarna pitanja, jedinica mere neodređenosti je bit. Sistemom postavljanja pitanja u vezi određivanja date nepoznate mogućnosti, indukuje se stablo na osnovu koga se može jednoznačno odrediti kod

svake od mogućnosti sistema S . Dužine kodova su približno jednake meri neodređenosti $\log|S|$. Ako su korišćena binarna pitanja, dobija se binarni kod.

Minimizacija očekivane vrednosti broja postavljenih pitanja nije trivijalan zadatak. Rešenje ovog problema je poznato i veoma se koristi u teoriji kodovanja. Ključna ideja ovog postupka je da se skup svih mogućnosti ne deli na podskupove jednake kardinalnosti, već jednakih verovatnoća. Iz teorije kodovanja je poznato da je očekivani broj pitanja aproksimativno iznosi:

$$-\sum_{i=1}^m p_i \log p_i \quad (1)$$

Ova veličina se naziva entropija i predstavlja meru neodređenosti količine neodređenosti verovatnosnih sistema mogućnosti. Ukoliko je neka od verovatnoća $p_i = 0$, usvaja se konvencija da je $0 \cdot \log 0 = 0$, što sledi iz činjenice da je $\lim_{x \rightarrow 0} x \log x = 0$.

DEFINICIJA 1. Neka je (S, P) verovatnosni sistem mogućnosti. Tada je količina neodređenosti u njemu definisana entropijom:

$$H(P) = -\sum_{i=1}^m p_i \log p_i \quad (2)$$

U definiciji (1) postoji određeni stepen proizvoljnosti, budući da nije definisana osnova logaritma. Baza k logaritma odgovara broju mogućih odgovora u eksperimentu postavljanja pitanja i dobijanja odgovora. Promena baze menja samo jedinicu mere. Binarni slučaj je najpopularniji i njemu odgovara jedinica bit.

Budući da je entropija fundamentalan pojam teorije informacija, navešćemo njene osnovne karakteristike. Uбудуće ćemo za sistem mogućnosti nad verovatnoćama $P = \{p_1, p_2, \dots, p_m\}$ označavati entropiju tog sistema sa $H(p) = H(p_1, p_2, \dots, p_m)$. Prvo ćemo povezati opšti pojam entropije sa pojmom mere neodređenosti, za sisteme mogućnosti u kojima nisu definisane verovatnoće pojedinih izbora. Ako je P uniformno raspodeljeno po svih n mogućnosti, tada je:

$$H\left(\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n}\right) = \sum_{i=1}^n \frac{1}{n} \cdot \log \frac{1}{n} = -\log \frac{1}{n} = \log n = h(n) \quad (3)$$

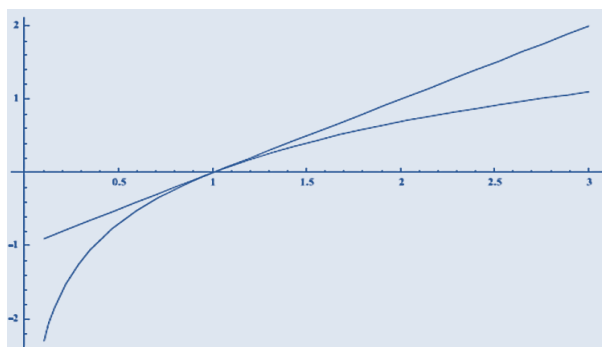
Entropija uniformne raspodele verovatnoća n mogućnosti jednaka je meri neodređenosti odgovarajućeg sistema mogućnosti bez definisanih verovatnoća mogućih ishoda. Ovaj rezultat je u tesnoj vezi sa poznatim Laplasovim principom nedovoljnog razloga, koji tvrdi da ukoliko ne znamo ništa određeno o nekoj pojavi, razumno je pretpostaviti jednake verovatnoće mogućih ishoda. U tom pogledu entropija je primenljiva kao mera neodređenosti i na one sisteme mogućnosti u kojima nam nisu poznate verovatnoće pojedinih ishoda.

Za zadati sistem mogućnosti S u kome je $|S| = n$, imamo maksimalnu neodređenost ukoliko ne poznajemo verovatnoće ishoda ili ako su one međusobno jednake, odnosno važi $H(P) \leq h(S) = \log n$. Ovaj stav možemo i formalno dokazati. Neka su $p_1, p_2, \dots, p_m$ i $q_1, q_2, \dots, q_m$ dve raspodele verovatnoća nad istim brojem mogućnosti m . Tada važi:

$$\sum_{i=1}^m p_i \cdot \log q_i \leq \sum_{i=1}^m p_i \cdot \log p_i \quad (4)$$

pri čemu jednakost važi ako i samo ako je $p_i = q_i$.

Dokaz. Budući da je $\log x = \log e \cdot \ln x$, gde $\ln x$ označava prirodni logaritam za osnovu e , i budući da je $\ln x$ konveksna funkcija, važi da je $\ln x \leq x - 1$. Ovu nejednačinu dobijamo iz činjenice da uslov konveksnosti znači da se data funkcija uvek nalazi ispod njene tangente u bilo kojoj tački i ako za tu tačku izaberemo $x = 1$, upravo ćemo dobiti gornju nejednakost, prikazanu na slici ispod.



Slika 5. Konveksnost funkcije $\ln_x$ i tangenta u tački $x=1$.

Stoga je:

$$\ln \frac{q_i}{p_i} \leq \frac{q_i}{p_i} - 1 \quad (5)$$

pa je:

$$\sum_{i=1}^m p_i \cdot \ln \frac{q_i}{p_i} \leq \sum_{i=1}^m q_i - \sum_{i=1}^m p_i = 1 - 1 = 0 \quad (6)$$

Na osnovu gornjeg zaključujemo da je:

$$\sum_{i=1}^m p_i \cdot \log q_i - \sum_{i=1}^m p_i \cdot \log p_i = \log e \sum_{i=1}^m p_i \cdot \ln \frac{q_i}{p_i} \leq 0 \quad (7)$$

Gornja nejednakost važi sa znakom jednakosti ako i samo ako je $\frac{q_i}{p_i} = 1$, za svako i odnosno ako je $p_i = q_i$. Dobijamo:

$$H(P) - \log m = -\sum_{i=1}^m p_i \cdot \log p_i + \sum_{i=1}^m p_i \cdot \log \frac{1}{m} \leq 0 \quad (8)$$

gde znak jednakosti važi ako i samo ako je $p_i = \frac{1}{m}$. Ovim smo ujedno dokazali sledeću teoremu.

$$\max H(P) = \log m \quad (9)$$

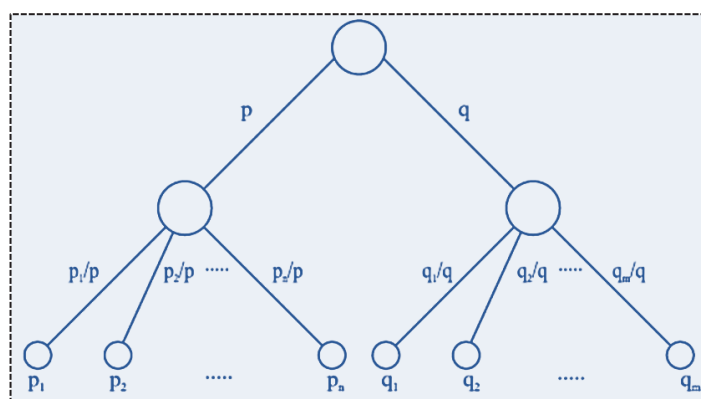
gde je maksimum tražen po svim raspodelama verovatnoća sistema sa m mogućnošću. Maksimum se dostiže samo na raspodeli koja odgovara jednakim verovatnoćama svih mogućnosti.

Ako su (S_1, P_1) i (S_2, P_2) dva verovatnosna sistema mogućnosti za koje važi $|S_1| = |S_2|$ i $P_1 = P_2$, tada je i $H(P_1) = H(P_2)$. Ovim se potvrđuje intuitivan stav da entropija zavisi samo od raspodele verovatnoće, a ne i od prirode mogućnosti e_i u nekom sistemu mogućnosti.

U važnosti je relacija:

$$H(p_1, p_2, \dots, p_m) = H(p_1, p_2, \dots, p_m, 0), \quad (10)$$

što potiče od konvencije da je $0 \cdot \log 0 = 0$. Mogućnosti čija je verovatnoća nula ne utiču na ukupnu količinu neodređenosti. To je sasvim u skladu sa našom intuicijom, pošto smo sigurni da se takve mogućnosti neće nikada realizovati.



Slika 6. Dvonivovski verovatnosni sistem mogućnosti

Posmatrajmo šemu na dva nivoa sa **Slika 6**. Na prvom nivou, biraju se dve mogućnosti sa verovatnoćama p i $q = 1 - p$. Ako se izabere prva mogućnost, na drugom nivou se bira n

mogućnosti sa verovatnoćama $\frac{p_i}{p}$. Ako se izabere druga mogućnost, tada se na drugom nivou bira m mogućnosti sa verovatnoćama $\frac{q_i}{q}$. Podrazumeva se da važi:

$$\sum_{i=1}^n p_i = p, \sum_{i=1}^n q_i = q \quad (11)$$

Primetimo da to implicira:

$$p_1 + p_2 + \dots + p_n + q_1 + q_2 + \dots + q_m = 1, \quad (12)$$

odnosno, $\{p_1 + p_2 + \dots + p_n + q_1 + q_2 + \dots + q_m\}$ je raspodela verovatnoća $n+m$ elemenata. Tada važi sledeća relacija između entropija na ova dva nivoa izbora:

$$H(p_1, p_2, \dots, p_n, q_1, q_2, \dots, q_m) = H(p, q) + p H\left(\frac{p_1}{p}, \frac{p_2}{p}, \dots, \frac{p_n}{p}\right) + q H\left(\frac{q_1}{q}, \frac{q_2}{q}, \dots, \frac{q_m}{q}\right). \quad (13)$$

3.1.1. Entropija

Verovatnosnom sistemu mogućnosti (S, P) možemo pridružiti konačnu slučajnu promenljivu X , koja uzima vrednosti e_i iz $S(e_1, e_2, \dots, e_m)$. Verovatnoća da je $X = e_i$ data sa $p_X(e_i)$, a P se naziva gustina verovatnoće slučajne promenljive X . Obrnuto, svaka konačna slučajna promenljiva indukuje jedan verovatnosni sistem mogućnosti. Stoga umesto verovatnosnog sistema mogućnosti možemo govoriti o slučajnoj veličini, a shodno tome i o entropiji slučajne veličine X sa vrednostima iz S , dato izrazom:

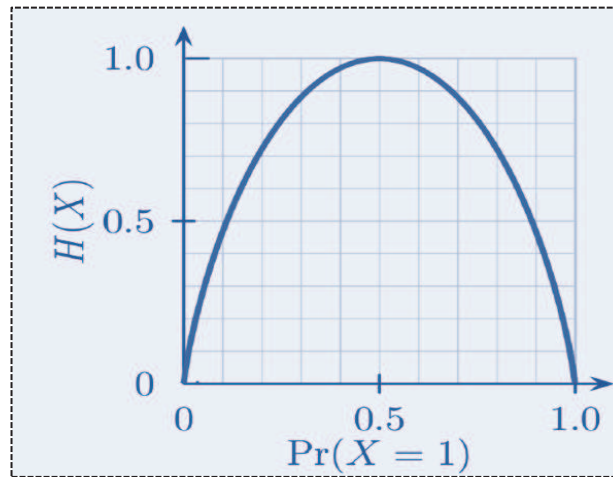
$$H(X) = -\sum_{x \in S} p_X(x) \log p_X(x).$$

Sada entropiju možemo interpretirati kao meru neodređenosti asociranu sa datom slučajnom veličinom X .

Ukoliko imamo verovatnosni sistem mogućnosti sa dva ishoda $X=1, X=0$, onda entropija slučajne veličine X zavisi od verovatnoće ova dva ishoda. Neka je verovatnoća prvog ishoda p , tada je verovatnoća drugog ishoda jednaka $1-p$. Tada je entropija data izrazom:

$$H = -p \log p - (1-p) \log(1-p). \quad (14)$$

Kada jedan od ishoda postaje verovatniji, tada entropija tog ishoda opada, **Slika 7**. Entropija ima maksimalnu vrednost 1 samo kada su verovatnoće ova dva ishoda jednake $p_{X=1, X=0} = 0,5$.



Slika 7. Entropija slučajne promenljive sa dva ishoda

Za optimalno kodovanje informacije koja se šalje kroz komunikacioni kanal, neophodno je opisati količinu informacionog sadržaja u svakoj pojedinačnoj poruci. Iz tog razloga, potrebno je znati način merenja količine informacije. Polazimo od tvrdnje da količina informacije zavisi od verovatnoće, a što je verovatnoća manja, količina informacije sadržane u poruci je veća.

Prvi zahtev je da mera koja će se koristiti za određivanje količine informacije, treba biti takva da monotonno opada sa porastom verovatnoće poruke. Drugi, da ako neko pošalje dve poruke, onda ukupna količina informacije treba biti jednaka zbiru pojedinačnih količina informacije. Funkcija koja zadovoljava oba zahteva je logaritam. Iz tog razloga, informacioni sadržaj ili količina informacije poruke X definisana je izrazom:

$$I_X = \log_2 \frac{1}{p_X} = -\log_2 p_X \quad (15)$$

gde je p_X verovatnoća poruke X .

Za opis informacionog sadržaja koristi se logaritam sa osnovom dva, jer u tom slučaju količina informacije koja je sadržana u poruci predstavlja broj bitova potrebnih za opis poruke. Kada se koristi logaritam sa osnovom dva, onda se količina informacije izražava u bitovima.

Uzajamna entropija – Bitno je razumeti kako se entropija združenog sistema mogućnosti ili višedimenzionih promenljivih odnosi prema entropiji komponentata sistema mogućnosti ili pojedinačnih promenljivih. Zatim, kako poznavanje izbora u jednoj komponenti sistema mogućnosti ili vrednosti jedne promenljive utiču na neodređenost preostalih komponenti sistema mogućnosti ili promenljivih.

Neka su zadati sistemi mogućnosti (S_1, P_2) uz koji je vezana slučajna promenljiva X i sistem (S_2, P_2) uz koji je vezana slučajna promenljiva Y .

Tada možemo definisati združeni sistem $(S_1 \times S_2, P)$, gde je P opisano združenom raspodelom verovatnoće $P_{X,Y}(x,y)$, definisanom nad parovima $(x,y) \in S_1 \times S_2$.

Tada možemo definisati tri entropije nad ova tri sistema:

$$H(X,Y) = - \sum_{x \in S_1} \sum_{y \in S_2} p_{X,Y}(x,y) \log p_{X,Y}(x,y) \quad (16)$$

$$H(X) = \sum_{x \in S_1} p_X(x) \log p_X(x) \quad (17)$$

$$H(Y) = \sum_{y \in S_2} p_Y(y) \log p_Y(y) \quad (18)$$

U kakvom su odnosu ove tri entropije? Za bilo koji par slučajnih veličina X i Y važi:

$$H(X,Y) \leq H(X) + H(Y) \quad (19)$$

Znak jednakosti važi ako i samo ako su X i Y nezavisne slučajne veličine, tj.:

$$p_{X,Y}(x,y) = p_X(x) \cdot p_Y(y) \quad (20)$$

Odavde vidimo da je neodređenost združenog sistema uvek manja od sume neodređenosti pojedinačnih sistema. Razlog potiče od međusobne zavisnosti (korelisanosti) X i Y , koja kao da rezultuje u zajedničkom udelu neodređenosti u obe promenljive. U slučaju nezavisnih X i Y , ove zajedničke neodređenosti nema, i važan je znak jednakosti.

Ova tvrdnja se može generalizovati na više od dve promenljive. U opštem slučaju, ako imamo slučajni vektor $X = (x_1, x_2, \dots, x_m)$, čija je raspodela verovatnoća data sa $p_X = (X)$, združena entropija entropija je data sa:

$$H(X) = \sum_x p_X(x) \log p_X(x) \quad (21)$$

Sve marginalne raspodele pojedinih komponenti vektora X su date sa:

$$p_X(x_i) = \sum_{x_1, \dots, x_{i-1}, \dots, x_{i+1}, \dots, x_m} p_X(x_1, x_2, \dots, x_{i-1}, \dots, x_{i+1}, \dots, x_m) \quad (22)$$

Slučajne veličine $x_1, x_2, \dots, x_m$ su uzajamno nezavisne, ako i samo ako važi:

$$p_X(x) = p_{X_1}(x_1) \cdot p_{X_2}(x_2) \dots p_{X_m}(x_m) \quad (23)$$

Za bilo koji slučajnu višedimenzionu promenljivu (slučajni vektor) $X = (x_1, x_2, \dots, x_m)$ važi:

$$H(X) \leq \sum_{i=1}^m H(X_i) \quad (24)$$

Jednakost važi samo u slučaju uzajamne nezavisnosti slučajnih varijabli $x_1, x_2, \dots, x_m$.

Uslovne entropije - Posmatrajmo slučajne veličine X i Y . Neka je opservirana realizacija $Y = y$. Kako ovaj događaj utiče na neodređenost veličine X ?

Opserviranjem $Y=y$ menja se raspodela $p_X(x)$ u uslovnu raspodelu:

$$p_{X|y}(x, y) = p_X(x) \quad (25)$$

$$H(X|y) = -\sum_x p_X(x) \log p_X(x) = H(X) \quad (26)$$

Odnosno, kada su X i Y nezavisne, opservacija promenljive Y ne menja neodređenost promenljive X .

Pored ove uslovne entropije možemo definisati očekivanu uslovnu entropiju:

$$H(X|Y) = \sum_y p_Y(y) H(X|y) \quad (27)$$

Što je u stvari očekivana vrednost $H(X|y)$ u odnosu na y . Do sada smo se susreli sa tri tipa uslovnih entropija:

$H(X|y)$ – Uslovna entropija promenljive X u odnosu na događaj ili opservaciju $Y = y$.

$H(X|E)$ – Uslovna entropija promenljive X u odnosu na događaj E .

$H(X|Y)$ – Očekivana vrednost uslovne entropije.

Za par slučajnih promenljivih X i Y uvek važi:

$$H(X, Y) = H(Y) + H(X|Y) \quad (28)$$

ili

$$H(X, Y) = H(X) + H(Y|X) \quad (29)$$

Za svaki par slučajnih veličina X i Y važi:

$$H(X|Y) \leq H(X) \quad (30)$$

Jednakost važi ako i samo ako su X i Y nezavisne. Primenom pravila lančanja, dobijamo:

$$H(X|Y) = H(X, Y) - H(Y) \leq H(X) + H(Y) - H(Y) = H(X) \quad (31)$$

pri čemu jednakost važi ako i samo ako su X i Y nezavisne.

Relativne entropije - Posmatrajmo dva verovatnosna sistema mogućnosti nad istim skupom mogućnosti S . Ako su X i Y korespondentne slučajne promenljive, tada se veličina:

$$K(P_X, P_Y) = \sum_{x \in S} p_X(x) \log \frac{p_X(x)}{p_Y(x)} \quad (32)$$

naziva relativna entropija ili Kulbak Lejblerova divergencija (Kulbak-Lejblerovo rastojanje) između raspodele dve slučajne veličine X i Y . Iako se u jednom od naziva koristi rastojanje, primetimo da ne važi uslov simetričnosti, naime:

$$d(P_X, P_Y) \neq d(P_Y, P_X). \quad (33)$$

Divergencija je jednaka nuli:

$$K(P_X, P_Y) = K(P_Y, P_X) = 0, \quad (34)$$

Ako i samo ako su:

$$P_X = P_Y. \quad (35)$$

Pošto se usvaja konvencija da je $0 \log(0/0) = 0$, $0 \log(0/q) = 0$, $p \log(p/0) = \infty$, $K(P_X, P_Y) = \infty$, ukoliko postoji x za koje je $p_X(x) > 0$, a $p_Y(x) = 0$.

Posmatrajmo združeni verovatnosni sistem mogućnosti $(S \times S, P)$ kome je asociran par slučajnih veličina $X, Y \in S$. Marginalni sistemi (S, P_X) i (S, P_Y) su indukovani marginalnim raspodelama verovatnoća:

$$\begin{aligned} p_X(x) &= \sum_y p_{X,Y}(x, y), \\ p_Y(y) &= \sum_x p_{X,Y}(x, y). \end{aligned} \quad (36)$$

Označimo sa $P_X \cdot P_Y$ raspodelu verovatnoća $p_X(x) \cdot p_Y(y)$.

Tada važi:

$$I(X; Y) = I(Y; X) = K(P, P_X \cdot P_Y). \quad (37)$$

Relativna entropija definisana je izrazom:

$$I(X; Y) = \sum_{x,y} p_{X,Y}(x,y) \log \frac{p_{X,Y}(x,y)}{p_X(x)p_Y(y)} \quad (38)$$

Na osnovu ovih uvida nam je jasno da uzajamna informacija $I(X; Y)$ meri stepen divergencije združene raspodele verovatnoće para X, Y od raspodele verovatnoće ovog para u slučaju njihove nezavisnosti. Stoga se često sa opravdanjem uzajamna informacija uzima kao mera međuzavisnosti dve slučajne veličine koje uzimaju vrednosti nad istim skupom.

3.1.2. Informacija i njena mera

Opservacija slučajne veličine ili događaji vezani za jednu slučajnu veličinu predstavljaju informaciju. Količina informacije dobijena na osnovu opservacija slučajne veličine ili događaja se meri na osnovu rezultujuće promene neodređenosti. Stoga zaključujemo da su entropija i mera informacije tesno povezani.

Šta je informacija i kako se ona meri?

Osnovna ideja je da je ona vezana za promenu neodređenosti. Stoga, mera informacije mora biti srazmerna količini promene neodređenosti. Počnimo naša razmatranja od sistema mogućnosti (S, P) , reprezentovanog slučajnom promenljivom X , koja uzima vrednosti $x \in S$, sa verovatnoćom $p_X(x)$. Ova slučajna promenljiva opisuje jedan eksperiment u kome su ishodi neodređeni. Mera neodređenosti u vezi sa ovim eksperimentom je data entropijom:

$$H(X) = - \sum_{x \in S} p_X(x) \log p_X(x) \quad (39)$$

Kada se izvrši eksperiment, mi opserviramo određenu vrednost $x \in S$. Nakon toga više nema neodređenosti. Prethodna neodređenost $H(X)$ redukovana na aposteriornu neodređenost 0. Razlika $H(X) - 0 = H(X)$ jednaka je količini informacije dobijene izvođenjem ovog eksperimenta. Stoga je entropija jedne slučajne veličine jednaka količini informacije dobijene opserviranjem jedne konkretne njene realizacije.

Ova ideja nam nameće dva važna zapažanja:

- Budući da je informacija jednaka promeni entropije, ona se meri u istim jedinicama kao i entropija, dakle ako je u 2 osnova logaritma, jedinica je jedan bit.
- Dobijena količina informacije je ista za sve moguće realizacije slučajne veličine X , bez obzira da li je verovatnoća neke realizacije velika ili mala.

Binarna slučajna promenljiva

Neka je zadata binarna slučajna veličina X , koja prima vrednosti 0 i 1 sa verovatnoćama p i $q = 1 - p$, respektivno. Opserviranjem jednog ishoda binarnog eksperimenta rezultuje u dobijanju količine informacije jednake:

$$H(X) = -p \log p - q \log q \quad (40)$$

Na primer, u fer bacanju novčića, opserviranjem ishoda, bacanjem se dobija količina informacije od 1 bita.

Generalizujemo sada ovo razmatranje. I dalje ćemo posmatrati slučajnu promenljivu X vezanu za sistem mogućnosti (S, P) .

Količina neodređenosti je i dalje $H(X)$. Sada ćemo obaviti eksperiment samo parcijalno. Nećemo opservirati direktno slučajnu veličinu X , već samo neki događaj $E \subseteq S$. I ovo predstavlja neku informaciju. Šta je sada njena mera?

Opservacija događaja E menja slučajnu veličinu X u uslovnu promenljivu $X|E$, koja se odnosi na novi verovatnosni sistem mogućnosti (E, P_E) . Ovde P_E označava uslovne verovatnoće.

$$p_{X|E} \frac{p_{X,Y}(x)}{p_X(E)}, x \in E \quad (41)$$

Ova nova situacija, nastala opservacijom događaja E , ima neodređenost jednaku uslovnoj entropiji:

$$H(X|Y) = -\sum_{x \in E} p_{X|E}(X) \log p_{X|E}(x) \quad (42)$$

Prema tome opserviranje događaja E , menja neodređenost sa početne $H(X)$ u $H(X|E)$. Količina informacije dobijena tom prilikom je jednaka $H(X) - H(X|E)$. Videćemo da ovo ne mora biti u opštem slučaju dobitak, budući da $H(X|E)$ može biti i veća od $H(X)$, tako da opserviranje događaja E može da poveća neodređenost, što vodi definiciji negativne informacije.

Kada se E svodi na opservaciju tačne vrednosti jedne od mogućnosti X , tada je $H(X|E) = 0$, pa je:

$$H(X) - H(X|X) = H(X) - 0 = H(X) \quad (43)$$

količina informacije dobijena tom prilikom jednaka entropiji.

Zato se često entropija neke slučajne veličine naziva i informacija o samoj sebi (eng. *self-information*). Odnosno, entropija jedne slučajne veličine je u srednjem jednaka količini informacije potrebne za opis te slučajne veličine. Da bi smo nastavili sa razvojem koncepta informacije, nužno je da uvedemo dva nova pojma: relativnu entropiju i uzajamnu informaciju.

Posmatrajmo združeni verovatnosni sistem mogućnosti $(S_1 \times S_2)$ i pridružene slučajne veličine X i Y . Pretpostavimo da smo opservirali jednu vrednost $Y = y \in S_2$. Kao što smo videli ova situacija menja neodređenost u pogledu promenljive Y na nulu, odakle sledi da ova opservacija sadrži:

$$i(y; Y) = H(Y) \quad (44)$$

bita informacije, relativno u odnosu na prethodna znanja o Y . Ovu veličinu ćemo zvati lokalna uzajamna informacija, odnosno koliko jedna realizacija slučajne veličine y govori o slučajnoj veličini Y u celini.

Medjutim, opservacija $Y = y$, menja i raspodelu verovatnoće slučajne veličine X , tako što ona sada postaje uslovna raspodela:

$$p_{X|Y}(x, y) \frac{p_{X,Y}(x, y)}{p_Y(y)}, x \in S_1 \quad (45)$$

Ovo znači da se takodje menja i neodređenost u pogledu slučajne promenljive X . Stoga opservacija $Y = y$, sadrži informaciju u odnosu na X . Količina ove informacije iznosi:

$$i(y; X) = H(X) - H(X | y) \quad (46)$$

Ovu veličinu ćemo zvati lokalna uzajamna informacija slučajne veličine Y i X , a govori o tome koliko jedna opservacija $Y = y$ daje informacije o drugoj slučajnoj veličini X .

Predimo sada sa pojedinačnih opservacija na očekivane vrednosti lokalnih uzajamnih informacija.

$$I(X; Y) = \sum_y p_Y(y) i(y; X) = \sum_y p_Y(y) (H(X) - H(X | y)) = H(X) - H(X | Y) \quad (47)$$

Veličina $I(X; Y)$ se naziva uzajamna informacija između slučajnih veličina X i Y . Ona predstavlja jedan od centralnih pojmova teorije informacija. Striktno govoreći, ona nije informacija, već očekivana vrednost količine informacije koju dobijamo o X kada opserviramo vrednosti Y .

Na osnovu Korolarija koji tvrdi da je uvek:

$$H(X | Y) \leq H(X) \quad (48)$$

Teorema:

$$I(X; Y) \geq 0 \quad (49)$$

dok je $I(X; Y) = 0$, ako i samo ako su X i Y međusobno nezavisne slučajne veličine. Podsetimo se da lokalna uzajamna informacija $i(y; X)$ može biti i negativna, dok je

uzajamna informacija kao njeno matematičko očekivanje uvek pozitivna. Na osnovu svojstva $H(X|Y) = H(X, Y) - H(Y)$, dobijamo da važi:

$$I(X; Y) = H(X) + H(Y) - H(X, Y) \quad (50)$$

Iz simetrije izraza iznad sledi sledeća teorema:

$$I(X; Y) = I(Y; X) \quad (51)$$

Sumiraćemo sada najznačajnija svojstva uzajamne informacije i njen odnos prema entropiji:

$$\begin{aligned} I(X; Y) &= H(X) - H(X|Y) \\ I(X; Y) &= H(Y) - H(Y|X) \\ I(X : Y) &= H(X) + H(Y) - H(X, Y) \\ I(X; Y) &= I(Y; X) \\ I(X; X) &= H(X) \end{aligned} \quad (52)$$

3.1.3. Grafička prezentacija informaciono-teorijskih mera

Slučajne promenljive o kojim govorimo su diskretne. Za dve slučajne promenljive X i Y , Šenonove informacione mere su: $H(X)$, $H(Y)$, $H(X|Y)$, i $I(X; Y)$. U radu [35] napravljena je nova prezentacija informaciono-teorijskih mera. Pogledajmo $\tilde{X}$ i $\tilde{Y}$ promenljive, koje odgovaraju promenljivoj X i Y , respektivno. $\tilde{X} \cup \tilde{Y}$ biće univerzalni set Ω , autor je u radu definisao stvarnu meru μ^* nad σ - poljem:

$$F\{(\tilde{X} \cup \tilde{Y}), \tilde{X}, \tilde{Y}, (\tilde{X} \cap \tilde{Y}), (\tilde{X} \cap \tilde{Y}^c), (\tilde{X}^c \cap \tilde{Y}), (\tilde{X} \cap \tilde{Y})^c, \emptyset\} \quad (53)$$

za:

$$\begin{aligned} \mu^*(\tilde{X} \cup \tilde{Y}) &= H(X, Y) \\ \mu^*(\tilde{X}) &= H(X) \\ \mu^*(\tilde{Y}) &= H(Y) \\ \mu^*(\tilde{X} \cap \tilde{Y}) &= I(X; Y) \\ \mu^*(\tilde{X} - \tilde{Y}) &= H(X|Y) \\ \mu^*(\tilde{Y} - \tilde{X}) &= H(Y|X) \\ \mu^*((\tilde{Y} - \tilde{X})^c) &= H(X|Y) + H(Y|X) \end{aligned} \quad (54)$$

i:

$$\mu^*(\phi) = 0 \quad (55)$$

Primećujemo da prvih šest jednakosti sa leve strane u izrazu (54), možemo dobiti sa izrazima sa desne strane sa kojima su predstavljene Šenonove informacione mere za X i Y , urađivši zamenu simbola:

$$\begin{aligned} H / I &\rightarrow \mu \\ , &\rightarrow \cup \\ ; &\rightarrow \cap \\ | &\rightarrow -. \end{aligned} \quad (56)$$

Ovo je formalna zamena simbola. Ne postoji značajna razlika između entropije i zajedničke informacije. Entropija se nekad odnosi na samu informaciju. Tako da za dve slučajne promenljive X i Y možemo koristiti Šenonove informacione mere na polju F . Mera μ^* je isto kao I - mera za dve slučajne promenljive X i Y . Lako je uvideti da za bilo koji teorijski identitet informacije X i Y , možemo dobiti odgovarajući teorijski identitet formalnom zamenom simbola.

Na primer, za informaciono teorijski identitet:

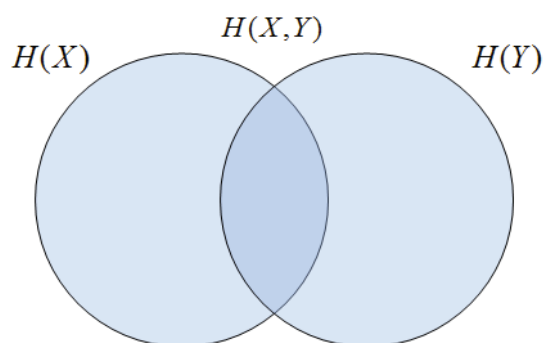
$$H(X, Y) \equiv H(X) + H(Y) - I(X; Y), \quad (57)$$

Sledi set teorijskog identiteta:

$$\mu(\tilde{X} \cup \tilde{Y}) \equiv \mu(\tilde{X}) + \mu(\tilde{Y}) - \mu(\tilde{X} \cap \tilde{Y}). \quad (58)$$

U nastavku sledi grafička prezentacija Šenonovih mera, preuzeta iz rada [35].

1. Zajednička entropija dve diskretne slučajne promenljive X i Y , ili njihova entropija nakon udruživanja definisana je sa $H(X, Y)$. Ovo znači da ako su promenljive nezavisne, njihova entropija biće zbir njihovih pojedinačnih entropija $H(X) + H(Y) = H(X, Y)$, prikazano na, (Slika 8).



Slika 8. Dve slučajne promenljive

Pod uslovom da je:

$$H(X, Y) \leq H(X) + H(Y) \quad (59)$$

i:

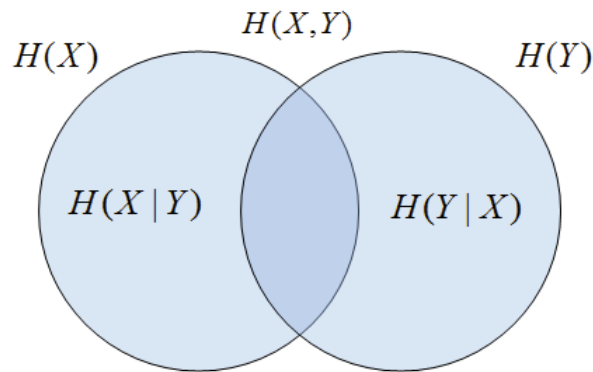
$$0 \leq H(X) \leq \log 2n \quad (60)$$

Gde n predstavlja broj mogućih vrednosti X .

2. Uslovna entropija ili uslovna nezavisnost promenljive X kada poznajemo promenljivu Y (često se naziva ekvivokacija) je prosečna uslovna entropija nad Y . Osnovna osobina ovog oblika entropije je:

$$H(X | Y) = H(X, Y) - H(Y). \quad (61)$$

Sledi grafička prezentacija ove veličine, **Slika 9**.

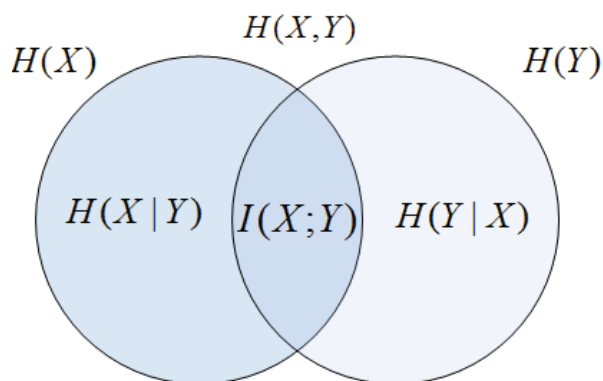


Slika 9. Uslovna entropija

Pod uslovom da je:

$$\begin{aligned} H(X | Y) &\leq H(X), \\ H(Y | X) &\leq H(Y). \end{aligned} \quad (62)$$

3. Uzajamna informacija meri količinu informacije koju je moguće dobiti o jednoj promenljivoj, kada je opservirala ili kada je poznata druga. Ova veličina je važna u komunikacijama, gde se može koristiti u cilju povećavanja razmenjene količine informacije između poslatih i primljenih signala. Sledi grafička prezentacija ove veličine, **Slika 10**.



Slika 10. Uzajamna informacija

3.1.4. Šenonova matematička teorija komunikacije

Klod Šenon (*Claude E. Shannon*) je u radu [33] formulisao teoriju za kompresiju podataka. Šenon je utvrdio da postoji fundamentalni limit za kompresiju podataka bez gubitaka. Ovo ograničenje se naziva stopa entropije i označava sa H . Tačna vrednost entropije zavisi od izvora informacija, tačnije, od statističke prirode izvora. Šenon tvrdi da je moguće kompresovati poruke na izvoru bez gubitaka, uz stepen kompresije blizu H i da je matematički nemoguće uraditi bolje od H .

Koncept ima mnogo primena u teoriji informacija i komunikacija, a Šenon ga je primenio i da bi izrazio količinu aktuelnih informacija u engleskom tekstu, rečima: "Entropija je statistički parametar koji meri u određenom smislu, koliko se mnogo informacija proizvodi u proseku za svako slovo teksta u jeziku. Ako se jezik prevede u binarne digitalne znakove-bitove (0 ili 1) u najefikasnijem slučaju, entropija H je srednji broj bita koji se zahteva po slovu originalnog jezika."

Takođe, Šenon je razvio teoriju o gubicima prilikom kompresije podataka. Ovo je više poznato kao stopa distorzije u teoriji. U kompresiji sa gubicima, dekompresovani podaci ne moraju biti isti kao originalni podaci. Umesto toga, neka vrednost distorzije D , može da se toleriše. Šenon je pokazao da za određeni izvor sa poznatim statističkim osobinama i datom merom distorzije, postoji funkcija $R(D)$ koja se naziva funkcija za stopu distorzije. Teorija kaže da ako je D tolerantna količina distorzije, onda je $R(D)$ najbolji stepen kompresije.

Kada se radi o kompresiji bez gubitaka (tj. bez distorzije ili $D = 0$), najbolji mogući stepen kompresije je $R(0) = H$ (odnosi se na alfabet, kao konačni izvor poruka). Drugim rečima, najbolja moguća stopa kompresije bez gubitaka je stopa entropije H .

Teorija o kompresiji podataka bez gubitaka i stopa distorzije zajedno čine teoriju o kodovanju informacionog izvora. Teorija o kodovanju izvora postavlja fundamentalna ograničenja na performanse svih algoritama za kompresiju podataka. Teorija, sama po sebi ne definiše kako de se dizajniraju i sprovedu u delo ovi algoritmi. On međutim, daje neke savete i smernice za postizanje optimalnih performansi. U nastavku poglavlja analiziraćemo kako je Šenon

modelovao izvor informacija u pogledu slučajnog procesa i njegovu teoremu za kodovanje izvora bez gubitaka.

Modelovanje izvora

Zamislamo da odemo u biblioteku "Amazon". Ta biblioteka ima veliki izbor knjiga, negde oko 50 miliona knjiga. Svaka knjiga ima u proseku 50 miliona znakova (slova). Izabraćemo jednu knjigu na slučajan način. U tom momentu ta knjiga je kompresovan informacioni izvor u nekom ZIP formatu, da bismo knjigu preneli do kuće na disku ili preko Interneta. Matematički, knjigu koju izaberemo označavano na sledeći način:

$$X = \{X_1, X_2, X_3, \dots\}. \quad (63)$$

gde X predstavlja celu knjigu, prvi znak u knjizi je X_1 , drugi znak u knjizi je X_2 i tako redom. Sada ćemo da definišemo konačan skup simbola ili poruka A (alfabet) koje generiše knjiga kao informacioni izvor.

$$A = \{a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p, q, r, s, t, u, w, x, y, z, \text{razmak}\}$$

Stavićemo sada nas u ulogu nekog inženjera koji treba da dizajnira algoritam za kompresiju izabrane knjige. Iz naše perspektive, karakteri u knjizi su slučajne promenljive koje uzimaju vrednosti iz alfabeta A . Cela knjiga je beskonačan niz slučajnih promenljivih (slučajan proces). Postoji nekoliko načina kako možemo modelovati statističke osobine knjige.

Šenonovi modeli aproksimacije

Model aproksimacije nultog reda - Svaki karakter je statistički nezavistan od svih drugih karaktera i verovatnoće svih 27 karaktera su jednake. Ako je ovaj model tačan, onda bi niz karaktera iz ove knjige izgledao ovako (svi primeri su preuzeti iz rada [33]) :

xfoml rxkhrjffjuj zlpwcfwkcyj ffjeyvkcsghyd qpaamkbzaacibzlhjqd

Dobijeni tekst nije proizvod jednog inteligentnog bića. Rezultat više deluje kao posledica nasumičnog udaranja po tastaturi.

Model aproksimacije prvog reda - poznato je da se u prirodnim jezicima neka slova javljaju češće od drugih. Na primer, u engleskom jeziku slova „a“ i „e“ su češći nego „k“ i „z“. Dakle, u ovom modelu, karakteri su i dalje nezavisni jedan od drugog, ali postoji verovatnoća distribucije karaktera koja odgovara jednom jeziku. Tipičan tekst za ovaj model izgleda ovako:

ocroh hli rgwr nmiewis eu ll nbnesebya th eei alhenhttpa oobttva nah brl

Model aproksimacije drugog reda - kod prethodna dva modela pretpostavlja se statistička nezavisnost od jednog do sledećeg karaktera. Distribucija karaktera ne odražava u potpunosti strukturu jezika. Sa druge strane, postoje karakteri koji se češće nalaze uz neke karaktere. Ovo znači da postoji zavisnost između karaktera. U ovom modelu, sadašnji karakter zavisi od prethodnog karaktera, ali je uslovno nezavisan od svih prethodnih karaktera. Na primer, slovo "u" se retko javlja (verovatnoća = 0.022). Međutim, obzirom da je prethodni karakter "q", verovatnoća karaktera "a" je sada mnogo veća (verovatnoća = 0.995). Tipičan primer za ovaj model će izgledati ovako:

**on ie antsoutinys are t inctore st be s deamy achin d ilonasive tucoowe at teasonare fuso
tizin andy tobe seace ctisbe**

Model aproksimacije trećeg reda - treći model je nastavak drugog modela aproksimacije. Ovde karakter zavisi od prethodna dva karaktera, ali je uslovno nezavisan od svih prethodnih karaktera. U ovom modelu, distribucija varira od toga šta predstavlja. Tipičan tekst za ovaj model će izgledati ovako:

**in no ist lat whey cratiet froure birs grocid pondenome of demonstures of the reptagin is
regoactiona of cre**

Sada je sličnost sa obični engleskim tekstom poprilično uočljiva na svakom od navedenih koraka.

Opšti model aproksimacije - U ovom modelu, knjiga je proizvoljni stacionarni slučajni proces. Statističke osobine ovog modela su previše kompleksne da bi se razmatralo praktično. Ovaj model je interesantan samo sa teorijskog stanovišta.

Entropija izvora u odnosu na aproksimaciju modela

Entropija izvora je broj koji zavisi samo od statističke prirode izvora. Ako izvor ima jednostavan model, onda se ovaj broj može lako izračunati. Sledi jedan proizvoljan izvor:

$$X = \{X_1, X_2, X_3, \dots\}, \quad (64)$$

čije će se vrednosti odnositi na engleski jezik.

Model aproksimacije nultog reda - karakteri su statistički nezavisni, svako slovo alfabeta podjednako verovatno može doći u obzir. Neka m bude veličina alfabeta. U ovom slučaju, entropija izvor je određena sa izrazom:

$$H = \log 2m, \quad (65)$$

bitova po karakteru. Za engleski tekst, veličina alfabeta $m = 27$. Dakle, ako je ovo bio precizan model za engleski jezik, onda entropija izvora je $H = 4.75$ bitova po karakteru.

Model aproksimacije prvog reda - karakteri su statistički nezavisni, ukupan broj karaktera odgovara veličini alfabeta i verovatnoća svakog i-tog karaktera je poznata. Entropija izvora u tom slučaju je:

$$H = \sum_{i=1}^m p_i \log_2 p_i \quad (66)$$

Koristeći aproksimaciju prvog reda, entropija engleskog teksta je 4.07 bitova po karakteru, ako bi ovo bio odgovarajući model izvora.

Model aproksimacije drugog reda - $P_{j|i}$ neka bude uslovna verovatnoća sadašnjeg karaktera j , budući da je poznat prethodni karakter i . Entropija je:

$$H = \sum_{i=1}^m P_i \sum_{j=1}^m P_{j|i} \log_2 P_{j|i}, \quad (67)$$

Koristeći aproksimaciju modela drugog reda, entropija engleskog teksta će biti 3.36 bita po karakteru, pod uslovom da je model ispravan.

Model aproksimacije trećeg reda - $P_{k|i,j}$ neka bude uslovna verovatnoća sadašnjeg karaktera k , budući da su poznata prethodna dva karaktera i, j . Entropija je:

$$H = \sum_{i=1}^m P_i \sum_{j=1}^m P_{j|i} \sum_{k=1}^m P_{k|i,j} \log_2 P_{k|i,j}, \quad (68)$$

Opšti model aproksimacije - B_n predstavlja prvi n karakter. Entropija u opštem slučaju data:

$$H = \lim_{n \rightarrow \infty} \frac{1}{n} \sum p(B_n) \log_2 p(B_n) \quad (69)$$

gde je suma veća od m^n mogućih vrednosti B_n . Praktično je nemoguće izračunati entropiju sa gore navedenom funkcijom. Koristeći metod intuitivnog, Šenon je bio u stanju da proceni entropiju za 27 karaktera engleskog alfabeta, gde je entropija po jednom karakteru 2.3 bita. Više detalja o ovome, moguće je pronaći u Šenonovom zborniku [36].

Šenonova teorema za kodovanje bez gubitaka

Šenonova teorema za kodovanje izvora zasniva se na konceptu blokovskog kodovanja. Da bi ilustrovali ovaj koncept, uvodimo poseban izvor informacija u kojem se alfabet sastoji samo od dva slova:

$$A = \{a, b\} \quad (70)$$

Blok kod za aproksimaciju nultog reda – u ovom modelu svaki karakter je mapiran u jedan bit.

B_1	$p(B_1)$	Kodna reč
a	0.5	0
b	0.5	1
R=1 bit		

Primer:

Originalni podaci	a	a	a	a	a	a	a	a	a	a	b	b	b	b	b	a	a	a	a	a
Kompresovani podaci	0	0	0	1	1	0	1	0	0	0	1	1	1	1	1	1	0	0	0	0

U ovom slučaju koristili smo 20 bitova za predstavljanje 20 karaktera. Prosečni broj bitova po jednom karakteru je 1 bit.

Blok kod aproksimacija drugog reda – u ovom modelu svaki karakter može da bude predstavljen sa jednim, dva ili tri bita.

B_2	$p(B_2)$	Kodna reč
aa	0.45	0
bb	0.45	10
ab	0.05	110
ba	0.05	111
R=0.71 bitova		

Primer:

Originalni podaci	aa	aa	aa	aa	aa	bb	bb	ba	aa	aa
Kompresovani	0	0	0	0	0	10	10	111	0	0

podaci									
--------	--	--	--	--	--	--	--	--	--

U ovom slučaju koristili smo 14 bitova za predstavljanje 20 karaktera. Prosečni broj bitova po jednom karakteru je 0.7 bita.

Blok kod aproksimacija trećeg reda – po tri karaktera su predstavljena binarnom sekvencom dužine od jednog do šest bitova.

B_3	$p(B_3)$	Kodna reč
aaa	0.405	0
bbb	0.405	10
aab	0.045	1100
abb	0.045	1101
bba	0.045	1110
baa	0.045	11110
aba	0.005	111110
bab	0.005	111111
R= 0.55 bitova		

Primer:

Originalni podaci	aaa	aaa	aaa	abb	bbb	aaa
Kompresovani podaci	0	0	0	1101	10	0

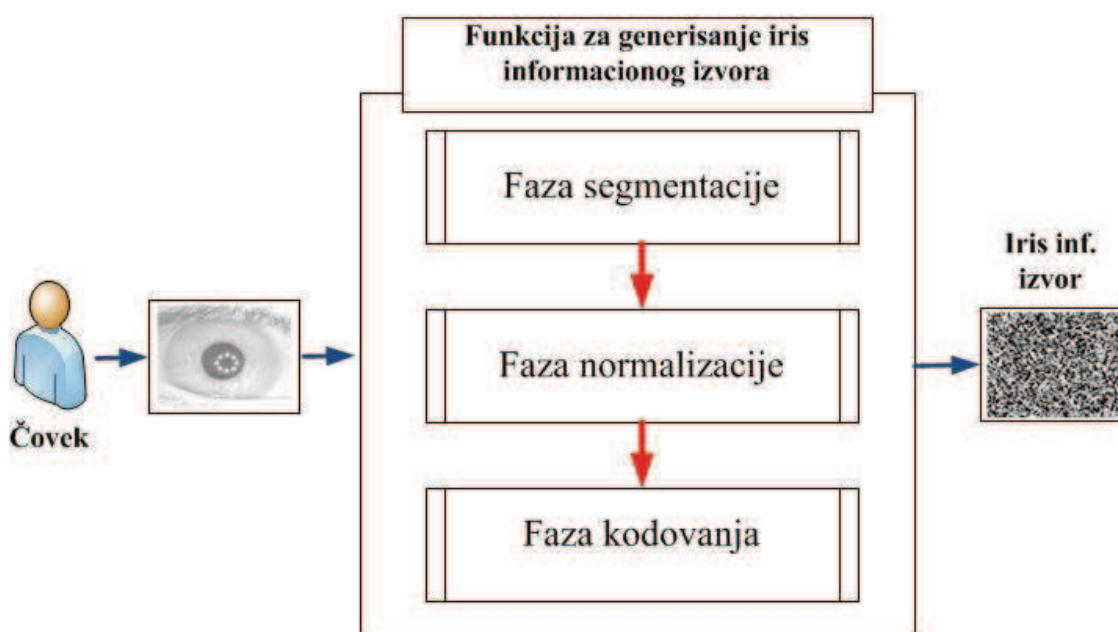
U ovom slučaju koristili smo 10 bitova za predstavljanje 18 karaktera. Prosečni broj bitova po jednom karakteru je 0.55 bitova.

Zahvaljujući Šenonovom viđenju matematičke teorije komunikacija, ovu metodologiju primenićemo u teorijskoj informacionoj analizi iris biometrijskog izvora. Cilj date analize u IV poglavlju ove disertacije je da primenom adekvatne aproksimacije odredi model iris

informacionog izvora, čime će biti postavljeni okviri za sintezu sistema za ekstrakciju čisto slučajnih nizova iz iris biometrije, koji će se koristiti za kriptografske potrebe.

3.2. Iris kao biometrijski informacioni izvor

U ovom delu rada proći ćemo kroz neophodne faze za stvaranje iris biometrijskog izvora. Ovaj proces stvaranja biometrijskog izvora počinje sa slikom oka i završava sa biometrijskim digitalnim obrascem koji predstavlja iris informacioni izvor. Informacija dobijena iz ovog informacionog izvora može biti korišćena u sistemima za prepoznavanje pojedinaca na osnovu nekih jedinstvenih karakteristika ili korišćena za ekstrakciju kriptoloških ključeva koji će biti čvrsto vezani za identitet jedne osobe. **Slika 11.**, predstavlja šemu za generisanje iris informacionog izvora.



Slika 11. Šema za generisanje iris inf. izvora

3.2.1. Segmentacija

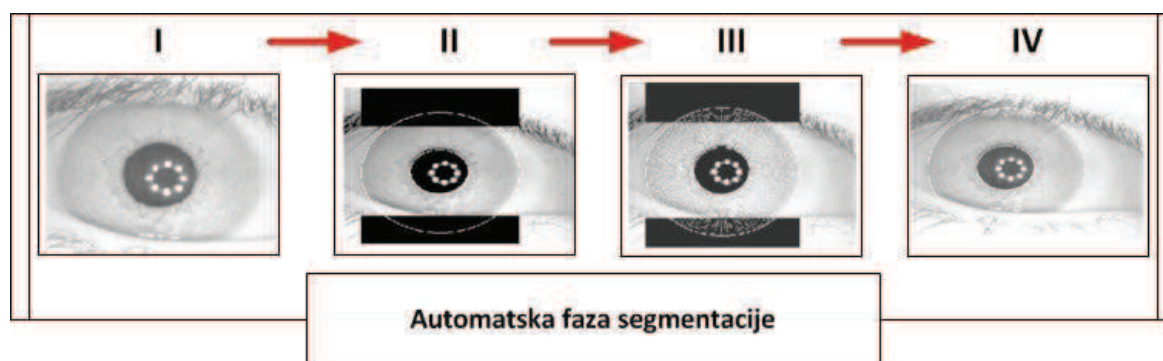
Faza segmentacije je automatski sistem koji se zasniva na Haf (Hough, 1962.) detekciji i sposoban je da lokalizuje kružni oblik irisa i zenicu, iako su tu smetnje koje stvaraju kapci, trepavice i refleksija. Iris region posmatramo između dva kruga, spoljni krug do beonjače i unutrašnji krug do zenice. Spoljašnji krug uglavnom je jednim delom sa gornje i donje strane prekriven kopcima i trepavicama. Iz tog razloga je potrebna takva tehnika koja će da izoluje ili izuzme ovakve objekte i prepozna kružni deo irisa.

Klasična tehnika Haf detekcije je korisna ako se malo zna o položaju krive koja čini granicu, ali se pretpostavlja da se može opisati parametarskom jednačinom. U slučaju prave $y = mx + c$, formira se prostor vrednosti parametara m i c , akumulatorska matrica $A(c; m)$ se inicijalizuje nulama, a onda se za svaku tačku gradijenta slike, čiji intenzitet prelazi neki

prag uradi inkrementacija $A(c, m) := A(c, m) + 1$ za m i c tako da je $c = -mx + y$ dužina te linije.

Uspešnost faze segmentacije prvenstveno zavisi od kvaliteta slike oka. Slike oka koje se koriste u fazi segmentacije su visokog kvaliteta jer ne sadrže refleksije zbog upotrebe infracrvene svetlosti za rasvetljavanje irisa. Problemi se javljaju kod slika osoba sa tamnim pigmentom zbog postojanja manjeg kontrasta što fazu segmentacije čini težom i rizičnijom. Faza segmentacije je po redosledu prva i iz tog razloga kritična za uspeh kompletnog procesa za formiranje iris informacionog izvora. Loš rezultat postignut na početku, dalje bi se negativno odrazio na kompletan sistem za prepoznavanje identiteta osobe ili sistem za generisanje kriptoloških ključeva na osnovu biometrije irisa. Drugim rečima, parametar FAR ili stopa lažnog odbacivanja bila bi veća.

Za kompletan proces formiranja iris informacionog izvora (generisanje biometrijskog obrazaca) u fazi segmentacije koristimo Haf automatsku detekciju. Haf detekcija predstavlja algoritam koji je dosta primenljiv u kompjuterskoj grafici za prepoznavanje jednostavnih geometrijskih oblika kao što su linije i krugovi. Konkretno, ovde se koristi kružna Haf detekcija za automatsku segmentaciju irisa, razvijena i predložena od autora [37].



Slika 12. I- Slika oka , II- detekcija kapaka i trepavica, III-označavanje iris regiona, IV- segmentiran iris

Model automatske segmentacije pokazao se kao uspešan (**Slika 12**). Eksperimentalna baza podataka „CASIA“ koja je korišćena u istraživanju u ovoj disertaciji obezbeđuje dobru segmentaciju. Stepenn uspešnosti faze segmentacije nad „CASIA“ biometrijskom bazom podataka bio je veći od 85%.

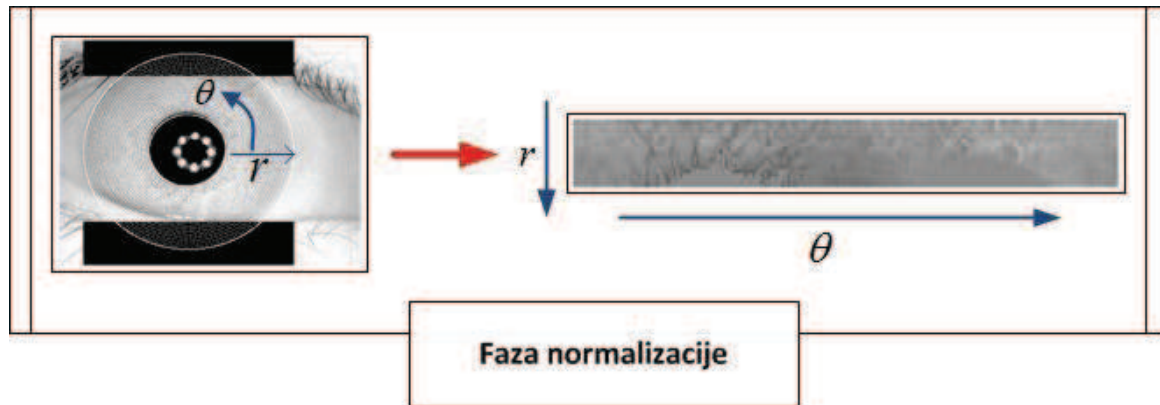
3.2.2. Noramalizacija

Nakon faze segmentacije, sledi faza normalizacije. Faza normalizacije predstavlja transformaciju segmentiranog iris regiona u zadate fiksne dimenzije pravougaonika, koje su određene unapred prema sledećoj fazi. Često nastanu manje varijacije kod segmentiranog iris regiona, zbog istezanja zenice koje nastaje promenom nivoa osvetljenja. Postoje i drugi razlozi, a to su: udaljenost između kamere i oka, rotacija kamere i rotacija oka u očnoj duplji.

Fazom normalizacije formira se iris region koji uvek ima iste dimenzije u cilju održavanja karakterističnih osobina za isto oko kroz više slika koje su kreirane u različitim uslovima.

Pored navedenih problema koji se rešavaju u fazi normalizacije, bitno je i napomenuti da iris nije uvek koncentričan ili prstenastog oblika, a jako je bitno da ima stalan radijus.

Model koji se koristi je *Daugmanov Rubber sheet model* (**Slika 13**), koji ima za cilj da mapira sve tačke unutar segmentiranog iris regiona na par polarnih koordinata (r, θ) , gde je r na intervalu od $[0,1]$ i θ je ugao $[0, 2\pi]$.



Slika 13. Daugmanov Rubber sheet model

Centar oka (zenica) predstavlja referentnu tačku za radijalne vektore koji prolaze kružno kroz region irisa (**Slika 13**). Broj tačaka koji se nalazi na radijalnim linijama definisan je kao radijalna rezolucija - r . Broj radijalnih linija (kružnica) definisan je kao ugaona rezolucija - θ . Imajući u vidu da kružni region irisa može biti nekoncentričan, potrebno je primeniti metod konfiguracije i ponovo mapirati tačke u iris regionu.

Ponovno mapiranje tačaka u iris regionu (x, y) vrši se preko kartezijanskih koordinata nad normalizovanom nekoncentričnom polarnom reprezentacijom koja je modelovana kao:

$$I(x(r, \theta), y(r, \theta)) \rightarrow I(r, \theta) \quad (71)$$

sa:

$$\begin{aligned} x(r, \theta) &= (1-r)x_p(\theta) + rx_l(\theta), \\ y(r, \theta) &= (1-r)y_p(\theta) + ry_l(\theta), \end{aligned} \quad (72)$$

gde je $I(x, y)$ slika iris regiona, (x, y) su originalne kartezijanske koordinate, (r, θ) su korespondentne normalizovane polarne koordinate, x_p, y_p i x_l, y_l su koordinate zenice i ivice irisa duž smera θ .

Nakon izvršene faze normalizacije iris regiona, dobijamo iris region pripremljen za sledeću fazu. Iris region je sada predstavljen dvo-dimenzionalnom slikom, dubina piksela 8 bitova (eng. *Gray-scale*). Međutim, nikada nismo u mogućnosti da precizno rekonstruišemo iris regione zbog navedenih varijacija.

3.2.3. Kodovanje

Nakon faze normalizacije, sledi faza kodovanja. Faza kodovanja ima za cilj da iz normalizovanog regiona irisa ekstrahuje konzistentne informacije koje su sadržane u irisu. U zavisnosti od izbora konzistentnih informacija zavisice kasnije faza prepoznavanja i faza za generisanje kriptoloških ključeva na osnovu irisa.

Informaciona konzistentnost podrazumeva kodovanje samo važnih karakteristika u irisu, a važne karakteristike su vezane samo za identitet jedne osobe. Za fazu kodovanja koriste se vejevlet transformacije (eng. *Wavelet transformation*). Rezultat faze kodovanja predstavljen je sa digitalnim biometrijskim obrascem.

Biometrijski obrazac irisa koji je generisan u fazi kodovanja, spreman je za fazu prepoznavanja. Za preciznu fazu prepoznavanja potrebno je uvesti mernu jedinicu za određivanje stepena podudaranja. Ova vrsta metrike daje dva opsega vrednosti, jedan za identične irise i drugi za različite irise. Poželjno je da ne postoji poklapanje između ova dva opsega vrednosti, u drugom slučaju sumnjamo da neka od faza nije protekla kako treba. Kod ovakvih sistema dozvoljena su minimalna odstupanja koja su određena sa FAR i FRR parametrom. Od vrednosti ovih parametara zavisice prihvatljivost ovog sistema.

Vejevlet transformacija

Vejevlet transformacija se koristi za razlaganje podataka po komponentama koje se pojavljuju na različitim rezolucijama unutar slike iris regiona. Vejevleti imaju značajnu prednost u odnosu na Furijeovu transformaciju kod koje je frekvencija podataka lokalizovana. Broj vejevlet filtra ili kako se nazivaju drugačije „banka vejevleta“ je primenjena na 2D slici iris regiona. Izlaz iz ove funkcije za kodovanje je konzistentna informacija koja predstavlja određeni stepen neodređenosti (različitosti) jedne osobe u odnosu na druge osobe. Kod biometrijske kriptografije, neophodno je izabrati adekvatnu vejevlet transformaciju i konfigurisati istu u cilju dobijanja kvalitetnog biometrijskog obrasca sa maksimalnom entropijom izvora (**Slika 11**).

Gaborov filtr

Sliku je, programski gledano, lakše posmatrati i obrađivati u frekventnom domenu, gde lako možemo detektovati nevažne informacije i odstraniti ih iz signala. U ovoj disertaciji izabran je Gabor filtr [38] za kodovanje koji zbog svojih svojstava precizno razlikuje bitne informacije u slici irisa koja sadrži nekorisne ili štetne podatke prouzrokovane raznim smetnjama (na slici irisa), što je idealno za ove potrebe.

Gaborov filtr ima mogućnost da obezbedi optimalne količine konzistentnih informacija u podatku unutar jednog signala u prostornom i frekventnom domenu. Jezgro Gaborovog filtra je dvo-dimenzionalna funkcija koju čini kompleksni sinusoidni talas modelovan Gausovom funkcijom.

Realni deo funkcije definisan je funkcijom:

$$g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = \exp\left(\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \cos\left(2\pi \frac{x'}{\lambda} + \psi\right) \quad (73)$$

imaginarni deo:

$$g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = \exp\left(\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \sin\left(2\pi \frac{x'}{\lambda} + \psi\right) \quad (74)$$

gde su:

$$\begin{aligned} x' &= x \cos \theta + y \sin \theta, \\ y' &= -x \sin \theta + y \cos \theta; \end{aligned} \quad (75)$$



Slika 14. Realni i imaginarni deo impulsivnog odziva

Centar frekvencije filtra je definisan frekvencijom sinus i kosinus talasa, kao i propusnim opsegom filtra koji je modelovan Gausovom širinom.

Daugman u svom radu koristi 2D Gabor filtr za kodovanje 2D iris obrazaca. 2D Gabor filtr u domenu slike se reprezentuje kao:

$$G(x, y) = e^{-\pi[(x-x_0)^2/\alpha^2 + (y-y_0)^2/\beta^2]} e^{-2\pi[u_0(x-x_0) + v_0(y-y_0)]} \quad (76)$$

Gde je (x_0, y_0) određena pozicija u slici, (α, β) je određena efektivna dužina i širina i (u_0, v_0) je određena modulacija koja ima prostornu frekvenciju ω_0 .

Daugmanova demodulacija na izlazu Gaborovog filtra daje kompresovane podatke. Ovo se postiže kvantifikacijom fazne informacije na četiri nivoa za svaki kvadrant. Ovaj metod umesto amplitudne informacije daje najznačajnije informacije unutar slike. Uzimajući vrednosti po fazama, omogućeno je diskriminantno kodovanje iris informacije. Odbacivaće redundantnu informaciju, kao što je osvetljenje koje predstavlja amplitudnu komponentu.

Ova četiri nivoa se predstavljaju pomoću 2-bitnih podataka, tako da svaki piksel u irisu odgovara samo na dva bita informacije. Ukupno se generiše 2048 bitova podataka na osnovu jedne slike irisa i određen broj bita sa kojim se maskiraju (prekrivaju) greške izazvane raznim greškama u fazi nastanka slike, segmentacije i normalizacije. Na ovaj način dobijamo kompaktan obrazac irisa pogodan za prepoznavanje i primenu u sistemima za biometrijsku kriptografiju.

Daugmanov sistem koristi polarne koordinate za normalizaciju, određene u polarnom filtru kao:

$$H(r, \theta) = e^{-i\omega(\theta-\theta_0)} e^{-i(r-r_0)^2/\alpha^2} e^{-i(\theta-\theta_0)/\beta^2} \quad (77)$$

Gde su (α, β) , isto kao u prethodnoj formuli, (r_0, θ_0) određuju centar frekvencije filtra. Demodulacija i proces fazne kvantifikacije može biti predstavljen kao:

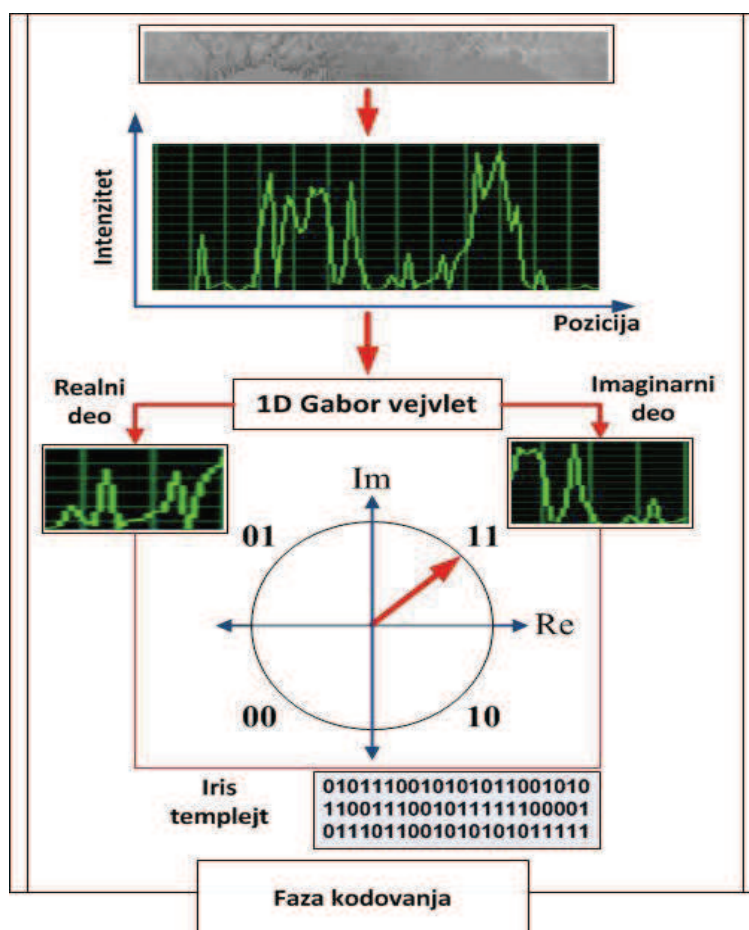
$$h_{\{Re, Im\}} = sgm_{\{Re, Im\}} \int \int_{\rho, \phi} I(\rho, \phi) e^{-i\omega(\theta-\theta_0)} e^{-i(r-r_0)^2/\alpha^2} e^{-i(\theta-\theta_0)/\beta^2} \rho d\rho d\phi \quad (78)$$

gde $h_{\{Re, Im\}}$ se može posmatrati kao složena realna vrednost i imaginarna komponenta zavisna od 2D integrala, $i(\rho, \phi)$ je iris osnovna slika u bezdimenzionalnom polarnom koordinatnom sistemu.

1D Gabor filtr

Faza kodovanja sprovedena je nad irisom nakon faze normalizacije sa 1D Log-Gabor vejvletom. 2D slika irisa je konvertovana u 1D signal i dobijeni signal je obrađen sa 1D Gaborovim vejvletom. Redovi 2D slike irisa su pretvoreni u 1D signal, a svaki red odgovara jednom kružnom prstenu iris regiona. U ovom slučaju uzima se ugaoni umesto radijalnog smera, koji odgovara kolonama normalizovane slike irisa, jer u ugaonom smeru se javlja maksimalna nezavisnost.

Intenzitet šumnih delova normalizovane slike irisa postavljen je na prosečne vrednosti okolnih piksela da bi se sprečio uticaj šuma na krajnji rezultat filtriranja. Izlaz funkcije filtriranja je kvantifikovana faza sa četiri nivoa koja koristi Daugmanov metod. Sa svakim filtrom se dobijaju dva bita podatka iz jednog fazora (kvadranta). Izlaz iz svakog fazora je podešen tako da kad se ide iz jednog u drugi, menja se samo 1 bit. Ovakva kvantifikacija, obezbediće preciznije prepoznavanje i konzistentniju informaciju jer smanjuje korelisanost između različitih irisa (**Slika 15**).



Slika 15. Faza kodovanja

Proces kodovanja generiše veliki broj bitova informacije, odgovarajuću masku koja označava šumne delove iris obrazaca i tako proglašava neke bitove u obrascu nevažećim. Uloga maske je da delove iris obrazaca koji se smatraju šumnim, proglasi nevažećim na takav način da neće uticati na krajni rezultat u fazi prepoznavanja. Međutim, ukoliko maska neutrališe više delova iris obrazaca, to će svakako uticati na preciznost, zbog smanjenja konzistentne informacije (karakteristike) koja je vezana za određenu osobu. U nekim delovima, fazna informacija će biti nekorisna, a to su delovi gde je amplituda jednaka nuli. Takođe, ovi delovi će biti označeni i proglašeni nevažećim sa maskom. Po strukturi, maska je identična iris obrascu. Ukupan broj bitova u obrascu je proizvod ugao i radijalne rezolucije i sve to puta dva bita (9600 bitova).

3.2.4. Jedinostvenost iris biometrije

Testiranje iris obrasca na jedinstvenost je najbolji test za korišćene filtre u fazi kodovanja, jer je cilj da se generišu potpuno nezavisni iris obrasci za različite osobe. Međutim, statistička nezavisnost nije postignuta. Jedinostvenost korišćenih filtra je potvrđena sa Hamingovim rastojanjem između različitih očiju. Detaljne rezultate analize moguće je pronaći u radu [39].

Sa teorijskog aspekta, merenjem Hamingovog rastojanja između različitih iris dobija se vrednost 0.5, a to je očekivana vrednost za dva potpuno slučajna niza. Verovatnoća za "0" ili

"1" jednaka je verovatnoći bacanja novčića. Ovakav rezultat ukazuje na potpunu nezavisnost između irisa različitih očiju. Sada je potpuno razumno da će dva preklapljena iris obrasca dati vrednost za Hamingovo rastojanje 0.5.

Jedinstvenost se takođe, određuje merenjem broja stepeni slobode – DOF (eng. *degree of freedom*) predstavljenih obrazaca. Ova veličina daje meru kompleksnosti iris obrasca i može se izračunati aproksimacijom čitave kolekcije različitih irisa preko svih međusobnih Hamingovih rastojanja kao binomijalna distribucija. Broj stepeni slobode DOF definiše se izrazom:

$$DOF = \frac{m(1 - m)}{\sigma^2} \quad (79)$$

gde je m aritmetička sredina, a σ standardna devijacija date distribucije.

Dobijeni rezultat za broj stepeni slobode - DOF kod "CASIA" eksperimentalne baze podataka iznosi 1068 (broj stepeni slobode), što predstavlja izuzetno dobar rezultat i garantuje jedinstvenost i nezavisnost između različitih iris obrazaca. Takođe, izračunati broj stepeni slobode - DOF pokazuje složenost iris biometrijskog obrasca.

U ovoj disertaciji za istraživanje je od izuzetne važnosti nivo poznavanja iris biometrijskog izvora (iris obrasca) za sintezu sistema za generisanje kriptoloških ključeva na osnovu irisa. Uspešna sinteza sistema direktno zavisi od stepena neodređenosti (entropije). Iz tog razloga, navešćemo neke detalje o rezultatima iz Daugmanove analize za iris kod nakon faze kodovanja u kojoj je korišćen 2D Gaborova demodulacija.

Tabela 1. Analiza vejevlet filtra korišćenih u fazi kodovanja

Veličina iris koda	Vejevlet filtr			Broj stepeni slobode - DOF
bit	Istraživač	Aritmetička sredina	Standardna devijacija	bit
2048	2D Gabor demodulacija, [1]	0.499	0.0163	249
9600	1D Gabor demodulacija, [39]	0.4993	0.0153	1068

U njegovom radu [1] je utvrđeno na velikim uzorcima da iris kod ima 249 stepeni slobode – DOF, a veličina iris koda je 249 stepeni slobode. Uporedno posmatrajući rezultat koji je dobijen sa 1D Gabor demodulacijom, rezultati su procentualno po jednom bitu identični (11%) (**Tabela 1**).

3.3. Zaštitni kodovi

Prenos informacija podrazumeva ili njihovo prenošenje iz jedne tačke u drugu ili njihovo prenošenje kroz vreme. Primer prenošenja od jedne tačke do druge je komuniciranje između dva računara na Internetu, dok je primer prenošenja kroz vreme memorisanje nekog sadržaja na nekom memorijskom mediju, a zatim njegovo iščitavanje u nekom budućem trenutku.

U oba slučaja, susrećemo se sa problemom pouzdanog prenosa informacija, kako usled delovanja šumova na kanalu za prenos, tako i u nedostatku memorijskog prostora ili njegovog kvara. Rešenje je u kodovanju i postoje tri osnovna razloga zašto nešto kodujemo:

- Kodovanje radi kompresije (izbacuje se redundantna);
- Kodovanje u cilju kvalitetnog prenosa preko šumnih kanala (dodaje se redundantna);
- Kodovanje radi tajnosti (poruke su dostupne samo autorizovanim korisnicima).

Tehnike kodovanje o kojima ćemo u disertaciji da govorimo su namenjene za sveprisutni šumove na kanalima prenosa koji mogu delovati na jedan od sledećih načina:

- U uslovima izuzetno visokog šuma nije uopšte moguće obaviti pouzdan prenos poruka;
- Nivo šuma je takav da je moguće preneti poruke sa prihvatljivim nivoom grešaka u prenosu;
- Moguće je preneti poruke sa verovatnoćom greške koja se može učiniti po želji malom primenom zaštitnih kodova (eng. *Error Correcting Code - ECC*). Osnovna ideja zaštitnog kodovanja je dodavanje redundanse u kodovane poruke, tako da i pored grešaka u prenosu preostaje dovoljno nepromenjenih informacija u kodovanim porukama na osnovu kojih se može rekonstruisati poslata poruka sa zadatim nivoom greške rekonstrukcije (**Slika 16**).



Slika 16. Model komunikacionog kanala

Komunikacioni kanal, ili kraće kanal, reprezentuje sve što se može dogoditi sa porukom u toku njenog prenosa od predajnika do prijemnika. Poruka je niz simbola. Simbol je element konačnog skupa. Ulazna sekvenca $X_1, X_2, \dots$ (sekvenca koju treba preneti) je određena u potpunosti informacionim izvorom. Izlazna sekvenca $Y_1, Y_2, \dots$ (primljena sekvenca) je određena uslovnom verovatnoćom izlaza za poznati ulaz. Stoga je u matematičkom smislu kanal određen uslovnim verovatnoćama različitih primljenih sekvenci, pod uslovom poznatih sekvenci koje se šalju, tj $P(y_1, y_2, \dots, y_n | x_1, x_2, \dots, x_n)$.

Mera sposobnosti jednog komunikacionog kanala da prenosi informacije, naziva se kapacitet kanala. Kapacitet C diskretnog kanala bez memorije jednak je:

$$C = \max_{p_X} I(X; Y), \quad (80)$$

gde je X ulaz u kanal, opisan raspodelom verovatnoće $p_X(x)$, dok je Y izlaz kanala. U ovom slučaju kapacitet kanala je maksimalna zajednička informacija između X i Y .

Kapacitet binarnog simetričnog kanala određen je po definiciji:

$$C = \max_{p_X} (H(Y)) - \tilde{h}(p), \quad (81)$$

Kako je Y binarna slučajna promenljiva, važi $H(Y) \leq \log 2$, odnosno $H(Y) \leq 1$ bit. Ovo je moguće postići ukoliko je X uniformno raspodeljeno. Tada imamo:

$$p_Y(0) = p \cdot p_X(1) + (1-p) \cdot p_X(0) = 0.5p + 0.5(1-p) = 0.5 \quad (82)$$

Što znači da je Y takođe, uniformno raspodeljeno, odnosno da je $H(Y) = 1$ bit. Stoga je $\max_X H(Y) = 1$ bit, odnosno, konačno:

$$C = 1 - \tilde{h}(p). \quad (83)$$

Brzina prenosa (za osnovu b) koda kojim se koduje diskretan izvor U sa $|v_U|$ poruka, čije su kodne reči fiksne dužine n je definisana sa:

$$R_b = \frac{\log_b |v_U|}{n} \quad (84)$$

Primitimo da je $|v_U|$ istovremeno i broj mogućih kodnih reči determinističkih nesingularnih kodova. Po pravilu se osnova b za računanje brzine koda, uzima da je jednaka arnosti koda.

Razmotrimo prenos poruka brzinom R preko DMC-a (eng. *discrete memoryless channel*) bez povratne sprege, čiji je kapacitet C . Za sve $R < C$ i sve $\varepsilon > 0$, postoji kod za ispravljanje grešaka brzine R i greške dekodovanja $P_b < \varepsilon$. Uopšteno govoreći ova važna teorema tvrdi da ako izaberemo odgovarajući način kodovanja informacije za zadati kanal, može se postići po želji mala greška dekodovanja, odnosno prenosa informacija.

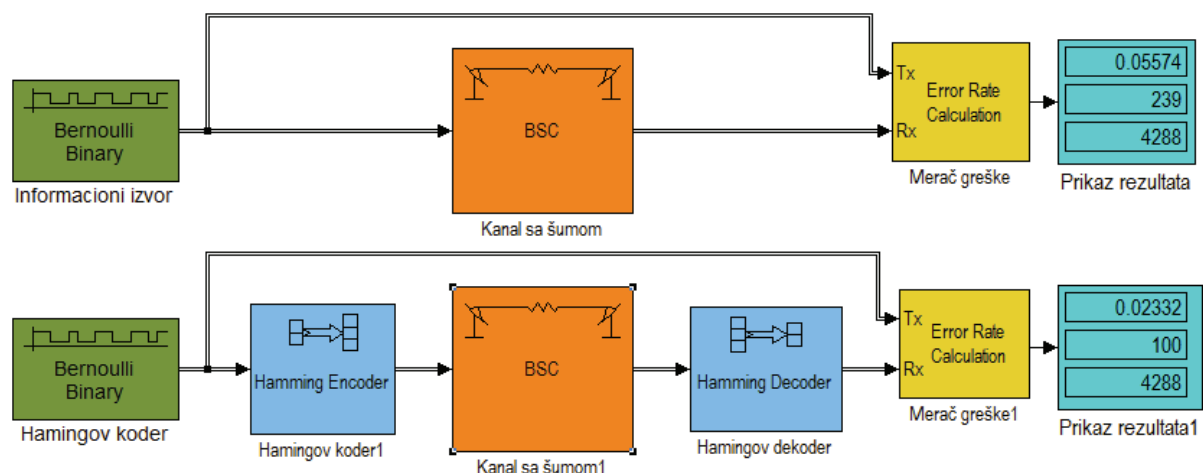
Data teorema je bila veliko iznenađenje u radu [33]. Pre ovog rezultata smatralo se da u cilju pouzdanog prenosa, moramo smanjivati brzinu prenosa ili ekvivalentno povećavati odnos signal/šum.

Šenon je srušio ovaj mit pokazujući da ako je brzina prenosa ispod kapaciteta kanala, povećanje pouzdanosti prenosa se može u potpunosti ostvariti konstrukcijom kompleksnijih sistema kodovanja i dekodovanja, bez potrebe da se menja odnos signal/šum.

Teorema nema praktičnu dimenziju. U njoj nisu dati recepti kako efikasno konstruisati dobre kodove. Na primer, ništa se ne kaže kolika je dužina kodnih reči n u funkciji parametra ε .

Stoga je nastala čitava disciplina koja se zove “Teorija kodovanja” (eng. *Coding Theory*), čiji je cilj dizajniranje efikasnih zaštitnih kodova, odnosno kodova što veće brzine i što manje greške dekodovanja.

Pionir u ovoj oblasti je američki matematičar Ričard Hamming (Richard Hamming) koji je osmislio prvi kod za ispravljanje grešaka (ECC), poznat kao Hamingov kod. Takođe, i danas memorijski kontroleri tradicionalno koriste Hamingov kod za ispravljanje grešaka. Iz ovog razloga pokažaćemo u sledećoj simulaciji performanse ovog koda za zadate parametre.



Slika 17. Primena Hamingvog blokovskog zaštitnog koda

Slika 17., predstavlja simulaciju jedne primene Hamingovog blokovskog koda (7,4). U gornjem delu ilustracije prikazan je prenos poruka preko kanala sa šumom bez primene zaštitnog koda. Verovatnoća greške u kanalu je postavljena na $p = 0.05$. U donjem delu ilustracije prikazan je identičan kanal sa dodatim Hamingovim koderom na predajnoj strani i Hamingovim dekodrom na prijemnoj strani. Sada kada pogledamo krajnji rezultat (prikazan na krajnjoj desnoj komponenti u simulaciji) vidimo da je verovatnoća greške smanjena sa $p = 0.05574$ na $p = 0.02332$. Preko oba kanala je preneseno ukupno 4288 bitova poruke. U prvom kanalu je bilo 239 pogrešnih bitova ($X \neq Y$), a u drugom kanalu 100 bitova. Kompletna simulacija je realizovana u MATLAB i Simulink programskom paketu koji nam je omogućio praktičnu demonstraciju svih zaštitnih kodova koji su od značaja za ovu disertaciju.

3.2.1. Klase zaštitnih kodova

U ovom poglavlju govorićemo o klasama zaštitnih kodova (ECC). Podela je napravljena na dve klase:

- Linearni kodovi;

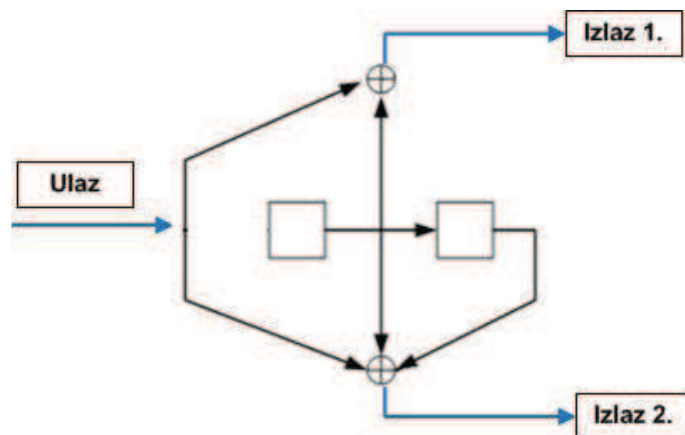
- Konvolucioni kodovi.

Linearni kodovi - predstavljaju jednu od najznačajnijih klasa kodova zbog njihove jednostavnosti i lake implementacije. U teoriji kodovanja linerani kod je kod za ispravljanje grešaka za koga je bilo koja linearna kombinacija kodnih reči takođe, kodna reč. Linerani kodovi su tradicionalno podeljeni na blokovske i konvolucione kodove, mada turbo kodovi se mogu posmatrati kao hibrid ova dva tipa. Linearni kodovi omogućavaju dizajniranje efikasnijih algoritama za kodovanje i dekodovanje, za razliku od drugih kodova.

Linearni kodovi se koriste u FEC i primenjuju se u metodama za prenošenje simbola (bitova) preko komunikacionog kanala, tako da, ako dođe do greške na nekom bitu u komunikaciji, te greške mogu biti detektovane i ispravljene na strani primaoca. Kodne reči u lineranom blokovskom kodu su blokovi simbola u kojima je kodovana originalan vrednost. Linerani kod dužine n , prenosi blok koji sadrži n simbola. Na primer, Hamingov kod (7,4) o kom smo govorili u prethodnom poglavlju, predstavlja 4 bita poruke za jednom kodnom reči čija je dužina 7. Sa kodom (7,4) moguće je otkriti dve greške i samo jednu grešku ispraviti. Ovaj kod ima ukupno $2^4 = 16$ kodnih reči.

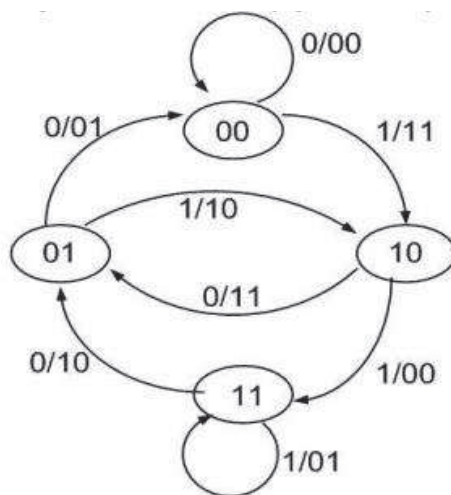
Druga vrsta linearnog blokovskog koda je ciklični kod. Ciklični kod je takav da ako je Z_i kodna reč sa n simbola $Z_{i,1}, Z_{i,2}, \dots, Z_{i,n}$, tada je i reč $Z_{i,2}, Z_{i,3}, \dots, Z_{i,n}, Z_{i,1}$ koja se dobija cikličnom permutacijom (poznatom i kao pomeranje u levo) simbola kodne reči Z_i , takođe kodna reč datog cikličnog koda. Ovo praktično znači da je svaka ciklična permutacija kodne reči takođe kodna reč. Ciklični kodovi su važna klasa linearnih kodova, pošto je u ovoj kasi moguće konstruisati efikasne procedure kodovanja i dekodovanja.

Konvolucioni kodovi - ne pripadaju klasi blok kodova. Mehanizam kodovanja pamti prethodno kodovane simbole. U izvesnom smislu se mogu tretirati kao neograničeni blok kodovi. Proces kodovanja i dekodovanja se značajno razlikuje. Daleko su efikasniji od blok kodova i imaju memoriju. Na ulaz konvolucionog koda se dovodi k bitova, a na izlazu iz njega se pojavljuje n bitova. U ovom zaštitnom kodu, izlazni bitovi ne zavise samo od trenutnih bitova na ulazu, već i od prethodno pristiglih bitova.



Slika 18. Primer konvolucionog koda

Slika 18., predstavlja jedan primer konvolucionog koda. Na ulazu koda je 1 bit ($k=1$), a na izlazu su 2 bita ($n=2$). Koder sadrži dva pomeračka registra koji imaju mogućnost pamćenja prethodnih bitova na ulazu ($m=2$). Rad ovog konvolucionog koda predstavljen je dijagramom stanja (**Slika 19**).



Slika 19. Dijagram stanja konvolucionog koda

Na ovom dijagramu, konvolucionni koder ima četiri moguća stanja (00,10,01,11). Promena stanja u koderu predstavljena su usmerenim strelicama. Sa odgovarajućim granama u dijagramu stanja, naznačeni su bitovi koji ulaze i bitovi koji izlaze iz koda. Na primer, „0/11“ označava da iz koda izlaze bitovi „11“ kao rezultat ulaznog bita „0“, pri čemu je stanje koda bilo „10“.

3.2.2. Rid-Solomon kodovi

Rid-Solomon - RS (eng. *RS-Reed-Solomon*) kodovi ispravljaju greške na nivou simbola, a ne bitova, pogodni su za ispravljanje koncentrisanih paketskih (eng. *burst*) grešaka. Rid-Solomon kodovi se obično izučavaju na svim akademskim institucijama u svetu. Oni imaju široku primenu u sistemima digitalnih komunikacija i sistemima za čuvanje podataka. Rid-Solomon kodovi su podgrupa BCH kodova i klasa linearnih, nebinarnih cikličnih kodova. Rid-Solomon kod je blok kod koji se predstavlja kao $RS(n,k,t)$ preko polja Galoa $GF(2^n)$.

Promenljiva n je veličina kodne reči (veličina bloka) sa jedinicom simbola, k je broj simbola podataka (informacije), $2t$ je broj simbola parnosti i svaki simbol sadrži m bitova (veličina simbola). Korelacija veličine simbola m i veličine kodne reči n je određena kao $n=2^m-1$ za m bitova u jednom simbolu. Postoji ukupno 2^m-1 različitih simbola u jednoj kodnoj reči, osim onog sa svim nulama. Sposobnost ispravljanja grešaka ili maksimalan broj ispravljenih slučajnih grešaka je dat sa:

$$t = \frac{(n-k)}{2} \quad (85)$$

Na primer, $RS(63,59)$ ima veličinu kodne reči n , a to je 63 simbola, broj simbola podatka je k , a to je 59 simbola. Maksimalan broj pogrešnih simbola koje $RS(63,59)$ dekodirer može da ispravi je jednako t , a to je ukupno 2 simbola.

$RS(n,k)$ kodovi postoje za svako n i k gde je $0 < k < n < 2^m + 2$. Većina konvencionalnih $RS(n,k)$ kodova je određena notaciom:

$$(n,k) = (2^m - 1, 2^m - 1 - 2t), \quad (86)$$

gde Rid-Solomon kod može biti proširen sa $n = 2^m$ ili $n = 2^m + 1$, ali ne i dalje.

Rid-Solomon kodovi postižu najveći mogući kod minimalnog rastojanja u odnosu na bilo koji drugi linerani kod. Minimalno rastojanje koda je minimalno Hamingovo rastojanje između bilo koje dve različite kodne reči. Minimalno rastojanje jednog koda je od centralnog značaja u teoriji kodova za ispravljanje grešaka. Za nebinarne kodove rastojanje između bilo koje dve kodne reči je određeno kao broj simbola u sekvenci koji su različiti od Rid-Solomon koda i taj kod najmanjeg rastojanja je određen notaciom:

$$d_{\min} = n - 1 + 1 \quad (87)$$

Kod je nadležan za ispravljanje bilo koje kombinacije t ili manjeg broja grešaka, a ovo se može dobiti sa sledećim izrazom za sposobnost ispravljanja grešaka:

$$t = \frac{(d_{\min} - 1)}{2}. \quad (88)$$

RS kod funkcioniše sa polinomijalnom konstrukcijom koja je konstruisana iz podataka, čime je generisana redundantna informacija. Ova redundantna informacija omogućava ispravljanje greške tokom prenosa ili skladištenja. Količina i vrsta greške koja može biti ispravljena zavisi od karakteristika Rid-Solomon koda. Optimalni kodovi koji su ostvareni za ispravljanje grešaka, nazvani su MDS (eng. *maximum distance separable*) kodovi. Rid-Solomon kodovi su vodeći član po broju i korisnosti klase MDS kodova. Parametri Rid-Solomon kodova su sumirani dole.

Dužina kodne reči:

$$n = 2^m - 1. \quad (89)$$

Broj simbola podataka u kodnoj reči:

$$k = 1, 3, \dots, n - 2. \quad (90)$$

Minimalno rastojanje koda:

$$d_{\min} = n - k - 1 \quad (91)$$

Sposobnost RS koda za ispravljanje grešaka:

$$t = \frac{(n - 1)}{2} \quad (92)$$

Sledeći dijagram (Slika 20) pokazuje tipičnu RS kodnu reč (poznatu kao sistematski kod zato što je podatak sa leve strane nepromenjen, a simboli parnosti su dodati sa desne).



Slika 20. RS kodna reč

U radu [40] istražene su performanse Rid-Solomon koda kao jednog od fleksibilnijih kodova. Posmatrano je više kriterijuma za dobre performanse RS koda. Merena je stopa greške – BER (eng. *bit-error-rate*). Primetno poboljšanje performansi Rid-Solomon koda je ostvareno povećanjem redundanse, čime je sposobnost koda za ispravljanje grešaka porasla.

3.2.3. Tehnike zaštitnog kodovanja

U poglavlju primene zaštitnih kodova govorićemo o popularnim tehnikama za kodovanje, od kojih ćemo navesti neke:

- ARQ (eng. *Automatic Repeat Request*) automatsko izdavanje zahteva za ponavljanjem poruke;
- FEC (eng. *Forward Error Correction*) dekodovanje primljene sekvence kao najbliže kodne reči;
- Hibridni ARQ;
- Konkatenacija kodova.

ARQ - Zaštitno kodovanje se obavlja tako da kod sam po sebi ne ispravlja, već samo detektuje greške. Kada kod detektuje greške, šalje se zahtev za ponavljanjem poruke. U pogledu greške dekodovanja, ARQ je bolji od FEC, budući da se primljene kodne reči prihvataju samo ako su prenete bez greške (dakle greška dekodovanja je nula), pod uslovom da se detekcija grešaka uvek obavlja ispravno. Ovo se ne dobija besplatno. Cena je plaćena snižavanjem brzine prenosa. Dodatna prednost ARQ je daleko jednostavnije dekodovanje u poređenju sa FEC.

ARQ zahteva dodatni povratni kanal, koji zavisno od primene i scenarija komuniciranja nije uvek na raspolaganju. U pogledu ometanja, ARQ je osetljiv, budući da ometanje može biti tako dizajnirano da se uvek dešava po neka greška u svakoj kodnoj reči. Optimalna situacija

sa tačke gledišta ometača je da neprekidno forsira jednu grešku u jednoj kodnoj reči, čime bi se ta kodna reč neprekidno reemitovala. Postoje dva tipa ARQ:

- Stani i čekaj;
- Neprekidni.

Stani i čekaj ARQ predajnik, šalje jednu kodnu reč i čeka pozitivnu potvrdu ispravnog prijema (ACK) ili negativni potvrdu (NAK), pre nego što pošalje narednu kodnu reč. Dobra strana stani i čekaj ARQ je potreba za poludupleksnim kanalom. Negativna strana stani i čekaj ARQ je trošenje vremena na čekanje potvrde, što snižava ekvivalentnu brzinu prenosa. Neprekidni ARQ zahteva dupleksni kanal, tako što se kodne reči po jednom kanalu neprekidno šalju sve dok se po drugom kanalu ne primi NAK. Postoje dva načina obrade NAK-a:

- ARQ sa slanjem unazad - N kada se ponovo šalje kodna reč u kojoj je detektovana greška plus sve kodne reči unazad. Parametar N se određuje na osnovu ukupnog kašnjenja kanala u oba smera. Na primer, kod geostacionarnih satelitskih kanala, kašnjenje može biti vrlo veliko, do 540 msec. Prijemnik mora da memoriše N kodnih reči primljenih u okviru ovog vremena, što zahteva veliku memoriju.
- Selektivni ARQ, kod koga se prenosi ponovo samo kodna reč u kojoj je detektovana greška, čime se dobija veća brzina prenosa nego u slučaju ARQ sa slanjem unazad - N . Oba tipa neprekidnog ARQ daju veću brzinu prenosa u odnosu na stani i čekaj ARQ, na račun većih memorijskih zahteva.

FEC - je pogodan za primene u kojima korisnik mora dobiti tačnu poruku iz prvog slanja. Primer su jednosmerni i difuzioni (broadcast) kanali. Današnji kodovi za ispravljanje grešaka pripadaju dvema osnovnim kategorijama: blok kodovima i konvolucionim kodovima.

Najpopularniji blok kodovi su Rid Solomon kodovi (RS-Reed-Solomon) koji rade na nivou m -bitnih simbola, a ne na nivou bita. Budući da RS kodovi ispravljaju greške na nivou simbola, a ne bita, pogodni su za ispravljanje koncentrisanih, paketskih (eng. *burst*) grešaka. Na primer – simbolni RS kodovi sa 8 bitova po simbolu, mogu korigovati sve pakete grešaka od 16 bita ili manje. Danas se koriste u JTIDS – novom standardu za duboki kosmos i za kompakt disk (CD) plejere.

Konvolucionni kodovi uvode ulazni niz bita poruke u pomerački registar dužine r . Snaga ovih kodova zavisi od veličine r . Veće r daje efikasnije kodove, ali to znatno usložnjava dekodovanje. Postoji više metoda dekodovanja konvolucionih kodova, ali je danas najrasprostranjeniji Viterbijev algoritam, otkriven 1967. godine od strane Andrew Viterbija. Danas je raspoloživ u jednom čipu izrađenom VLSI tehnologijom. Viterbijev dekodler je neefikasan i nepraktičan za veliko K . Za veliko r , ($r < 9$) češće se koristi drugi algoritam dekodovanja koji se naziva sekvencijalni dekodler.

Treći postupak dekodovanja je dekodler sa povratnom spregom, pogodan za kanale sa paketnim greškama, dok je znatno slabijih performansi ako su greške na kanalu slučajno

raspodeljene. Generalno se može reći da konvolucioni kodovi daju bolje rezultate od blok kodova jednake kompleksnosti kodovanja i dekodovanja. Problem sa svim kodovima koje smo do sada razmatrali je potreba da se proširi propusni opseg za onoliko koliko zahtevaju dodati biti parnosti (redundantni biti), ako želimo da zadržimo brzinu prenosa nekodovanih poruka istom.

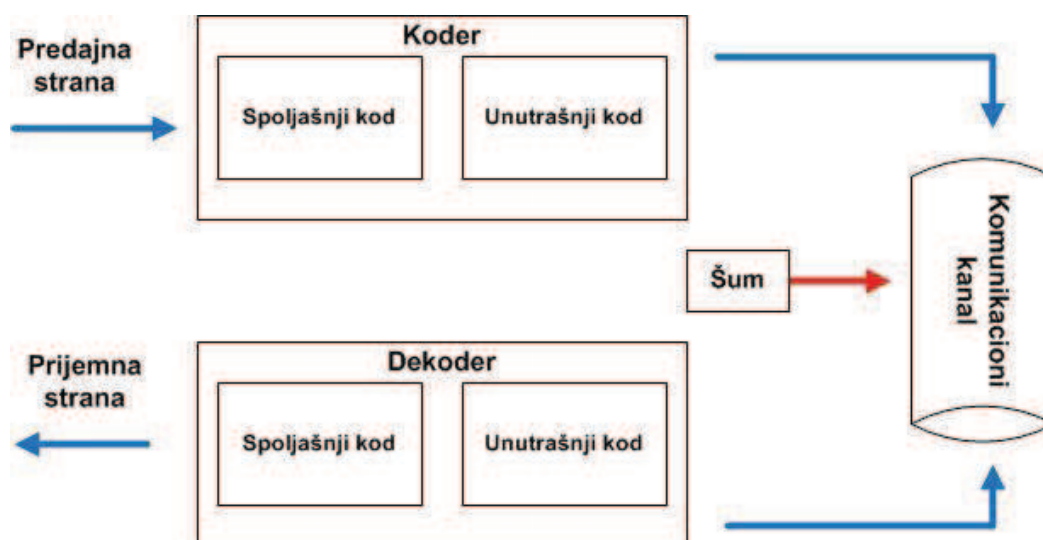
Hibridni ARQ - kombinuju kodove za detekciju i kodove za ispravljanje grešaka, čime se povećava efikasnost korišćenja kanala. U prijemniku se prvo pokušavaju ispraviti sve greške u primljenoj kodnoj reči. Ukoliko se u tome ne uspe, koristi se jedna od pomenute tri ARQ tehnike za retransmisiju.

Hibridni ARG I vrste šalje sve neophodne bitove parnosti za detekcije i ispravljanje grešaka.

Hibridni ARQ II vrste šalju samo bitove parnosti neophodne za detekcije grešaka i čuvaju u rezervi bitove parnosti za ispravljanje grešaka. Kada prijemnik detektuje grešku, zahteva slanje bitova za korekciju grešaka i na osnovu njih pokušava da ispravi greške pre nego što pošalje zahtev za retransmisiju celokupne kodne reči. Hibridni ARG II vrste je vrlo efikasan za kanale koji se karakterišu dobrim stanjima koja su većinska u odnosu na ređa loša stanja kanala.

Konkatenacija kodova - predstavlja dva koda u sirijskoj vezi. Po pravilu se koristi na onim kanalima na kojima jedan kod ne može da koriguje sve greške koje se mogu pojaviti.

Na primer (**Slika 21**), unutrašnji kod brzine $\frac{1}{2}$, $K = 7$, je konvolucioni kod sa Viterbijevim dekodovanjem, kojim se koriguje većina slučajnih grešaka. Spoljašnji kod je Rid-Solomon kod sa 16 simbola, kojim se ispravljaju sve paketske greške koje je propustio Viterbijev dekoader.



Slika 21. Šema konkatentnog koda

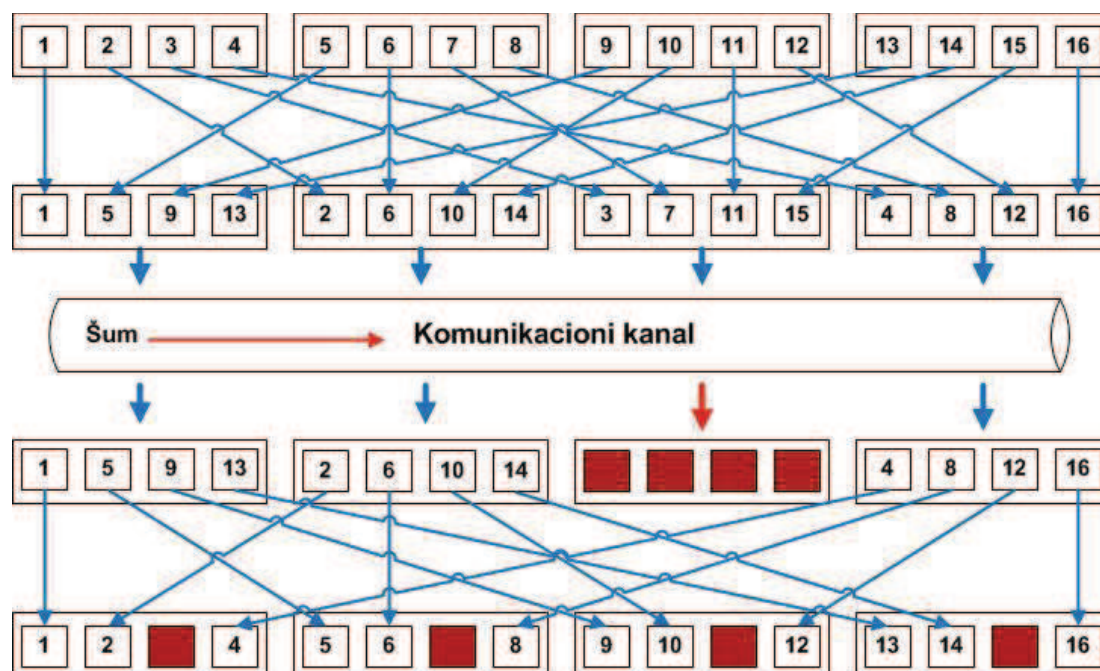
RS kod radi sa 8 bitnim simbolima i stoga je veoma efikasan u ispravljanju paketskih grešaka. Ukupna brzina konkatentnog koda jednaka je proizvodu brzina sastavnih kodova,

odnosno u ovom primeru $(1/2)(223/255) \cong 0.44$. Kodni dobitak konkatentnih kodova je veoma velik. Neki od njih rade bukvalno na nivou termalnih šumova i tada se pojavljuje problem sinhronizacije. I pored ovih dobrih svojstava, konkatentni kodovi se teško implementiraju zbog njihove složenosti i male brzine, što je veliki problem u kanalima sa ograničenim propusnim opsegom.

Najbolji primer je CIRC šema kodovanja koja je postala ekstremno uspešna. Unutrašnji kod ima minimalno rastojanje $d_{\min} = 5$, tako da može da koriguje najviše dva pogrešna simbola. Spoljašnji kod ima takodje minimalno rastojanje $d_{\min} = 5$ i služi ili za ispravljanje grešaka ili za ispravljanje brisanja, zavisno od toga koliko je unutrašnji kod bio uspešan u korigovanju paketskih grešaka. Ona je konvertovala kanal koji nije tolerisao trunku prašine u vrlo robustan praktičan sistem CD audio nosača koji je osvojio svet. Današnji CD može da toleriše 8 mm velike rupe probušene kroz disk, bez audio perceptivnih efekata.

3.2.4. Interliveri

Interliveri su jedan od najpopularnijih načina da se isprave paketske greške korišćenjem kodova koji dobro rade na slučajnim greškama, tako što se prethodno kroz postupak interlivinga paketske greške rasprše tako da izgledaju kao slučajno raspoređene. Proces interlivinga možemo opisati na sledeći način. Na predajnoj strani posle funkcije kodovanja, nad kodnim rečima primenjujemo funkciju interlivinga. Zatim poruku šaljemo kroz komunikacioni kanal u kom postoji verovatnoća za nastanak paketskih grešaka. Na kraju, na strani primaoca primenjuje se inverzna funkcija za deinterliving i dobijena poruka se prosleđuje funkciji za dekodovanje koja može da ispravi sve slučajno raspoređene greške. Slučajna raspodela grešaka postiže se primenom interlivinga (Slika 22).



Slika 22. Jedan primer blokovskog interlivinga

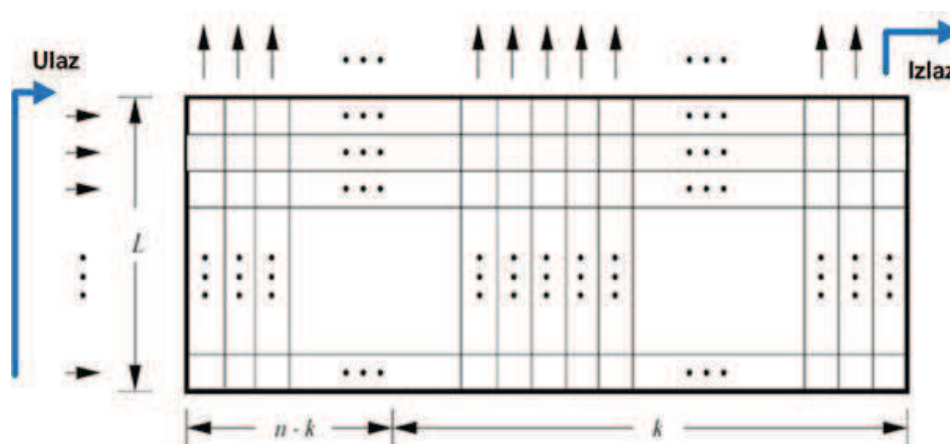
Postoje dva tipa interlivinga koji se danas najviše koriste:

- Blok interliving;
- Konvolucioni interliving.

Blokovski interliving - se popunjava sa L kodnih reči složenih po redovima. Svaka kodna reč je dužine n (Slika 23). Ovih L kodnih reči se zatim šalje po kolonama, sve dok se celokupni interliver ne isprazni. Nakon toga se interliver ponovo puni i započinje novi ciklus.

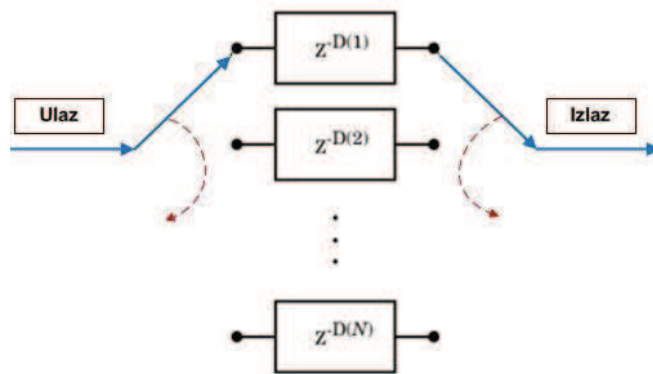
Na prijemu se kodne reči deinterlivinguju pre dekodovanja. Paketne greške dužine L ili manje bita ne izazivaju više od 1 bita greške u bilo kojoj kodnoj reči. Dekoder koji dobro radi na slučajnim greškama će sa velikom verovatnoćom ispraviti po jednu grešku u svakoj kodnoj reči umesto celokupan paket grešaka.

Parametar L se naziva stepen interlivera ili dubina interlivinga. L se bira prema najgorim uslovima prenosa datog kanala. Mora biti veći od najvećeg očekivanog paketa grešaka u datom kanalu. Osnovni nedostatak interlivinga je kašnjenje koje se unosi popunjavanjem redova. Konvolucioni interliving eliminiše ovaj problem, osim kašnjenja vezanog za početno punjenje. Ovim se redukuju memorijski zahtevi na polovinu u odnosu na blokovske interlivere. Zajednički nedostatak oba tipa interlivera je kašnjenje koje u nekim kanalima može iznositi i nekoliko sekundi, što može biti neprihvatljivo za neke aplikacije (recimo prenos govora).



Slika 23. Šema blokovskog interlivinga

Konvolucioni interliving - se sastoji od skupa pomeračkih registara, svaki sa ustaljenim kašnjenjem. Svaki novi simbol iz ulaznog vektora napaja susedni prenosni registar, a "najstariji" simbol u tom registru postaje deo izlaznog vektora. Konvolucioni interliver ima memoriju čija operacija ne zavisi samo od trenutnog simbola, već i od prethodnog. Šema (Slika 24), prikazuje strukturu konvolucionog interlivera, pokazujući skup smena registra i njihove vrednosti kašnjenja $D(1), D(2), \dots, D(N)$. KTH registar drži $D(k)$ simbol, gde je $k = 1, 2, \dots, N$.



Slika 24. Šema konvolucionog interlivera

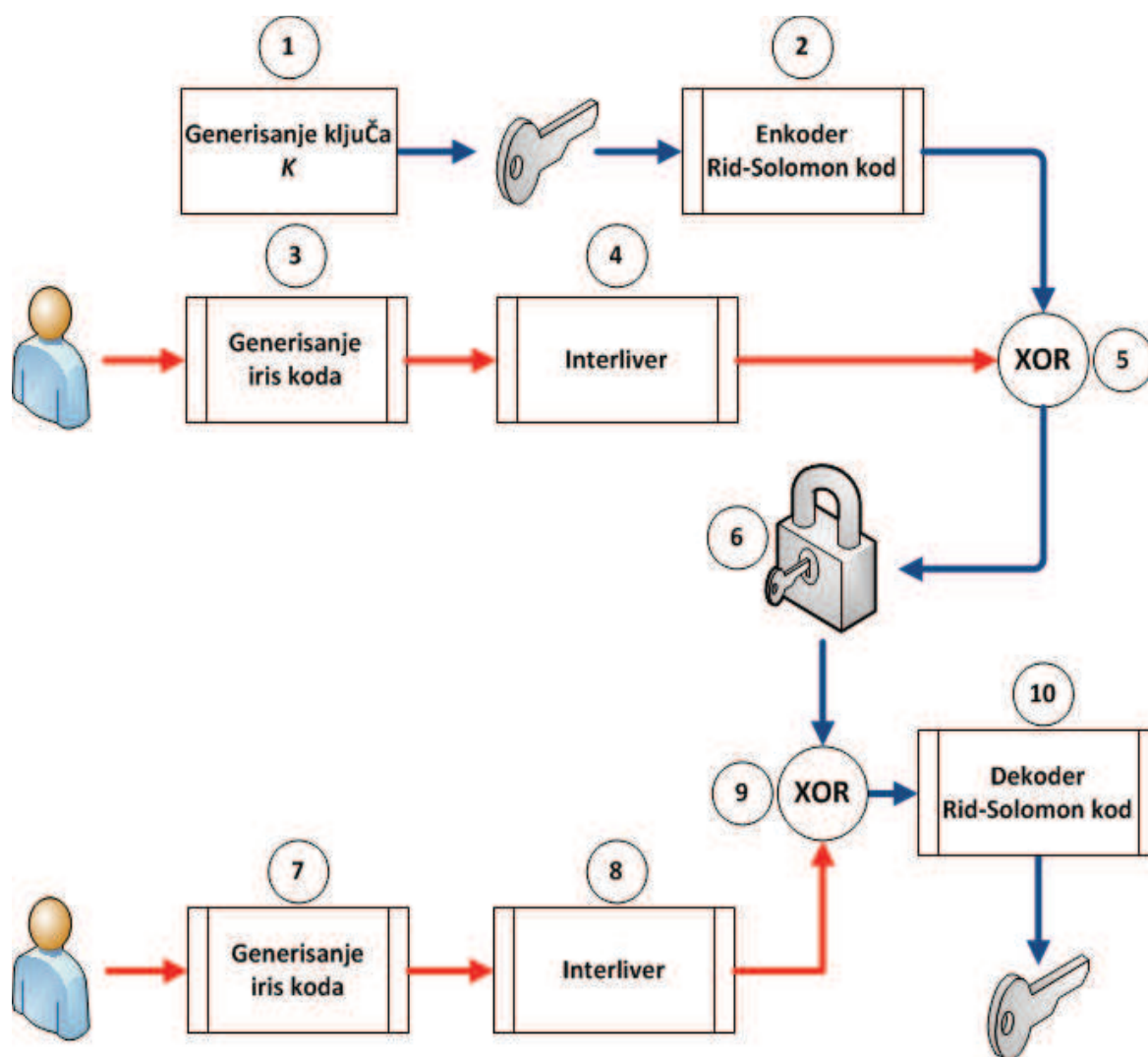
Pseudoslučajni interliveri – u ovoj tehnici interlivinga ideja je da se eliminišu regularni obrasci ili određene statističke osobine iz određene sekvence podataka i to tokom veoma dugog perioda. Pseudoslučajni interliver se često koristi kod turbo kodova. Uniformni slučajni interliver je stvarna apstrakcija nad statističkim verovatnoćama za neki interliver sa ogromnom periodom L . Ovakvi interliveri pokušavaju da uvedu element slučajnosti u neki dizajn koda bez povećanja kompleksnosti, pretpostavljajući korišćenje iterativnog dekodovanja dva interlivovana koda. U praksi pseudoslučajni interliveri su predvidljivi i invertibilni. Ponašaju se kao PRNG generatori koji imaju određenu periodu za svako izabrano unutrašnje stanje generatora. Perioda generatora određena je sa polinomom. Ukoliko su podaci matrice sa više redova, funkcija procesira pojedinačno kolone matrice.

4. Novi sistem za generisanje ključeva na osnovu biometrije irisa

U ovom delu rada data su uvodna razmatranja vezana za različite aspekte razvoja novog sistema za generisanje kriptoloških ključeva na osnovu biometrije irisa, kao i kompletan proces njegovog razvoja i integracije u standardne biometrijske sisteme. U poglavlju su date i komponente tradicionalnog biometrijskog sistema za prepoznavanje na osnovu irisa u kog će naš sistem biti integrisan, kao i izbor nosećih tehnologija napravljen na osnovu analize trenutnog stanja na polju biometrijske kriptografije i bezbednosti.

4.1. Generička šema predloženog rešenja

Za sintezu jednog kompleksnog sistema neophodno je napraviti generičku šemu sa detaljnim opisima svih komponenti koje na direktan ili indirektan način učestvuju u svojstvu gradivnih komponenti ovog novog predloženog rešenja (Slika 25.).



Slika 25. Generička šema predloženog rešenja

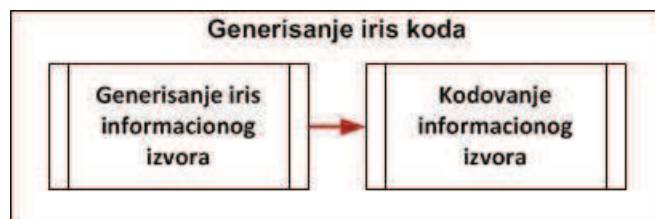
Na slici je predstavljena generička šema predloženog rešenja sa svim svojim podsistemima. Kompletan sistem sastoji se od dva podsistema: podsistem za kodovanje iris biometrijskog izvora koji je zasnovan na algoritmima tradicionalnih biometrijskih sistema i podsistemu za prenos poruka preko biometrijskog kanala sa šumom. U nastavku sledi detaljno objašnjenje pojedinačnih uloga svih navedenih komponenti koje su sadržane u ovom dijagramu procesa. Komponente su numerisane brojevima od 1 do 10.

Komponenta 1 - predstavlja proces u kom se generiše kriptološki ključ zadate veličine. Preporučeno je da ključ bude generisan na slučajan način preko informacionih izvora koje je moguće pronaći u prirodnim okruženjima. U drugom slučaju koji nije toliko popularan, moguće je ključ generisati i nekim pseudoslučajnim generatorom. Ovakav način generisanja ključa može da naruši bezbednost jednog simetričnog kripto sistema.

Komponenta 2 - predstavlja proces zaštitnog kodovanja za korekciju grešaka u kome se koriste Rid-Solomon kodovi. U ovom procesu kriptološki ključ K je kodovan Rid-Solomon koderom. Uloga ovog procesa je od suštinske važnosti za sistem, zapravo tehnika preuzeta iz polja komunikacija, na osnovu koje je nastala i čitava ideja o potpuno jednom novom naučnom istraživačkom polju pod nazivom "biometrijska kriptografija".

Komponenta 3 – predstavlja proces generisanja iris koda koji pripada podsistemu za kodovanje iris biometrijskog izvora. U našem slučaju specifičan proces koji u sebi sadrži dva podprocesa (**Slika 26.**). Prvi podproces se odnosi na standardnu proceduru za generisanje biometrijskog obrasca koji je opisan u poglavlju 3. Drugim rečima, možemo da kažemo da je ovo jedan složen proces za formiranje iris informacionog izvora na osnovu koga je kasnije urađena sinteza kompletnog rešenja o kome je reč. Drugi podproces predstavlja još jedan složen proces koji uvodi informacione mere i na osnovu rezultata dobijenih preko teorijsko-informacione analize koduje informacioni izvor u cilju izdvajanja najkonzistentnijih informacija.

Napomenućemo da je ovo kritična tačka u ovom sistemu, gde se dobijen loš rezultat prenosi domino efektom na preostale komponente. Zapravo, performanse čitavog sistema zavise od kvaliteta procesa kodovanja iris informacionog izvora. Iz tog razloga sprovedli smo detaljnu i rigoroznu teorijsko-informacionu analizu o kojoj ćemo kasnije govoriti u ovom poglavlju.



Slika 26. Proces generisanja iris koda

Komponenta 4 – predstavlja još jednu tehniku preuzetu iz teorije komunikacija. Interliver ima značajnu ulogu u sistemu. Koristi se za raspršivanje paketskih grešaka i smanjenje uzajamne informacije između različitih irisa. Uticaj interlivera na uzajamne informacije dat je

u podpoglavlju 4.3. U predloženom rešenju koristili smo slučajni interliver (eng. *Random interliver*). Tehniku interlivinga primenićemo nad kodovanim iris kodom.

Komponenta 5 - predstavlja logičku operaciju “XOR” u kojoj se bodovan kriptološki ključ čvrsto vezuje za originalni iris kod nad kojim je primenjena tehnika za interliving. Imajući u vidu da su oba podatka slučajna, dobijamo sistem koji po svim svojim osobinama, dosta podseća na “One-time pad” perfektni tajni kripto sistem.

Komponenta 6 - predstavlja podatak koji je rezultat “XOR” operacije. Dobijeni podatak u sebi čuva zaključanu originalnu biometriju i slučajno generisani ključ. Na osnovu ovog podatka nije moguće otkriti originalnu biometriju ili ključ. Ovaj podatak zadovoljava formu podatka za skladištenje sa aspekta zaštite privatnosti biometrijskih podataka.

Komponenta 7 - predstavlja proces u kome se započinje sa procedurom u kojoj regularni korisnik želi da na osnovu originalne iris biometrije regeneriše ključ koji je zaključan čvrstom vezom originalne iris biometrije i ključa. Ovaj proces je identičan procesima u komponenti 3.

Komponenta 8 - predstavlja proces koji je identičan procesu u komponenti 4. Tehnika interlivinga je primenjena i nad originalnom iris biometrijom koja služi za regenerisanje ključa K . Napominjemo da u oba slučaja interliver radi samo u jednom smeru bez inverznog režima. Tačnije, u ovom ekvivalentnom kanalu ne postoji potreba za invertibilnosti.

Komponente 9 i 10 - predstavljaju dva poslednja procesa u dijagramu procesa, preko kojih se regeneriše ključ koji je zaključan čvrstom vezom u komponenti 5 na dijagramu. Zatim, ponovo se primenjuje logička operacija “XOR” nad kodovanim iris kodom i podatkom koji u sebi čuva ključ. Zatim se primenjuje Rid-Solomon dekodirer koji koriguje sve nastale greške. Ukoliko je dobijeni rezultat posle dekodiranja jednak ključu koji je generisan u procesu (1), to znači da je kompletna procedura realizovana uspešno i da smo napravili sistem koji ima mogućnost da na osnovu istog iris biometrijskog podataka generiše kriptološke ključeve.

4.2. Postavka eksperimentalnog okruženja

Pored detaljno izložene generičke šeme potrebno je obezbediti i kvalitetno eksperimentalno-razvojno okruženje koje je standardizovano za oblast u kojoj je predmet istraživanja, kao i u odnosu na koga se može kasnije uraditi evaluacija predloženog sistema sa već postojećim sistemima slične namene. Osnovni cilj standardizacije u ovom slučaju je da nam omogući preciznije sagledavanje veličine doprinosa u ovom istraživačkom radu.

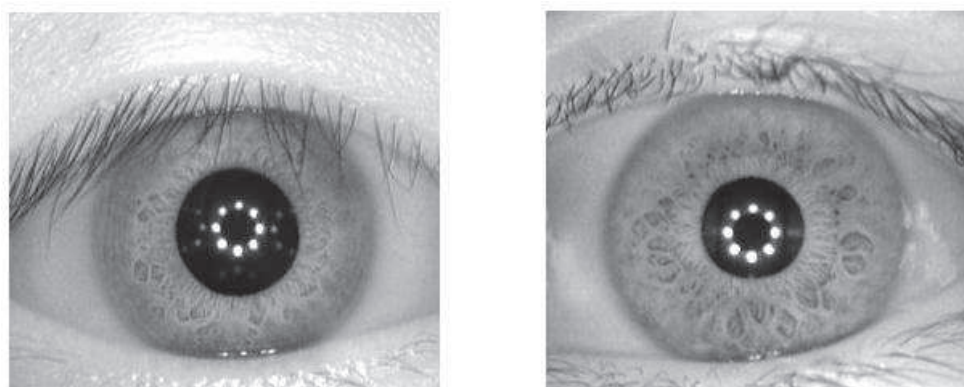
Eksperimentalno okruženje sastoji se iz dve komplementarne celine:

- programskog razvojnog okruženja;
- baze podataka sa eksperimentalnim podacima.

Za programsko-razvojno okruženje u ovom istraživačkom radu je izabran Matlab® programski paket. Matlab je nastao kao skraćenica za "MATrix LABoratory" ("laboratorija za matrice"). Izumeo ga je 1970-ih Kliv Moler (Cleve Moler), šef katedre za informatiku na Univerzitetu "Novi Meksiko". Omogućava pristup brzog razvoja aplikacija (eng. *RAD-rapid application development*) u cilju brzog generisanja eksperimentalnih rezultata. Matlab® je RAD okruženje i obezbeđuje odlično radno okruženje sa alatima za obradu podataka, vizualizaciju podataka i standardizovane algoritme koji se primenjuju u svim visokotehnološkim poljima. Pored mnogobrojnih algoritama nudi programiranje na visokom nivou. Svi programski kodovi nastali u ovom istraživačkom radu su razvijeni upravo u ovom softveru.

Eksperimentalne podatke koje smo koristili, dobili smo besplatno iz „CASIA“ baze podataka. „CASIA“ - baza podataka sadrži slike irisa, a tvorci ove baze su kineska istraživačka grupa sa Instituta za biometrijska istraživanja. Više verzija baze podataka je besplatno ponuđeno međunarodnoj zajednici za potrebe biometrijskih istraživanja. Više od 3000 korisnika iz 70 zemalja preuzelo je „CASIA“ bazu podataka sa slikama irisa i mnogo odličnih istraživačkih radova koji su priznati od međunarodne naučne zajednice su takođe nastali na osnovu slika ove baze podataka.

Od 1990-tih zabeležen je visok rast interesovanja na polju tehnologija namenjenih za prepoznavanja pojedinca na osnovu biometrije. Osnovni problem ili izazov je upotrebljivost i skalabilnost. Upotrebljivost je najveće usko grlo kod prepoznavanja na osnovu biometrije irisa. Prvi problem je realizovati jednostavnu (eng. *user friendly*) hardversku infrastrukturu za autentifikaciju korisnika. Drugi problem je široka primena ovih sistema koja zahteva visok stepen integracije, indeksiranje i skladištenje biometrijskih podataka u bazama podataka. Sve ovo vodi do problema skalabilnosti kod sistema za prepoznavanja korisnika na osnovu biometrije irisa. Do danas se pojavila i četvrta verzija baze podataka "CASIA-Iris V4" koja je korišćena za potrebe eksperimentalnog rada na ovoj disertaciji (**Slika 27**).



Slika 27. Primeri slika irisa iz „CASIA-Iris V4“ baze podataka

"CASIA-Iris V4" baza podataka u sebi sadrži ukupno 54.601. sliku irisa sa više od 1800. originalnih i 1000 virtuelnih subjekata. Sve slike irisa su 8-bitne u „JPEG“ formatu,

prikupljene akvizicijom specijalnih kamera sa infra-crvenim osvetljenjem. Bazu podataka može da koristi svako za potrebe istraživanja i edukativne svrhe.

4.3. Informaciono-teorijska analiza Iris biometrijskog izvora

U ovom podpoglavlju sprovedena je informaciono-teorijska analiza iris biometrijskog izvora. Polazeći od teorijskih osnova i Šenonovog koncepta perfektnih šifarskih sistema, izvršićemo rigoroznu analizu biometrije irisa u cilju kvalitetne sinteze optimizovanog sistema za ekstrakciju kriptoloških ključeva visokog kvaliteta na osnovu biometrijskih podataka. Predložena teorijsko-informaciona analiza biometrije kao svojevrsnog informacionog izvora davalala bi konkretna rešenja za konstruisanje biometrijskog kripto sistema sa teorijski garantovanim performansama. Iz tog razloga sprovedene su sledeće informacione analize:

- Analiza teksture slike irisa koja je dobijena nakon faze normalizacije;
- Analiza u cilju odabira optimalnih parametara vejevlet filtra u procesu kodovanja za formiranje kvalitetnog iris informacionog izvora;
- Analiza iris koda (biometrijskog obrazaca) u cilju odabira najkvalitetnijih segmenata irisa koji će da učestvuju u sintezi sistema;
- Modelovanje iris informacionog izvora primenom Šenonovih modela aproksimacije za modelovanje informacionih izvora;

4.3.1. Informaciona analiza teksture slike irisa

Analiza teksture slike irisa je sprovedena u prvom delu procesa za formiranje iris informacionog izvora (biometrijskog obrazaca). Osnovna ideja je da proverimo informacionu vrednost teksture slike irisa nakon faze normalizacije, ali pre faze kodovanja. Metod koji smo primenili na teksturu irisa koristi se za određivanje lokalne entropije. Metod je dobro poznat u oblasti za obradu slika (eng. *image processing*) i opisan u radu autora [41].

„Šenonova entropija“ je prvi put uvedena od strane Šenona 1948. Entropija u toj upotrebi termina se definiše kao očekivani informacioni sadržaj nekog događaja. Ako je poznata verovatnoća za neki događaj, onda se entropija može izračunati kao:

$$H = -\sum_{i=1}^l \log(p_i(x_i)) \quad (93)$$

Ili malo drugačijom notacijom:

$$H = -\sum_{i=1}^l \log(p_i(x_i)^{-1}), \quad (94)$$

gde se p odnosi na verovatnoće simbola histograma slike. U aplikacijama, koristi se logaritam osnove 2 zbog lakšeg rukovanja za binarnim informacionim sadržajem. Za posmatrane podatke dobijeni rezultat se može tumačiti kao relativna frekvencija događaja.

Lokalna entropija zavisi od izbora veličine prozora preko koga se određuje entropija svakog piksela. Broj prozora jednak je ukupnom broju piksela na slici (4800). Metod vraća niz entropijskih vrednosti H_i gde je $i=4800$. Svaki piksel je centar jednog prozora kome se posebno određuje vrednost entropije. Dobijene vrednosti izražene su preko logaritma sa osnovom 2, čime je određena jedinica mere za informaciju 1 bit. Izabrani metod nam dozvoljava da menjamo veličinu prozora koji je kvadratne dimenzije.

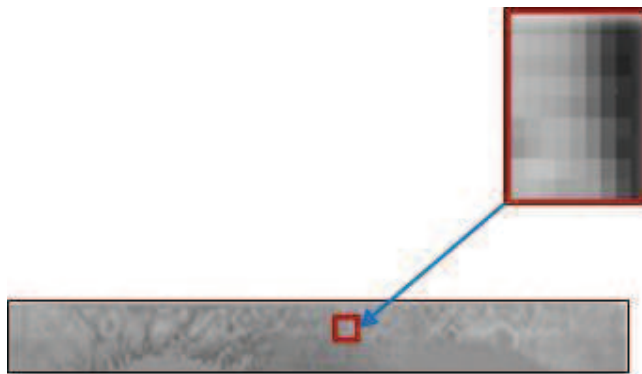
Menjanje veličine prozora dosta podseća na Šenonove modele aproksimacije koji se primenjuju za modelovanje informacionih izvora prirodnog jezika. Vrednost lokalne entropije se povećava povećanjem veličine prozora.

Međutim, sukcesivnom analizom je potvrđeno da entropija neko vreme raste nakon čega počinje da stagnira. Veličine prozora koje smo primenili su: $[3 \times 3]$, $[5 \times 5]$, $[7 \times 7]$ i $[9 \times 9]$ (**Slika 28**). Imajući u vidu rezoluciju slike, smatramo da nema efekta koristiti veću veličinu prozora od $[9 \times 9]$ iz razloga što se koristi simetrična dopuna pikselima koji će da reflektuju identične vrednosti entropije.

Za svaki segment određena je lokalna entropija primenom Šenonove aproksimacije prvog reda. Šenonova aproksimacija prvog reda podrazumeva određivanje frekvencija za intenzitete vrednosti piksela preko lokalnog histograma za 8-bitne slike irisa. Rezultat ove analize je entropijska vrednost pojedinačno svakog piksela na osnovu kojih tumačimo segmente površine iris koda.

Značaj ovo metoda i načina prikupljanja podataka na lokalnom nivou omogućiće nam prikazivanje 3D entropijskog modela slike irisa. Na ovaj način bićemo u mogućnosti da skoro vizuelno utvrdimo regione sa niskim entropijama koji ne pogoduju sintezi sistema za generisanje kriptološkog ključa.

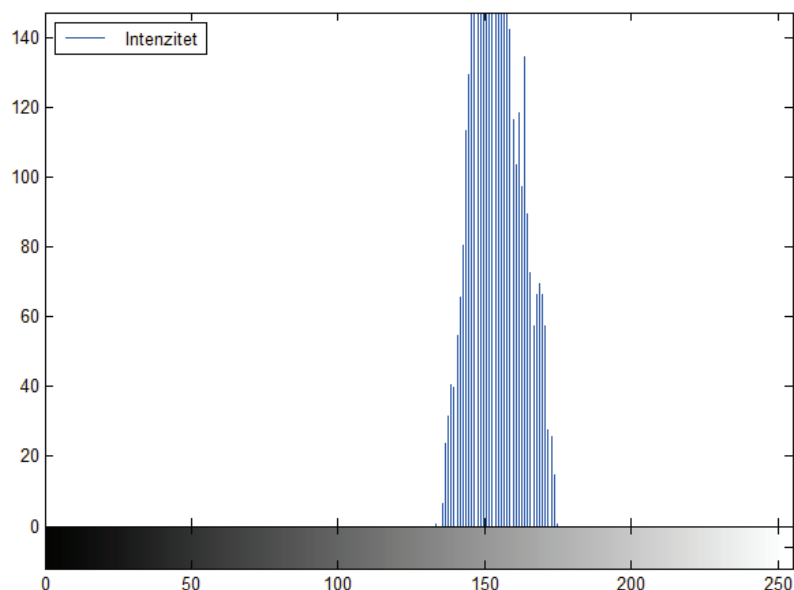
Slika 28., dat je razmotan iris kao rezultat faze segmentacije sa radijalnim parametrom za normalizaciju 20 piksela i ugaonim parametrom za normalizaciju 240 piksela, što čini rezoluciju slike 20×240 piksela. Dubina piksela je 8-bitna, time je svaki piksel predstavljen sa jednom vrednosti iz opsega od 0 do 255.



Slika 28. Računanje lokalne entropije nad slikom normalizovanog irisa

Metod ima mogućnosti primene u analizi više-dimenzijskih slika. Ukoliko tekstura slike ima više od dve dimenzije, pretpostavlja se da je u pitanju RGB standard, gde je dubina piksela 24-bita.

Da bismo se više približili razumevanju strukture irisa, izračunali smo za 100 slika irisa prosečnu vrednost intenziteta za korespondentne piksele na slici i prikazali rezultat na histogramu (**Slika 29**), kao i neke statističke pokazatelje, (**Tabela 2**).



Slika 29. Histogram prosečne vrednosti intenziteta piksela za sliku irisa

Ako pogledamo rezultate u tabeli, shvatićemo da je struktura irisa predstavljena sa intenzitetom piksela iz opsega od 134 do 175 i da raspodela poprilično odgovara Gausovoj normalnoj raspodeli.

Tabela 2. Statistički pokazatelji za dobijeni histogram

Statistički pokazatelji			
Minimalna vrednost	134	Aritmetička sredina	154.20
Maksimalna vrednost	175	Standardna devijacija	7.76

U sledećoj fazi računali smo lokalnu entropiju za dve veličine prozora $[3 \times 3]$ i $[9 \times 9]$ nad teksturom slike irisa. U merenju je učestvovao isti broj slika irisa kao prethodnom merenju. Nakon izračunatih lokalnih entropija za svaku sliku odredili smo prosečne vrednosti entropije za sve korespondentne piksele. Ukupno za svaki piksel ima 100-tinu korespondentnih

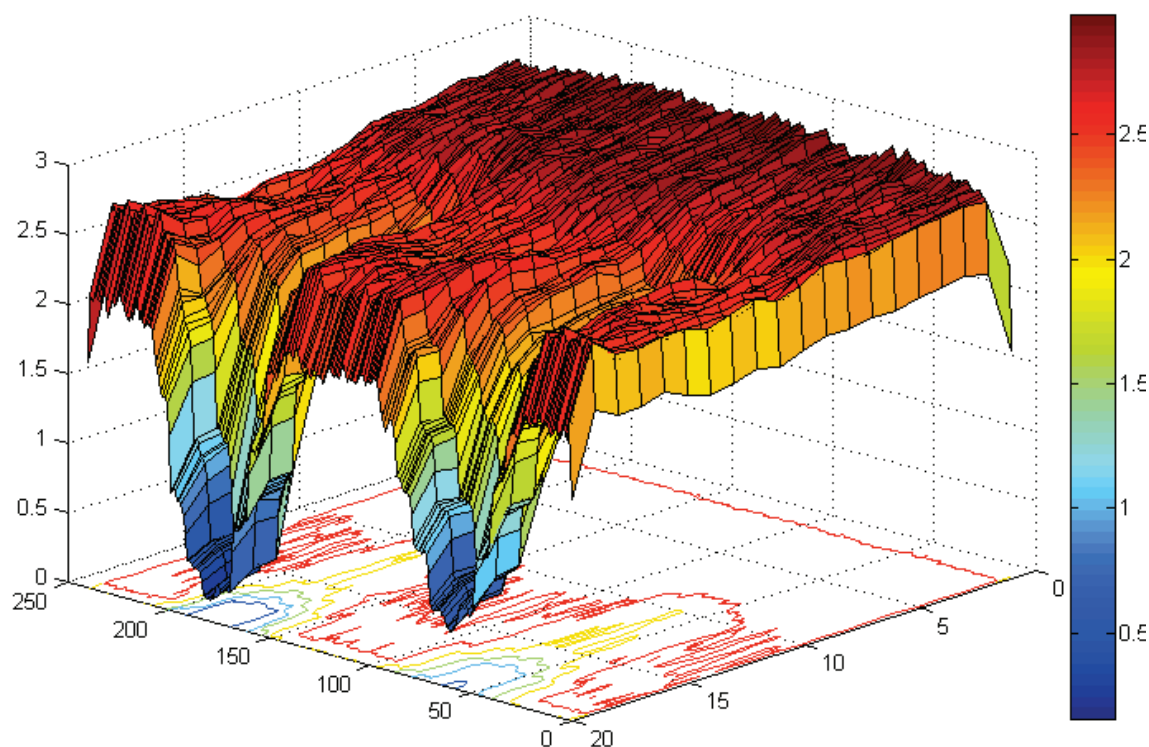
piksela. Rezultat merenja je predstavljen sa dve matrice dimenzije 20 x 240, vrednosti koje odgovaraju rezoluciji tekstone slike irisa (20 x 240).

Vrednosti u poljima matrice su prosečne entropijske vrednosti. Strukturu matrice smo iskoristili za modelovanje 3D entropijskog modela irisa. Rezultat ovog eksperimenta je predstavljen na plotovima (**Slika 30** i **Slika 31**).

Pravljenje 3D modela lokalne entropijske vrednosti je od velikog značaja za razumevanje informacionog kvaliteta površine irisa. Stručan naziv za ovaj prikaz rezultata je 3D kolor mapa (eng. *Heatmap*).

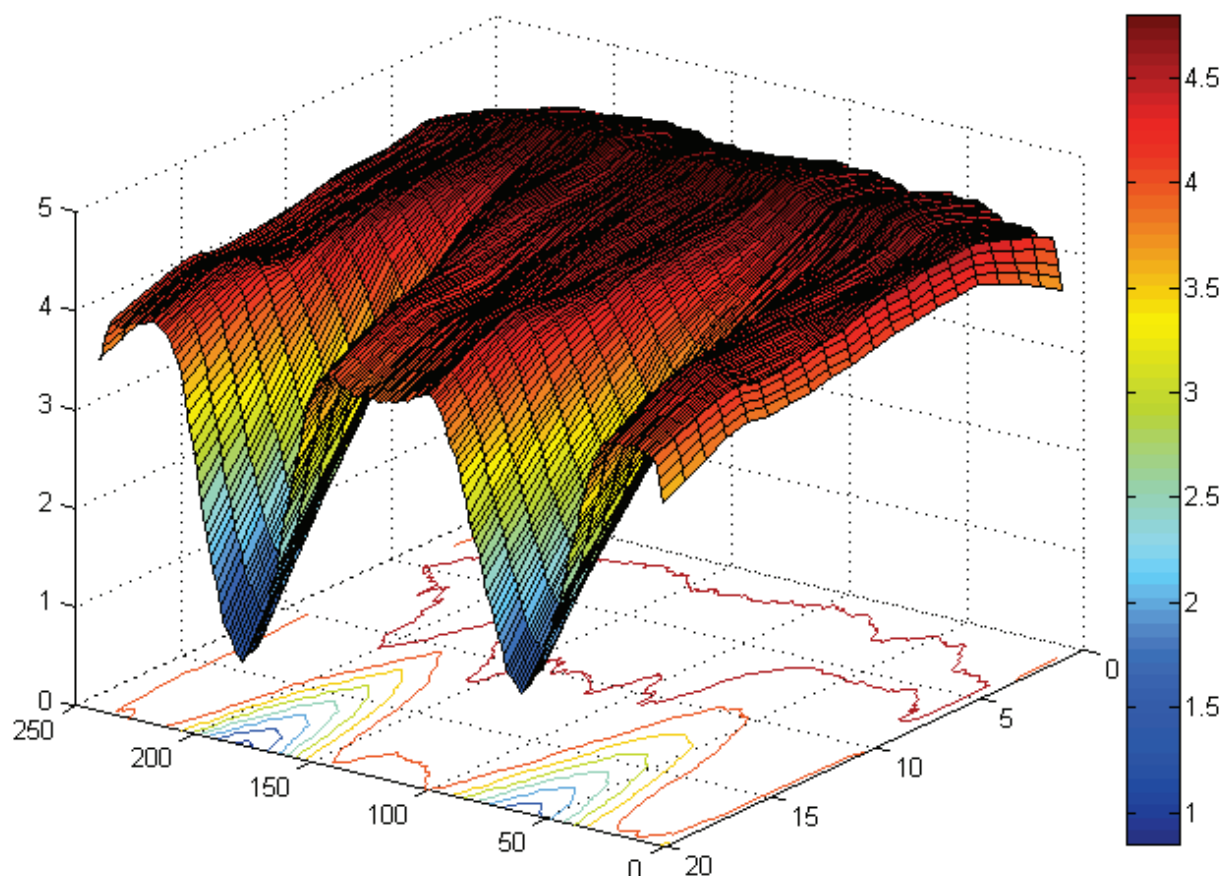
Prvi model predstavlja rigorozniju analizu jer je korišćena veličina prozora [3 x 3] i može da se tretira kao model na osnovu minimalnih entropijskih vrednosti, dok drugi model je manje rigorozan i modelovan na osnovu maksimalne entropije koja je zabeležena na teksturi slike irisa.

Na 3D entropijskom plotu (**Slika 30**), vidimo da prvi kružni segment irisa do zenice poseduje veću lokalnu entropiju od drugog kružnog segmneta irisa koji se graniči sa beonjačom. Posmatrajmo levu stranu na Y osi i videćemo da lokalna entropija nakon vrednosti 10 na osi Z naglo opada. Ovoj momenat je bio polazna osnova za dalju informacionu analizu tekstone irisa i njihovih međusobnih odnosa.



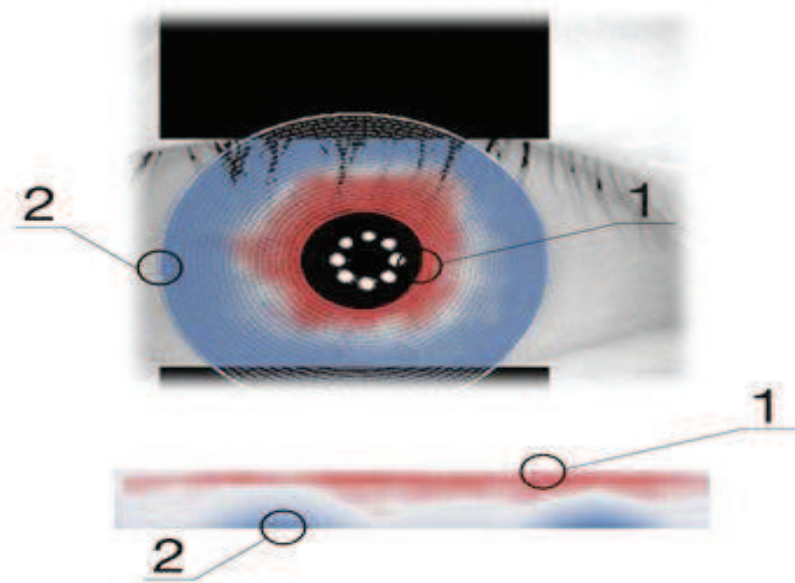
Slika 30. 3D informacioni model lokalne entropije (3x3)

Iako je prvi model dosta precizniji kada je upitanju lokalna entropija, drugi model (**Slika 31**) je dobijen na osnovu veličine prozora $[9 \times 9]$ za određivanje lokalne entropiju. Smatramo da ovaj model daje konačnu sliku o kvalitetu prvog segmenta u odnosu na kvalitet drugog segmenta. Na osnovu ovog modela možemo precizno da odredimo prostorne okvire prvog segmenta koji treba biti podvrgnut daljim informacionim analizama.



***Slika 31.** 3D informacioni model lokalne entropije (9x9)*

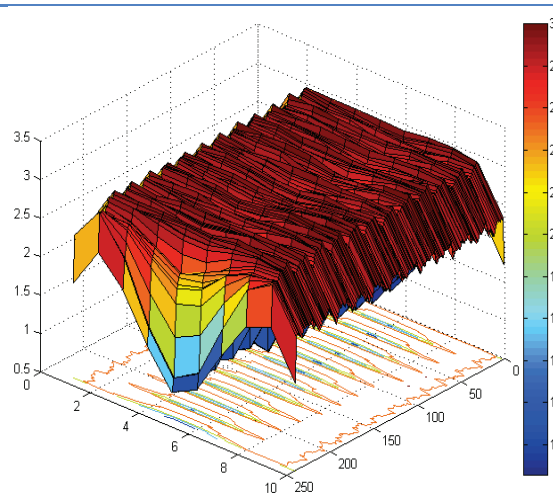
Na 3D entropijskom plotu (**Slika 31**), vidimo da je prethodna konstatacija tačna i da prvi kružni segment irisa do zenice poseduje veću lokalnu entropiju, što se jasno vidi na levoj strani na Y osi. Ovaj 3D model je dosta ilustrativan iz razloga što sada jasno vidimo segment irisa koji je značajan za nas. Ovu konstataciju potvrđujemo sa konturom koja se jasno ocrta u dvo-dimnzionalnom prostoru (X, Z). Kontura označena jakim crvenom bojom ocrta informaciono najkvalitetniji segment irisa, koji će kasnije biti ekstraktovan i korišćen u informacionoj analizi u cilju postavljanja teorijskih okvira ovog sistema za generisanje kriptoloških ključeva. Sa drsne strane plotu na osnovu toplotne skale možemo da se uverimo da su u tom segmentu zabeležene maksimalne vrednosti.



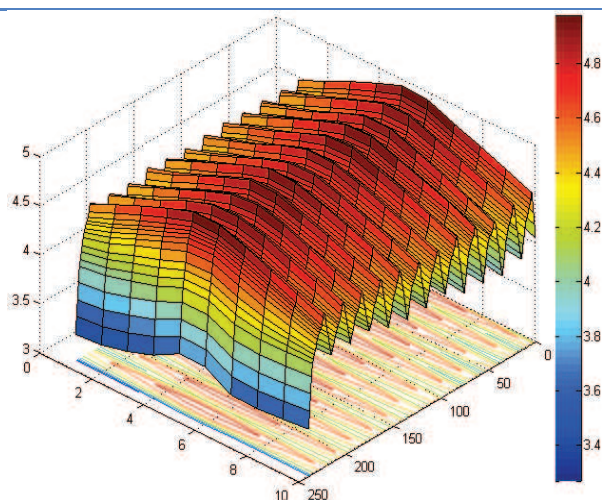
Slika 32. Vrednosti lokalne entropije po segmentima

Ovo saznanje je polazna osnova za nastavak informacione analize teksture irisa i njihovih međusobnih odnosa. U narednim analizama posebno su tretirani segmenti na osnovu podele koju smo do sada napravili (Slika 32). Prvi segment irisa predstavlja informativno najkvalitetniji segment, dok je u drugom delu zabeležen pad kvaliteta posmatran kroz rezultate dobijene preko lokalne entropije. Na kvalitet drugog segmenta dobrim delom utiče automatska faza segmentacije koja ne ispunjava uslov od 100% tačnosti. Iz tog razloga smatramo da ovaj segment ne može da predstavlja materijal za generisanje ključeva.

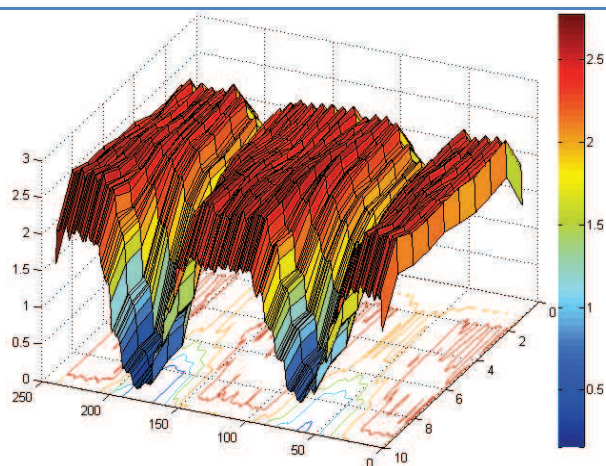
Dobijeni rezultati nas jasno usmeravaju na potrebu za podelom irisa po segmentima. U nastavku eksperimenta, sledi identična analiza koja se odnosi posebno na prvi i posebno na drugi segment irisa, čiji su informativni sadržaji prikazani u 3D entropijskom plotu.



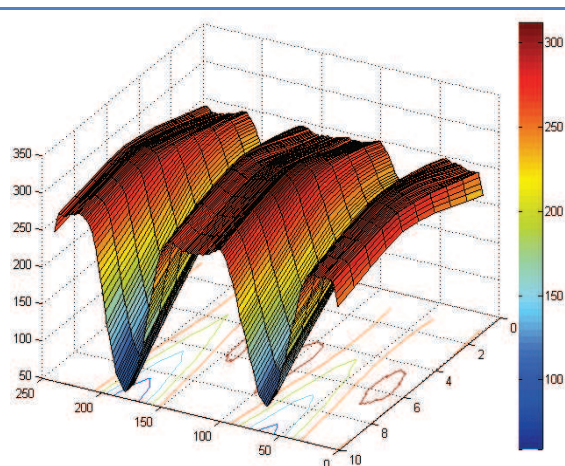
Slika 33. Lokalna entropija za prvi segment [3x3]



Slika 34. Lokalna entropija za prvi segment [9x9]



Slika 35. Lokalna entropija za drugi segment [3x3]



Slika 36. Lokalna entropija za drugi segment [9x9]

Tabela 3., sadrži dobijene rezultate za sve četiri veličine prozora na osnovu kojih je računata lokalna entropija. Prema dobijenim rezultatima zaključujemo da je veća entropija (neodređenost) izmerena u prvom segmentu, čija je maksimalna prosečna vrednost lokalne entropije po pikselu iznosila 4,4412 bita informacije, dok je u drugom segmentu prosečna vrednost entropije po pikselu dostigla 3,6020 bita informacije. U tabeli su prikazane vrednosti lokalne entropije.

Tabela 3. Prosečne vrednosti lokalne entropije za četiri veličine prozora

Veličina prozora (piksel)	[3 x 3]	[5 x 5]	[7 x 7]	[9 x 9]
Segment 1. bit/piksel	2.6540	3.6379	4.1267	4.4412
Segment 2. bit/piksel	2.1337	2.9149	3.3351	3.6020
Iris bit/piksel	2.4429	3.3526	3.8379	4.1405

Tabela 3. sadrži vrednosti lokalne entropije po segmentima i za ceo iris kod. Lokalna entropija je određena za sve četiri veličine prozora. Upotrebom najmanje veličine prozora dobijamo rigorozniju informacionu analizu, čime su vrednosti lokalne entropije znatno manje. Razlog za to je potpuno logičan, odabirom malog prozora analiza je bila detaljnija čime su u analizu uključeni uticaji podataka koji ne oslikavaju pravu strukturu iris koda. Vrednosti koje ćemo usvojiti za potrebe komparativnih odnosa između segmenata odnoseće se na veličinu prozora [9 x 9] (površina od 81 piksle). Povećanjem veličine prozora primećujemo da entropija polako stagnira i doseže svoje maksimume. Ukoliko bismo povećavali veličinu prozora, rezultat bi izgubio značaj jer bi tako veličina prozora prerasla veličinu slike irisa.

Tabela 4. Statistički pokazatelji po segmentima

Statistički pokazatelji				
Prvi segment	Minimalna vrednost	134	Aritmetička sredina	149.90
	Maksimalna vrednost	164	Standardna devijacija	5.59
Drugi segment	Minimalna vrednost	136	Aritmetička sredina	149.96
	Maksimalna vrednost	169	Standardna devijacija	5.38

Tabela 4. sadrži osnovne statističke pokazatelje za prvi i drugi segment. Prema podacima u tabeli ne dolazi do značajnijih odstupanja između ova dva segmenta osim kod standardne devijacije. U prvom segmentu je zabeležena nešto veća standardna devijacija. Standardna devijacija u uzorku nam govori koliko u proseku elementi uzorka odstupaju od aritmetičke sredine celokupnog uzorka.

4.3.2. Informaciona analiza iris koda

U narednoj informacionoj analizi posmatraćemo detaljno strukturu generisanog iris koda nakon faze kodovanja. U našem slučaju, data analiza predstavlja analizu iris informacionog izvora. Postoji nekoliko parametara od izuzetne važnosti u algoritmu za kodovanje irisa. Rezultat faze kodovanja predstavlja iris kod ili informacioni izvor. Za uspešnost ove faze potrebno je odrediti optimalne vrednosti parametara koji učestvuju u fazi kodovanja kako bi se obezbedila najveća entropija iris koda.

Ovi parametri podrazumevaju: radijalnu i ugaonu rezoluciju (r i θ) odnosno, parametri koji daju broj tačaka na slici irisa koje će se kodovati u svakom irisu, kao i parametri filtra za kodvanje samo jedinstvenih karakteristika sa slike irisa.

Parametri filtra uključuju: broj filtra N , talasnu dužinu λ_n (u pikselima) (eng. *wavelength*), opseg propusnosti filtra (eng. *bandwidths*) koji je predstavljen sa σ / f i multiplikovani faktor između vrednosti centara talasne dužine i uzastopnog filtra α .

Analiza koja sledi je sprovedena u delu procesa u kome se formira iris informacioni izvor (iris kod). Zapravo, ovo je faza generisanja iris biometrijskog obrasca koji je u našem slučaju posmatran kao iris informacioni izvor.

U ovom merenju smo menjali vrednosti parametra za talasnu dužinu λ_n filtra. Na osnovu uvida u detalje rezultata koji su dati u radu [39 str. 32], procenili smo da je ovaj parametar od izuzetne važnosti za naš sistem. Kada kažemo da je od izuzetne važnosti, tada

podrazumevamo da promenom ovog parametra možemo da utičemo na povećanje entropije izvora.

U cilju potvrđivanja ove teze, sprovedi smo iscrpnu analizu nad istim uzorkom iris kodova kao u prethodnim analizama. Za svaki iris iz grupe uzoraka je generisan iris kod za jednu od tri predložene vrednosti parametra $\lambda_n = \{12, 18, 24\}$.



Slika 37. Iris kod interpretiran kao crno-bela slika

Slika 37., predstavlja iris kod koji je generisan nakon faze kodovanja. Korišćena je maska za maskiranje delova iris koda koji sadrže greške, a koje su izazvane sa trepavicama, kopcima i drugim uzrocima.

Rezultati koji slede dobijeni su na osnovu slične metodologije koja je primenjena u prethodnoj informacionoj analizi. Sada iris kod posmatramo na nivou binarnih matrica (iris biometrijski obrazac) nad kojima računamo lokalnu entropiju. Dimenzija kompletnog iris koda je 20 x 480, a dimenzija po segmentima je 10 x 480.

Posmatraćemo iris kod u celini i po segmentima za izabranu vrednost parametara λ_n ovog filtra. Više o fazi kodovanja napisano je u delu rada 3.2.

Tabela 5. Prostiranje lokalne entropije u iris kodu za parametar filtra ($\lambda_n = 12$)

$\lambda_n = 12;$	Lokalna entropija H				Toplotne mape
	3x3	5x5	7x7	9x9	
Ceo iris kod	0.7005	0.7813	0.8176	0.8412	

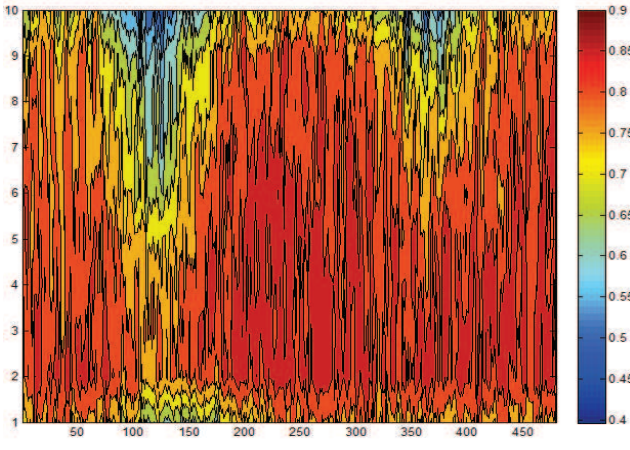
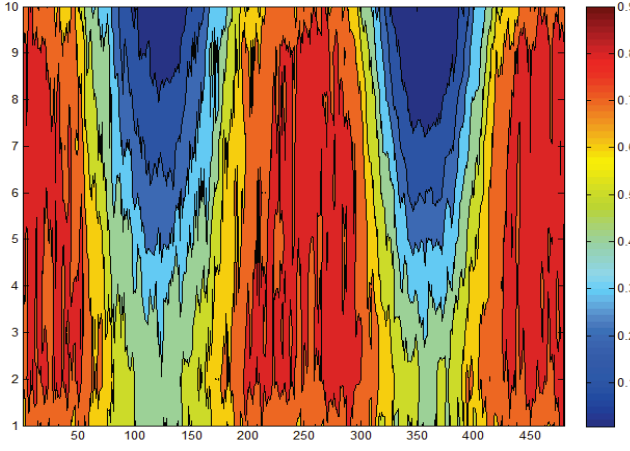
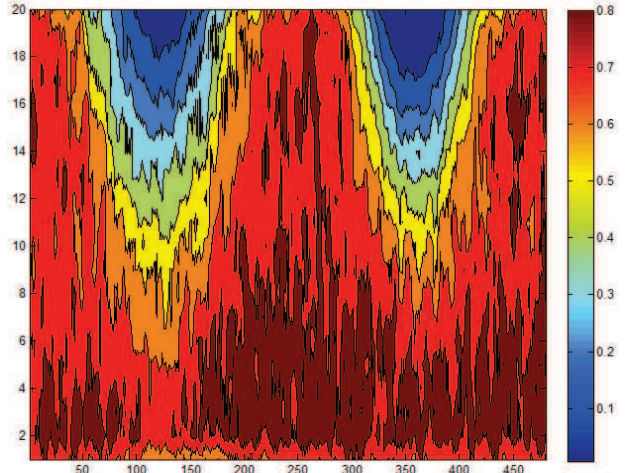
Prvi segment iris koda	0.7971	0.8828	0.9163	0.9351	
Drugi segment iris koda	0.5905	0.6647	0.7008	0.7251	

Tabela 6. Prostiranje lokalne entropije u iris kodu za parametar filtra ($\lambda_n = 18$)

$\lambda_n = 18;$	Lokalna entropija H				Toplotne mape
	3x3	5x5	7x7	9x9	
Ceo iris kod	0.6653	0.7515	0.7923	0.8189	

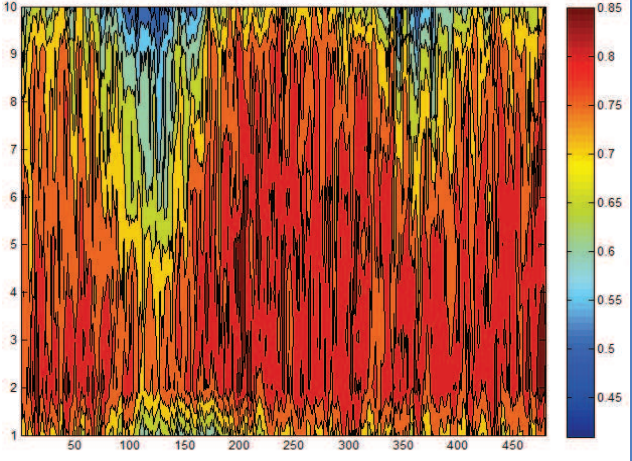
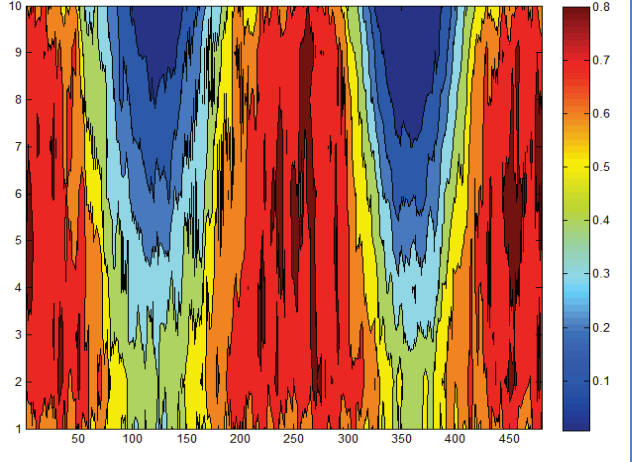
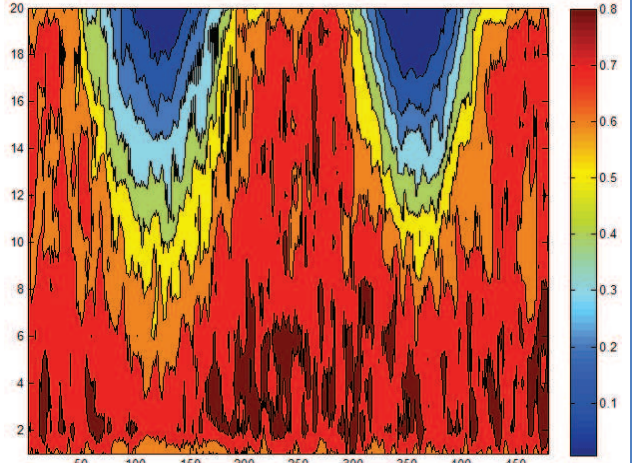
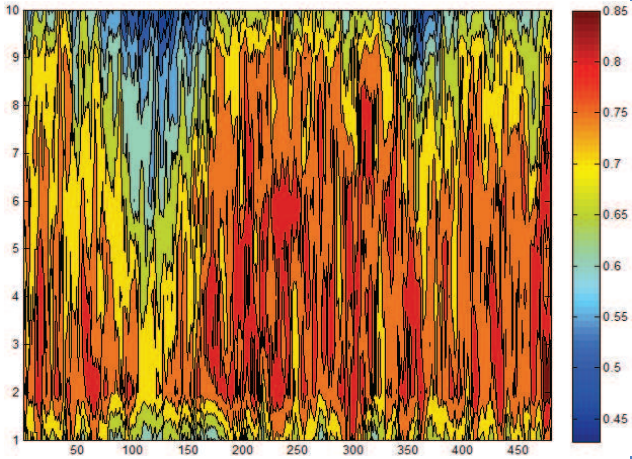
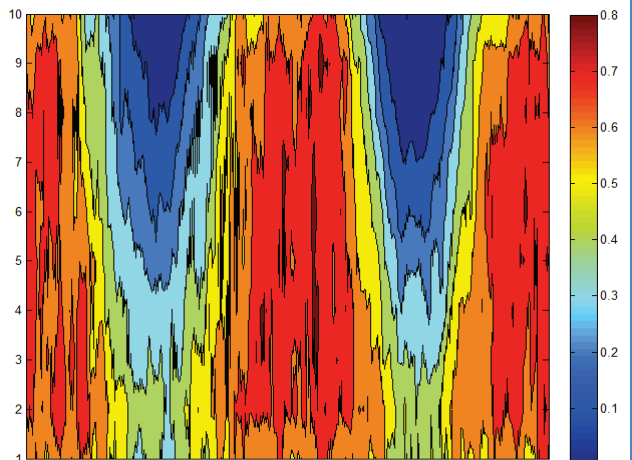
Prvi segment iris koda	0.7606	0.8523	0.8906	0.9125	
Drugi segment iris koda	0.5564	0.6335	0.6729	0.6997	

Tabela 7. Prostiranje lokalne entropije u iris kodu za parametar filtra ($\lambda_n = 24$)

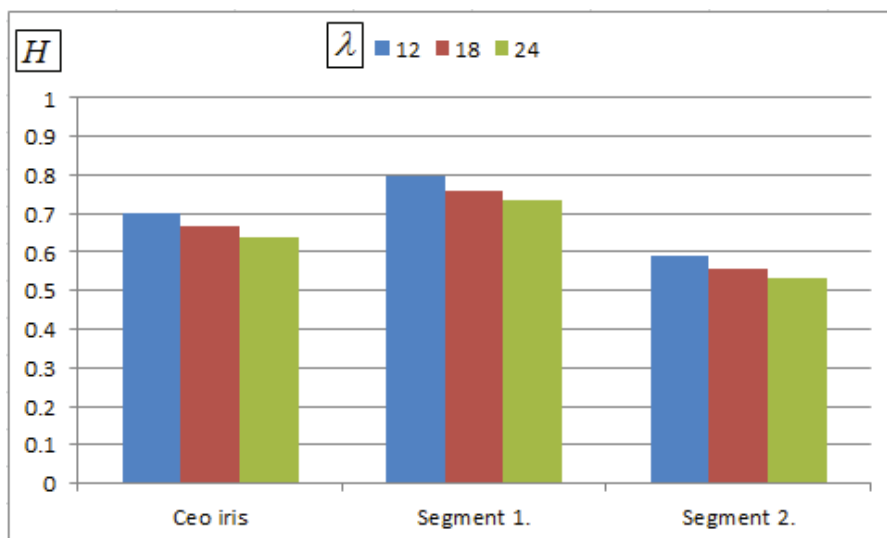
$\lambda_n = 24;$	Lokalna entropija H				Toplotne mape
	3x3	5x5	7x7	9x9	
Ceo iris kod	0.6390	0.7267	0.7698	0.7982	

Prvi segment iris koda	0.7330	0.8259	0.8663	0.8901	
Drugi segment iris koda	0.5316	0.6092	0.6498	0.6776	

U ovoj analizi, iris kod smo posmatrali kao 2D crno-belu sliku sa dubinom piksela 1 bit. Na osnovu njene tekture, računali smo lokalnu entropiju koja podrazumeva izračunate vrednosti svih korespondentnih piksela. Ova mera je za nas vrlo značajna iz razloga što želimo da utvrdimo stabilne ili konzistentne segmente irisa sa najvećom entropijom.

Pogledajmo sada rezultate predstavljene preko 2D toplotne mape koja predstavlja entropijski model prostiranja lokalne entropije za parametar $\lambda = 12$ (Tabela 5). U tabeli su date tri toplotne mape za tri segmenta i četiri veličine prozora preko kojih je računata lokalna entropija. U drugom redu tabele, prikazane su vrednosti entropije za prvi segment i vrednosti lokalnih entropija. Na toplotnoj mapi za prvi segment irisa primećujemo da je izabran informaciono najkvalitetniji segment sa ujednačenim vrednostima lokalne entropije preko cele površine.

Cilj ove analize je bio da izaberemo optimalni parametar filtra koji direktno utiče na povećanje ekstrakta entropije iris informacionog izvora. Za potrebe narednih analiza i sintezu celokupnog sistema koristićemo vrednost parametra vežvlet filtra $\lambda = 12$. U sledećoj analizi primenom Šenonovih modela aproksimacije utvrdićemo maksimalnu entropiju po segmentima u iris kodu.



Slika 38. Vrednosti lokalne entropije za različite vrednosti parametra λ

Slika 38., sadrži vrednosti lokalne entropije (veličina prozora [3 x 3]) za ceo iris kod, prvi segment i drugi segment iris koda. Na grafiku se vidi da je entropija u drugom segmentu imala najmanju vrednosti dok je u prvom segmentu postignuta najveća vrednost entropije. U pogledu izbora parametra vejavlet filtra, najveća vrednost entropije u iris kodu je postignuta sa parametrom filtra $\lambda = 12$.

4.3.3. Šenonovi modeli aproksimacije za modelovanje iris inf. izvora

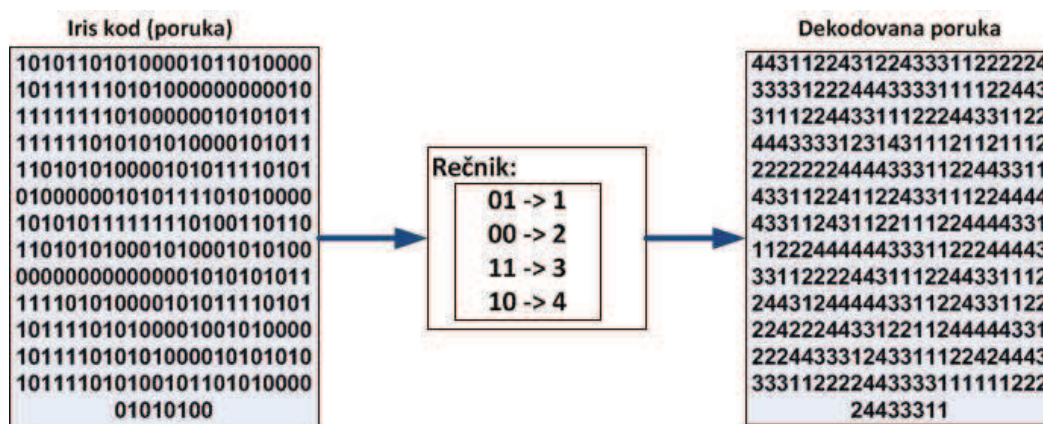
U ovom delu rada primenićemo Šenonove modele aproksimacije o kojima je bilo više reči u poglavlju 3.1. Celokupan metod uključuje algoritme koji su posebno razvijeni za aproksimacije od II do V reda. U ovom metodu vrlo je bitna priprema ulaznih podataka. Nakon kodovanja iris informacionog izvora, dobijen je binarni iris kod kao matrica sa 480 kolona i 20 redova. Obzirom da je ispitivan u ovoj analizi samo prvi segment, odbačena je druga polovina matrice i dobijena je nova matrica sa 480 kolona i 10 redova. U fazi kodovanja matrica je formirana iščitavanjem red po red. Prvi red predstavlja prvi prsten irisa do zenice i tako dalje do 20-tog prstena. Ovo je vrlo bitno jer podatak koji se posmatra kroz modele aproksimacije je 1D signal. Iz tog razloga pretvaranje prvog 2D dela matrice (segment 1.) u 1D signal treba obaviti po redovima matrice. Ovaj proces možemo definisati sledećim izrazom:

$$1D_{signal} = r1(2D_{signal}) + r2(2D_{signal}) + \dots + r10(2D_{signal}),$$

gde r predstavlja red 2D matrice iris koda.

Sada kada su spremni ulazni podaci, daćemo objašnjenje algoritma za aproksimaciju II reda koja je u potpunosti slična i sa narednim modelima aproksimacije. U ovoj fazi generisani 1D signal je tretiran kao kodovana poruka koju je potrebno dekodovati na osnovu zadatih kodnih reči. Svaka kodna reč odgovara jednoj poruci. Ceo algoritam pomalo podseća na Hafmanov algoritam za entropijsko kodovanje sa rečnikom koji je generisan na osnovu verovatnoće pojavljivanja simbola u poruci.

Za razliku od entropijskog kodovanja, u ovom algoritmu su pretpostavljani svi simboli sa jednakim verovatnoćama i obavlja se samo faza kodovanja. U aproksimaciji II reda ukupan broj poruka je neki skup $S = \{x_1, x_2, x_3, x_4\}$, gde je svaka poruka dekodovana sa kodnim rečima [00,10,11,01]. Sledi primer dekodovanja iris koda primenom aproksimacije II reda (Slika 39.).



Slika 39. Primer za aproksimaciju II reda

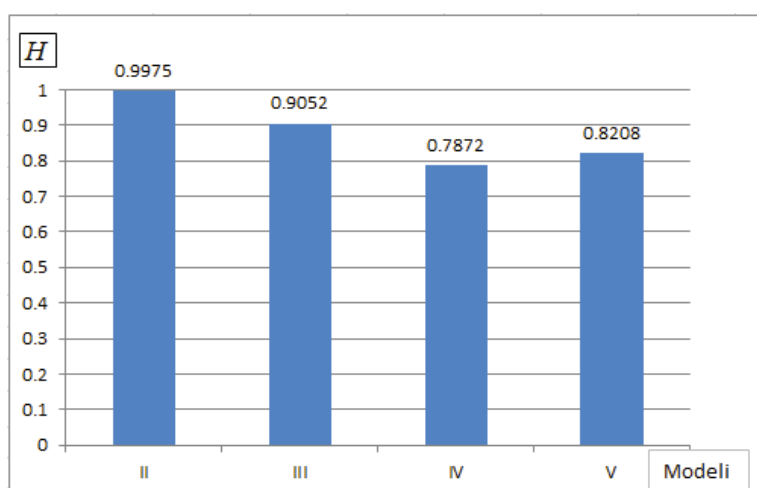
Tabela 8. Modelovanje iris informacionog izvora primenom Šenonovih modela aproksimacije

Modeli iris informacionog izvora na osnovu Šenonovih modela aproksimacija za prvi segment irisa (4800 bitova)							
Aprok. II reda	p	Aprok. III reda	p	Aprok. IV reda	p	Aprok. V reda	p
[00]	0.2827	[010]	0.2302	[0000]	0.1805	[01010]	0.1338
[10]	0.2446	[101]	0.2284	[1111]	0.1467	[10101]	0.1328
[11]	0.2445	[000]	0.1863	[1010]	0.1467	[00000]	0.1062
[01]	0.2282	[111]	0.1497	[0101]	0.1283	[11111]	0.0729
		[100]	0.0521	[0100]	0.0936	[11010]	0.0469
		[001]	0.0520	[0010]	0.0917	[10100]	0.0464
		[011]	0.0509	[1101]	0.0909	[00101]	0.0458
		[110]	0.0502	[1011]	0.0903	[01011]	0.0452
				[0001]	0.0068	[01000]	0.0404
				[1000]	0.0051	[00010]	0.0396
				[1110]	0.0037	[00001]	0.0395
				[1100]	0.0036	[11101]	0.0390
				[0111]	0.0035	[10000]	0.0386
				[0011]	0.0034	[10111]	0.0382
				[0110]	0.0024	[11110]	0.0381
				[1001]	0.0024	[01111]	0.0373
						[10010]	0.0110
						[01001]	0.0106
						[01101]	0.0101
						[10110]	0.0098
						[00100]	0.0041
						[11011]	0.0028
						[10001]	0.0020
						[01110]	0.0015
						[00110]	0.0011
						[11000]	0.0010
						[01100]	0.0010
						[11100]	0.0009

						[00011]	0.0009
						[10011]	0.0008
						[11001]	0.0008
						[00111]	0.0006
Entropija H po modelima aproksimacije izvora (bit)							
1.9955		2.7156		3.1490		4.1044	
Entropija H po 1 bitu iris koda (bit)							
1 bit = 0.9975		1 bit = 0.9052		1 bit = 0.7872		1 bit = 0.8208	
Ukupna entropija H u prvom segmentu irisa (bit)							
4788		4345		3779		3940	

Tabela 8., sadrži rezultate entropije za II, III, IV i V model aproksimacije informacionog izvora. Najmanja vrednost entropije po jednom bitu u iris kodu je zabeležena u IV-toj aproksimaciji $H_{IV} = 0.7872$ bita informacije. Utvrđeno je da prvi segment irisa nama najznačajniji segment ukupno raspolaže sa 3940 bitova informacije. **Slika 40.**, dat je bolji pregled dobijenih rezultata.

Pored entropijskih vrednosti, prikazane su verovatnoće kodnih reči p_i (od najveće do najmanje) za sve modele aproksimacije. Ako bolje pogledamo, videćemo da su kodne reči svih modela podeljene na dva dela na osnovu njihovih vrednosti verovatnoća. Prvi deo kodnih reči sadrži visoke verovatnoće, dok drugi deo i po nekoliko puta manje verovatnoće u zavisnosti od reda. Iz ovog razloga izabrali smo V-ti model aproksimacije, jer verovatnoće polovine poruka iznad V modela imaju vrednost 0.



Slika 40. Entropija prvog segmenta irisa po modelima aproksimacije

4.3.4. Analiza uzajamne informacije između istih i različitih irisa

Iz prethodne analize smo saznali vrednosti entropije u iris kodu, kao i optimalne vrednosti parametara filtra. Iris kod je prvo analiziran kao celina, a zatim po segmentima. U analizi koja sledi, želimo da odredimo uzajamnu informaciju između istih i različitih irisa za prvi segment (**Slika 32**). Ova analiza je vrlo značajna sa aspekta bezbednosti za sintezu sistema. O bezbednosnim aspektima ćemo više govoriti u narednim poglavljima.

Merenje će biti obavljeno na podacima koji predstavljaju dekodovani iris kod za sva četiri modela aproksimacije modelovanog iris informacionog izvora. Očekivani rezultat u ovoj analizi je da segment (prvi segment) sa najvećom lokalnom entropijom kod istih irisa sadrži maksimalnu uzajamnu informaciju, dok je kod različitih irisa očekivano da sadrži minimalnu uzajamnu informaciju.

Metod koji koristimo za merenje uzajamne informacije $I(irisX; irisY)$ između dva iris koda je definisan sledećim izrazom:

$$I(A, B) = \sum_{b \in B} \sum_{a \in A} p(a, b) * \log \left(\frac{p(a, b)}{p(a)p(b)} \right) \quad (95)$$

gde je:

- $p(a, b)$ - zajednička verovatnoća funkcije raspodele od A i B ;
- $p(a)$ - marginalna verovatnoća funkcije raspodele od A ;
- $p(b)$ - marginalna verovatnoća funkcije raspodele od B .

U pogledu teorije verovatnoće, relativna entropija sistema meri razdaljinu između dve različite funkcije verovatnoće. Uzajamnu informaciju sada u ovom aspektu možemo jednostavno definisati:

$$\begin{aligned} I(A; B) &= H(A) + H(B) - H(A, B); \\ I(A; B) &= H(A) - H(A|B) = \\ &= H(B) - H(B|A). \end{aligned} \quad (96)$$

gde je:

- $H(A)$ - marginalna entropija od A ;
- $H(B)$ - marginalna entropija od B ;
- $H(A|B)$ - uslovna entropija za A ;
- $H(B|A)$ - uslovna entropija za B ;
- $H(A, B)$ - zajednička entropija za A i B .

Za ove potrebe razvijen je algoritam za određivanje uzajamne informacije između dva irisa. Listing programskog koda je dat u dodatku A na kraju rada.

Tabela 9. Prosečna uzajamna informacija između različitih irisa

Prosečne uzajamne informacije između različitih irisa za zadate modele Šenonove aproksimacije				
Šenonovi modeli aproksimacije (II, III, IV, V)	Broj simbola poruke	Uzajamna informacija (simbol/bit)	Uzajamna informacija (bit/bit)	Uzajamna informacija (bit/segment 1. iris kod)
Aprok. II reda	4	0.0071	0.0035	16.8
Aprok. III reda	8	0.0474	0.0158	75.84
Aprok. IV reda	16	0.1083	0.0270	129.6
Aprok. V reda	32	0.5325	0.1065	511.2

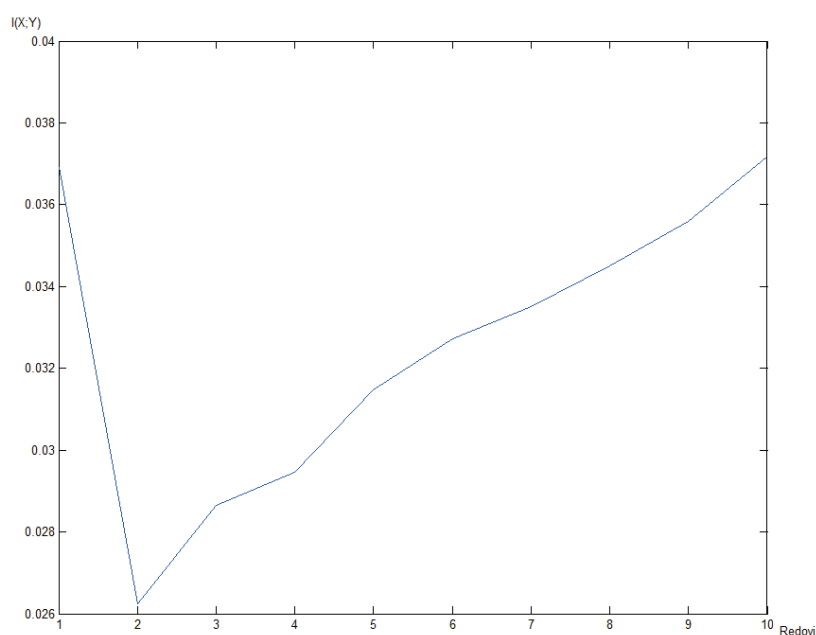
Tabela 9., sadrži prosečnu uzajamnu informaciju između različitih irisa. Uzajamna informacija je izračunata za svaki Šenonov model aproksimacije. U merenju je učestvovalo 100 irisa sa ukupno 4950 komparativnih merenja. U ovom slučaju, kao merodavan rezultat posmatračemo vrednost uzajamne informacije za aproksimaciju V-tog reda. Ukupna količina uzajamne informacije po jednom bitu između dva različita iris koda u prvom segmentu je u proseku $I(X;Y)=0.1065$, dok je ukupna količina uzajamne informacije u prvom segmentu u proseku $I(X;Y)=511$ bitova.

U sledećem delu eksperimenta merićemo uzajamnu informaciju između istih irisa, tačnije između irisa koji pripadaju jednoj osobi. „CASIA“ baza podataka obezbeđuje 10 slika irisa za jednu osobu. Merenje je obavljeno na 50 slučajno izabranih osoba. Dobijena uzajamna informacija je određena na osnovu komparativne analize 45 merenja za svaku osobu. Rezultati uzajamne informacije za svih 50 osoba su uprosečeni.

Tabela 10. sadrži prosečnu uzajamnu informaciju između irisa koji pripadaju istoj osobi. Uzajamna informacija je izračunata za svaki Šenonov model aproksimacije. U ovom slučaju kao merodavan rezultat posmatračemo vrednost uzajamne informacije za aproksimaciju V-tog reda. Ukupna količina uzajamne informacije po jednom bitu između dva iris koda u prvom segmentu irisa je u proseku $I(X;Y)=0.2078$, dok je ukupna količina uzajamne informacije na prvom segmentu u proseku $I(X;Y)=997$ bitova.

Tabela 10. Prosečna uzajamna informacija između istih irisa

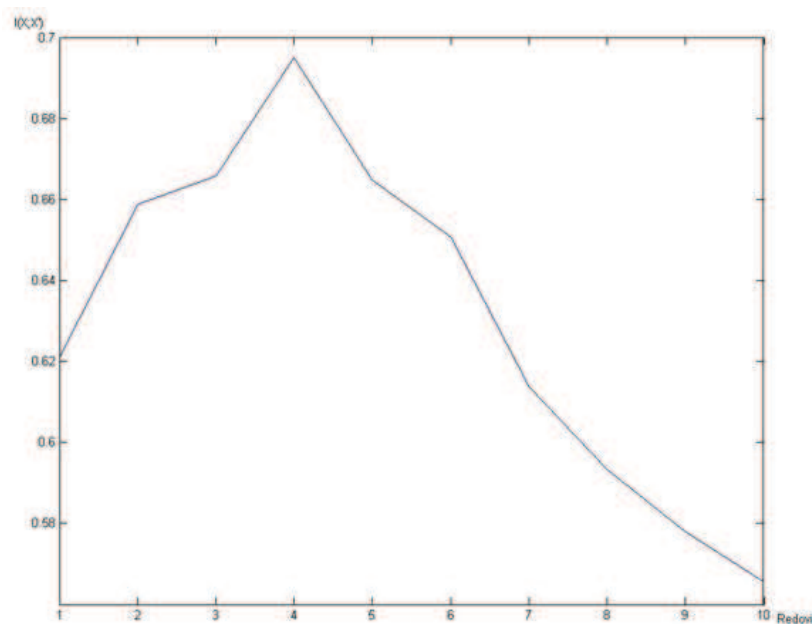
Prosečne uzajamne informacije između istih irisa za zadate modele Šenonove aproksimacije				
Šenonovi modeli aproksimacije (II, III, IV, V)	Broj simbola poruke	Uzajamna informacija (simbol/bit)	Uzajamna informacija (bit/bit)	Uzajamna informacija (bit/segment 1.)
Aprok. II reda	4	0.3001	0.1500	720.28
Aprok. III reda	8	0.3761	0.1253	601.82
Aprok. IV reda	16	0.5488	0.1372	658.65
Aprok. V reda	32	1.0391	0.2078	997.55



Slika 41. Uzajamna informacija po redovima u iris kodu za prvi segment (različiti irisi)

Slika 41., predstavlja uzajamnu informaciju između redova u prvom segmentu različitih irisa. Ukupan broj redova je 10. Svaki od redova odgovara jednom kružnom prstenu irisa. Minimalna uzajamna informacija je zabeležena između 2 i 5 reda. Pronalazak segmenta irisa u kome postoji najmanja zavisnost između susednih redova predstavlja vrlo značajan rezultat sa aspekta kriptanalize. Podsetićemo se da je 3D modelom lokalne entropije (**Slika 31**) detektovan segment sa najvećom lokalnom entropijom.

Odnos dobijenih vrednosti za uzajamne informacije, entropije i lokane entropije je informaciono-teorijski korektan. Sa ovom pretpostavkom zaključujemo da u delovima irisa u kojima je lokalna entropija izražena, uzajamna informacija će težiti minimalnim vrednostima.



Slika 42. Uzajamna informacija po redovima u iris kodu za prvi segment (isti irisi)

Slika 42., predstavlja uzajamnu informaciju između redova u prvom segmentu istih irisa. Ukupan broj redova je 10. Svaki od redova odgovara jednom kružnom prstenu irisa. Maksimalna uzajamna informacija je zabeležena između 2 i 6 reda. Odnos dobijenih vrednosti za uzajamne informacije i lokane entropije je informaciono-teorijski korektan. Sada se sa velikom verovatnoćom očekuje da ovaj deo iris koda sadrži podatke sa najkonzistentnijim informacijama. U svakom slučaju, dobijeni rezultati nakon evaluacije predloženog sistema će potvrditi ili odbaciti ovu pretpostavku.

Estimacijom entropije, lokalne entropije i uzajamne informacije, identifikovani su segmenti irisa koji su najpogodniji za ove potrebe. Izvršena je optimizacija parametara odgovarajuće vejevlet transformacije u cilju dobijanja što veće entropije i što niže uzajamne informacije u transformacionom domenu.

$H(K)$ - entropija ključa ili veličina ključa;

$I(Y;Y')$ - informacija između irisa iste osobe;

$I(X;Y)$ - informacija između irisa različitih osoba;

$H(X)$ - entropija iris koda;

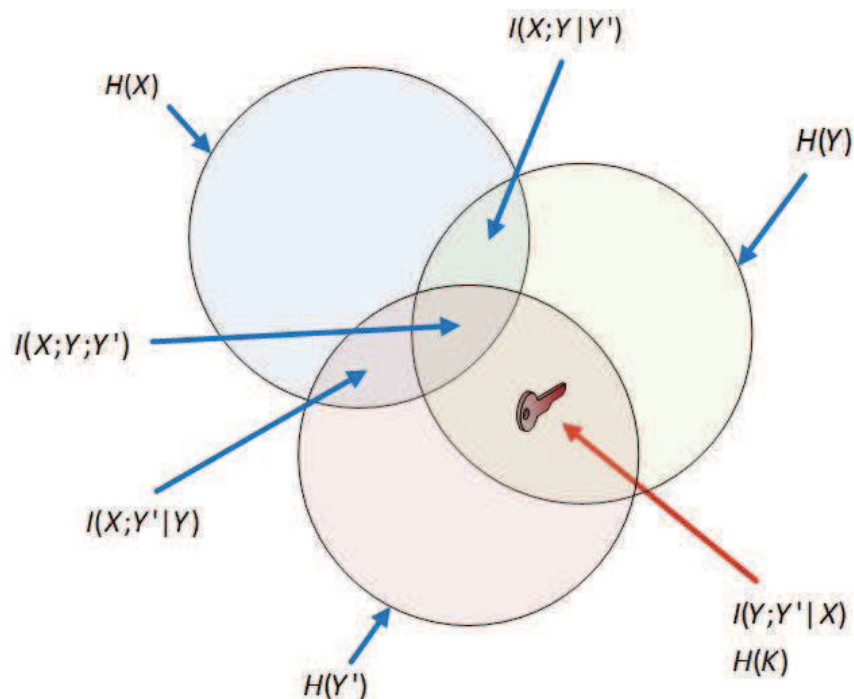
$H(X,Y)$ - entropija dva iris koda različitih osoba;

$H(Y, Y')$ - entropija dva iris koda iste osobe;

$H(K)_{\max}$ - maksimalna veličina ključa.

$$\begin{aligned}
 H(IRIS X) &= H(IRIS Y) = H(IRIS Y') = 3940; \\
 I(X; Y) &= 511; \\
 I(Y; Y') &= 997; \\
 H(X, Y) &= 7281; \\
 H(Y, Y') &= 6709; \\
 I(Y; Y' | X) &= 788; \\
 H(K) &= I(Y; Y' | X) = 788; \\
 I(X; Y; Y') &= I(Y; Y') - I(Y, Y' | X) = 997 - 788 = 209;
 \end{aligned}
 \tag{97}$$

Maksimalna entropija ključa iznosi $H(K) = 788$ bitova, čime su postavljeni okviri za sintezu sistema za ekstrakciju čisto slučajnih nizova iz iris biometrije.



Slika 43. Grafička prezentacija prostora biometrijskog ključa

Maksimalna veličina kriptološkog ključa je određena izrazom:

$$H(K)_{\max} = I(Y; Y' | X); \tag{98}$$

Slika 43., predstavlja grafičku prezentaciju pozicije biometrijskog ključa. Dijagram na slici sadrži tri slučajne promenljive X, Y i Y' . Važno je napomenuti da je uzajamna informacija $I(X; Y; Y')$ simetrična ukoliko su ove tri promenljive jednako nezavisne. Imajući u vidu da u ovom dijagramu postoji veća zavisnost između promenljive Y i Y' , time je uzajamna

informacija $I(Y;Y')$ veća. Na osnovu ove činjenice i informacija $I(Y;Y'|X)$ je veća. U ovom slučaju pojedinačne uzajamne informacije $I(X;Y)$ i $I(X;Y')$ su se smanjile, dok se zajednička informacija $I(X;Y;Y')$ i $I(Y;Y'|X)$ povećala. Na dijagramu koji sledi pokušali smo da prikazemo vrednosti u zadatoj proporciji.

4.3.5. Kompliksnost iris koda - DOF

Kompleksnost iris koda moguće je približno odrediti merenjem broja stepeni slobode – DOF (eng. *degree of freedom*) nad skupom iris kodova. Ova veličina daje meru kompleksnosti iris koda i može se izračunati aproksimacijom čitave kolekcije različitih iris kodova, preko svih međusobnih Hamingovih rastojanja kao binomijalna distribucija. Broj stepeni slobode DOF definiše se izrazom:

$$DOF = \frac{m(1 - m)}{\sigma^2} \quad (99)$$

gde je m aritmetička sredina, a σ standardna devijacija date distribucije.

DOF – broj stepeni slobode može biti definisan kao minimalni broj nezavisnih koordinata koje mogu u potpunosti odrediti stanje sistema.

Tabela 11. DOF – broj stepeni slobode iris koda

Broj stepeni slobode - DOF za iris kod									
	Ceo iris kod		DOF	Segment 1. iris kod		DOF	Segment 2. iris kod		DOF
λ_n	m	σ	bit	m	σ	bit	m	σ	bit
12	0.4760	0.0092	2946.8	0.4711	0.0106	2217.5	0.4623	0.0163	935.5
18	0.4718	0.0135	1367.3	0.4681	0.0136	1346.1	0.4548	0.0250	396.7
24	0.4665	0.0195	654.5	0.4646	0.0178	785.0	0.4458	0.0352	199.3

4.4. Izbor adekvatnog zaštitnog koda

Na početku izbora adekvatnog zaštitnog koda, potrebno je napraviti analizu grešaka u procesu projektovanja sistema koji je otporan na nastale greške. Kod ovakvih sistema greške mogu da budu uzrok otkaza.

Grešku (eng. *error*) možemo tretirati kao meru za odstupanje od tačnosti. Na primer, ukoliko je vrednost greške izuzetno visoka u sistemu za generisanje ključa na osnovu dva biometrijska podatka, onda to može da utiče na smanjenje količine uzajamne informacije između biometrijskih podataka što dovodi do otkaza sistema ili u drugom slučaju prouzrokuje pogrešan rezultat koji je u ovoj klasi sistema nepopustljiv. Dakle, vidimo da čak i ako postoji greška, ne mora da se ispoljava preko otkaza sistema, već može da daje loše rezultate.

Karakteristike grešaka možemo posmatrati sa sledeća tri aspekta:

- uzrok;
- vrednost;
- trajanje.

Uzrok greške može da nastane zbog problema u realizaciji na hardverskoj ili softverskoj platformi zbog lošeg izbora funkcionalnih komponenti.

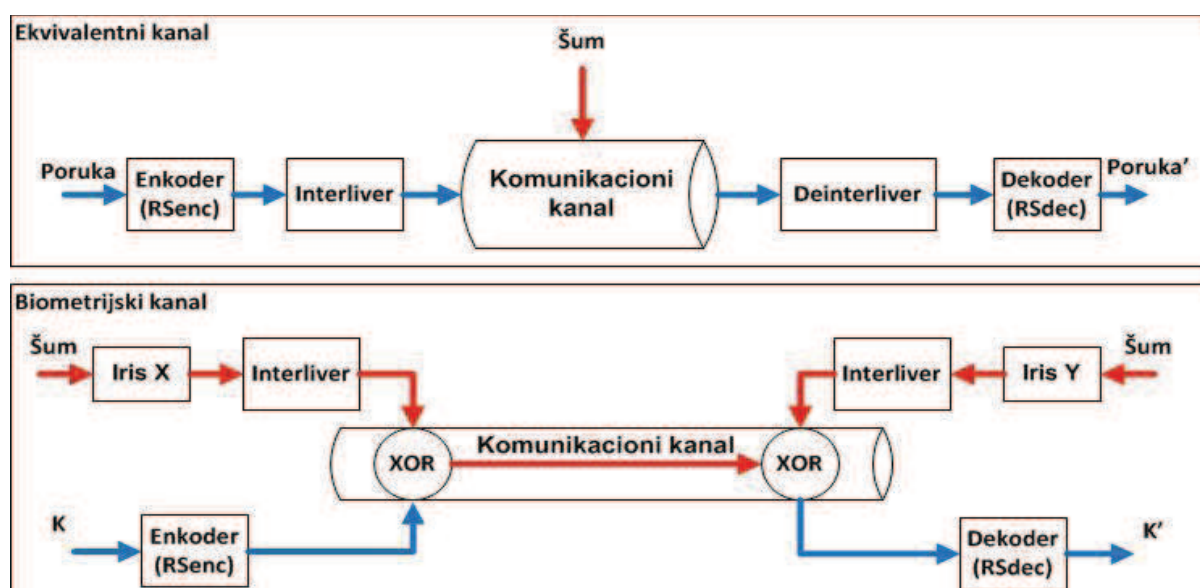
Vrednost grešaka generalno može da bude određena ili neodređena. Vrednost greške je određena ukoliko se ta vrednost ne menja bez spoljašnjeg uticaja. Greška koja ima neodređenu vrednost, njena vrednost se menja u različitim vremenskim trenucima.

Trajanje grešaka može da bude stalno, prolazno ili trenutno. U našem slučaju imamo stalne greške za koje smo merenjem odredili približne vrednosti koje su stalne i ispoljavaju se sve dok ne preduzmemo odgovarajuće zaštitne mere ili mere korekcije.

Pored ova tri navedena aspekta bitno je navesti i rasprostranjenost grešaka koja može biti okarakterisana kao lokana ili globalna. Rasprostranjenost grešaka u iris biometrijskom kodu može da bude lokalna i globalna.

4.4.1. Analiza ekvivalentnog kanala

Analiza ekvivalentnog kanala podrazumeva pronalaženje nekog klasičnog komunikacionog kanala iz teorije komunikacija koji utvrđuje uslove pod kojima je moguć pouzdan prenos informacija, a odgovara po svim svojim karakteristikama našem kanalu, ali sa različitim sadržajem. Potrebno je utvrditi maksimalnu brzinu pouzdanog prenosa po kanalu sa šumom uz upotrebu kodova za ispravljanje grešaka i maksimalnu količinu grešaka koja nastaje kada se informacija prenese većom brzinom od maksimalne.



Slika 44. Analiza ekvivalentnog kanala

Sada je jasno da ekvivalentni kanal predstavlja isti kanal koji je opisan na dva različita načina. **Slika 44.**, sadrži analizu dva ekvivalentna kanala. Gornji kanal na slici predstavlja ekvivalentni kanal koji se često može sresti u klasičnim komunikacionim kanalima. Šema kanala sadrži enkoder, dekoder, interliver, deinterliver i kanal sa šumom. Poruka koja se prenosi ovim kanalom, izložena je uticaju šuma koji se ispoljava u vidu paketskih grešaka u sadržaju poruke. Obzirom da se radi o većim paketskim greškama na predajnoj strani, primenjen je Rid-Solomonov kod i interliver čiji je zadatak da rasprši simbole kodovane poruke, a zatim se poruka šalje preko kanala. U toku prenosa poruka je izložena uticaju šuma koji utiskuje u sadržaj poruke paketske greške. Kada poruka stigne na prijemnu stranu, primenjuje se deinterliver koji simbole kodovane poruke vraća na iste indeksne pozicije koje su bile nakon kodovanja na predajnoj strani. Paketske greške koje su nastale sada su razbijene i ujednačeno raspoređene preko celog sadržaja poruke. U poslednjoj fazi, primenom Rid-Solomon dekodera koriguju se sve nastale greške.

Donji kanal na slici predstavlja biometrijski kanal sa prisustvom šuma. Šema kanala sadrži enkoder, dekoder, interliver i komunikacioni kanal. U ovom kanalu poruka je ključ K koji se prenosi preko šumnog kanal koji je indukovao sa dva iris. Šum koji je prisutan u kanalu, ispoljava se preko paketskih grešaka. U cilju postizanja visokih performansi na strani dekodera, primenjujemo interliver nad oba iris (jedan na predajnoj, a drugi na prijemnoj strani) koja zajedno indukuju komunikacioni kanal sa šumom. Zatim, ključ K kodujemo sa Rid-Solomon enkoderom i šaljemo kroz takav komunikacioni kanal nad kojim sada imamo kontrolu za određen opseg vrednosti paketskih grešaka. Poruka na prijemu ili ključ K se dekoduje Rid-Solomon dekodera i dobijena poruka predstavlja ključ K' koji je jednak ključu K na predajnoj strani.

Možemo sad da zaključimo da smo analizom ekvivalentnog kanala sami modelovali ovaj novi kanal, preko koga ostvarujemo prenos poruka sa maksimalnim kapacitetom. Ovaj kanal ima osobine binarnog simetričnog kanala sa šumom i na ovaj način smo potvrdili Šenonovu teoremu, koja je bila veliko iznenađenje u radu [33], u kome se tvrdi da ako izaberemo odgovarajući način kodovanja informacije za zadati kanal, može se postići po želji mala greška dekodovanja, odnosno prenosa informacija. U poglavlju 3. dato je više detalja vezanih za ovu teoremu.

4.4.2. Uloga zaštitnog koda i interlivera

Imajući u vidu da projektovani sistem pripada biometrijskoj kriptografiji, glavni izazov je integritet podataka. Greške koje nastaju nije moguće u potpunosti eliminisati, ali ih je moguće znatno smanjiti. U zavisnosti od stepena integracije, verovatnoća nastanka greške je veća ili manja. Sa aspekta projektovanja sistema, odnosno za postizanje visoke pouzdanosti u radu sistema koriste se tri osnovna pristupa:

- izbegavanje grešaka;
- ispravljanje grešaka;
- smanjenje grešaka;

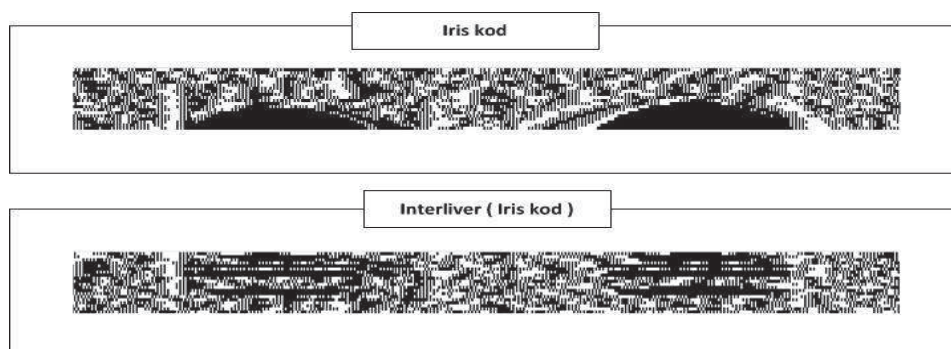
Izbegavanje grešaka podrazumeva sprečavanje nastanka grešaka nastalih na hardverskoj ili softverskoj platformi. Ovakve tehnike se odnose na izbor kvalitetnog hardvera (iris kamera) koji učestvuju u procesu akvizicije slike irisa, kao i adekvatan izbor algoritama koji učestvuju u obradi signala (kodovanje irisa). Izbegavanje grešaka podrazumeva i faze u kojima se vrše testiranja gde se preispituju pojedine komponente sistema.

Ispravljanje grešaka je tehnika koja ne sprečava nastanak greške. Ova tehnika uspešno detektuje i ispravlja nastale greške. Možemo da kažemo da ova tehnika ima mogućnost transformacione invertibilnosti nad podatkom sa greškom, u cilju dobijanja istog podatka bez greške. Jedna od važnijih komponenti ovog sistema su zaštitni kodovi koje primenjujemo nad podatkom (ključ K), koji je u toku prenosa izložen greškama. Na ovaj način je omogućen prenos u kritičnim situacijama za zadati stepen tolerancije.

U projektovanju sistema korišćeni je Rid-Solomon kod. Ovaj kod ispravlja greške na nivou simbola, a ne na nivou bitova, pogodni su za ispravljanje koncentrisanih paketskih (eng. *burst*) grešaka. Oni imaju široku primenu u sistemima digitalnih komunikacija i sistemima za čuvanje podataka. Rid-Solomon kodovi su podgrupa BCH kodova i klasa linearnih, nebinarnih cikličnih kodova. Rid-Solomon kod je blok kod koji se predstavlja kao $RS(n, k, t)$ preko polja Galoa $GF(2^n)$.

Rid-Solomon kod uvodi veliku redundantnu informaciju (preko polinomske strukture) i tako postiže visoke performanse za detekciju i ispravljanje grešaka. Da bi se ova cena platila neophodno je koristiti samo konzistentnu informaciju (jedinstvene karakteristike). Iz ovog razloga su urađene rigorozne informacione analize nad teksturom irisa, sa ciljem izbora upravo takvih segmenata koji će da sadrže većinom konzistentnu informaciju.

Smanjenje grešaka je uglavnom jedna od tehnika koja utiče na smanjenje nastalih grešaka koje nije moguće ispraviti. U našem slučaju izabrali smo za ovu vrstu tehnike interliver. Interliveri koriste metodu preplitanja vrednosti u iris kodu. Metod za preplitanje može da bude raznovrsan. Mi koristimo slučaju permutaciju koja je zadata sa nekom inicijalnom vrednošću koja predstavlja početno stanje pesudogeneratora. Ovakvi interliveri uvode element slučajnosti u neki dizajn koda bez povećanja kompleksnosti, pretpostavljajući korišćenje iterativnog dekodovanja dva interlivovana koda.



Slika 45. Binarna struktura iris koda

Slika 45., prikazana je binarna struktura iris koda, na osnovu koje je moguće vizuelno prepoznati efekat koji je dobijen primenom slučajnog interlivera.

Uloga slučajnog interlivera u našem dizajnu sistema je dvostruka. Njegova osnovna uloga je u klasičnim šumnim komunikacionim kanalima za raspršivanje paketskih grešaka, u našem sistemu slučajni interliver koristimo za razbijanje paketskih grešaka i za smanjenje uzajamne informacije između dva iris koda, koja nisu generisana preko istog informacionog izvora. Bitno je istaći da slučajni interliver ne utiče na entropiju sistema, ostavljajući je na istom nivou.

Tabela 12., sadrži uzajamne informacije iris kodova sa interliverom i bez interlivera i njihove entropije. Data analiza se odnosi samo na prve segmente iris koda.

Tabela 12. Uticaj slučajnog interlivera na uzajamne informacije i entropiju

Uticaj interlivera na entropiju i uzajamnu informaciju		
Merene veličine (uzajamna inf. i entropija)	bit/bit	Ukupno (bit/segment 1.)
$I(X; Y)$	0.1016	487.6
$I(X; \text{interliver}(Y))$	0.0670	321.6
$I(\text{interliver}(X); Y)$	0.0715	343.2
$I(\text{interliver}(X); \text{interliver}(Y))$	0.0683	327.8
$I(X; X')$	0.2078	997.4
$I(\text{interliver}(X); \text{interliver}(X'))$	$I(X; X')$	997.4
$H(X)$	0.8021	3850
$H(Y)$	0.8112	3893.7
$H(\text{interliver}(X))$	$H(X)$	3850
$H(\text{interliver}(Y))$	$H(Y)$	3893.7

Razmotrićemo prvo rezultate vezane za uzajamnu informaciju. X i Y simbolizuju irise od različitih osoba, X i X' irise od iste osobe, a funkcija $in()$ interliver. U prvom merenju određena je uzajamna informacija $I(X; Y)$, koja ima referentnu ulogu. U drugom merenju računali smo $I(X; in(Y))$. Rezultat koji je dobijen ukazuje na smanjenje $I(X; Y)$. U trećem merenju primenjen je interliver na obe promenljive i dobijena je slična vrednost koja odgovara proseku $I(in(X); Y)$ i $I(X; in(Y))$. U četvrtom merenju izračunali smo informaciju $I(X; X')$, zatim je primenjen interliver na obe promenljive i ponovo računali $I(in(X); in(Y))$. Dobijena je slična identična vrednost koja odgovara $I(X; X')$. Zaključujemo da primena interlivera na jednu ili na obe promenljive X i Y , jednako utiče na smanjenje uzajamne informacije koje iznosi približno 40%.

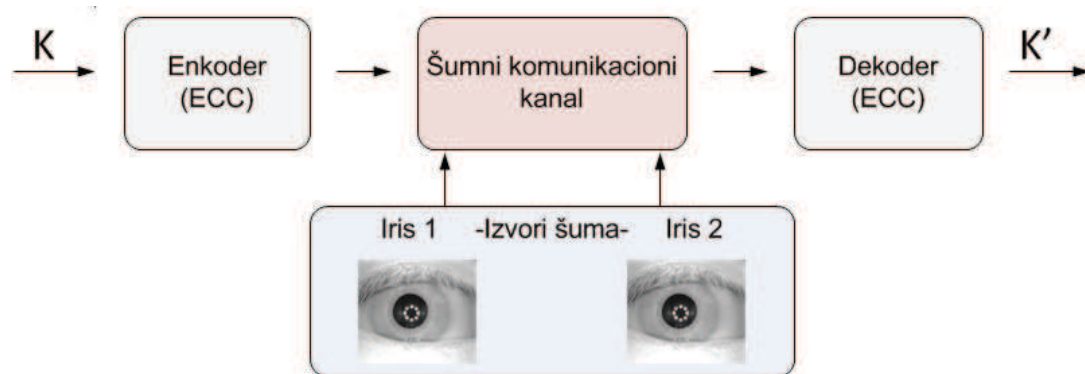
Zatim sledi analiza dobijenih rezultata koja se odnosi na entropiju. U prvom merenju izračunata je entropija $H(X,Y)$, koja ima referentnu ulogu. U drugom merenju imamo $H(\ln(X))$ i $H(\ln(Y))$. Dobijeni rezultati ukazuju na nepromenjeno stanje entropije $H(X)$ i $H(Y)$. Zaključujemo da primena slučajnog interlivera na X ili Y ne utiče na promenu entropije.

4.4.3. Motiv za uvođenje zaštitnog koda i interlivera

Primena zaštitnih kodova i interlivera nad irisom je stavila biometriju u paralelu sa klasičnim komunikacionim kanalima. Jedan od naših ciljeva je bio da pronađemo ekvivalentni kanal koji ima zajedničke osobine sa biometrijskim komunikacionim kanalom. Pronalazak ekvivalentnog kanala, davao bi garantovano dobre performanse celokupnog sistema koji je krajnji cilj.

Slika 46., predstavlja deo sistema u kome je biometrija poistovećena sa klasičnim komunikacionim kanalom sa prisustvom šuma. U ovom primeru dva irisa indukuju biometrijski komunikacioni kanal preko koga imamo cilj da prenesemo ključ K . Vrednosti greške u ovom kanalu variraju u zavisnosti od toga da li dva irisa pripadaju jednoj osobi ili su od različitih osoba. U ovom apstraktnom modelu biometrijskog kanala, naš zadatak je da ključ K prenesemo preko kanala bez grešaka. Tačnije, cilj je projektovati komunikacioni kanal čiji je maksimalni kapacitet definisan izrazom:

$$C = \max I(K; K') = 1 \quad (100)$$



Slika 46. Biometrijski komunikacioni kanal sa šumom

Razlozi za nastanak ovakvog motiva obrazložen je kroz sledeće činjenice:

- 1) Osnovni problem kod biometrijskih podataka je varijabilnost koja je prisutna svaki put u podacima nakon akvizicije biometrijskih podataka. Drugim rečima, svaki novi biometrijski podatak generisan na istom biometrijskom izvoru se uvek razlikuje 10% - 20% od svih prethodnih, dok se biometrijski podaci generisani na različitim biometrijskim izvorima razlikuju od 40% do 60%.
- 2) Kao odgovor na pomenuti problem varijabilnosti u biometriji irisa, nastala je ideja o ekvivalenciji sa drugim komunikacionim kanalima. Ova ideja je uvela primenu

zaštitnih kodova za ispravljanje grešaka (eng. *error correcting code*), a tehnika je preuzeta iz oblasti komunikacija (Slika 46).

- 3) Kapacitet koda za ispravljanje grešaka u ovom slučaju performanse dekodera, moraju biti takve da odvajaju originalnog korisnika od lažnog. Na primer, biometrijski kanal indukovao sa dva iris koda sa istog biometrijskog izvora, mogu biti u stanju da prenesu poruku bez grešaka u našem slučaju ključ K , dok kanal indukovao sa dva iris kod sa različitih biometrijskih izvora, ne sme da prenese ključ K , čak šta više poželjno je da ključ K bude uništen, tako da je uzajamna informacija $I(K;K')=0$.
- 4) Postizanjem pomenutih performansi, uspeh bi predstavili sa rezultatom $K=K'$, dok se K' koristi kao kriptološki ključ.
- 5) Pravljenjem kompromisa između ukupne količine informacije H iris koda i njihove uzajamne informacije I između istih i različitih kodova irisa, dobijamo okvire za maksimalnu veličinu ključa K .
- 6) Interliver je tehnika koja se koristi za raspršivanje paketskih grešaka. Analizom je utvrđeno da iris kod sadrži paketske greške. Ova tehnika se kod klasičnih komunikacionih kanala koristi za podizanje performansi zaštitnih kodova ili kodova za ispravljanje grešaka.

Sada možemo da zaključimo da sve ove činjenice nezaustavno utiču na stvaranje jedne nove oblasti, koja opravdano nosi ime biometrijska kriptografija.

4.5. Evaluacija sistema

U ovom delu rada ćemo prikazati evaluaciju naše implementacije sistema za generisanje kriptoloških ključeva na osnovu biometrije irisa. Evaluacija performansi rešenja bazirana je na rezultatima dobijenih preko eksperimenata čiji je model i struktura takođe objašnjen u ovom poglavlju. U nastavku ćemo diskutovati o rezultatima i napraviti poređenje sa drugim rezultatima. Na kraju ćemo navesti oblasti primene predloženog rešenja.

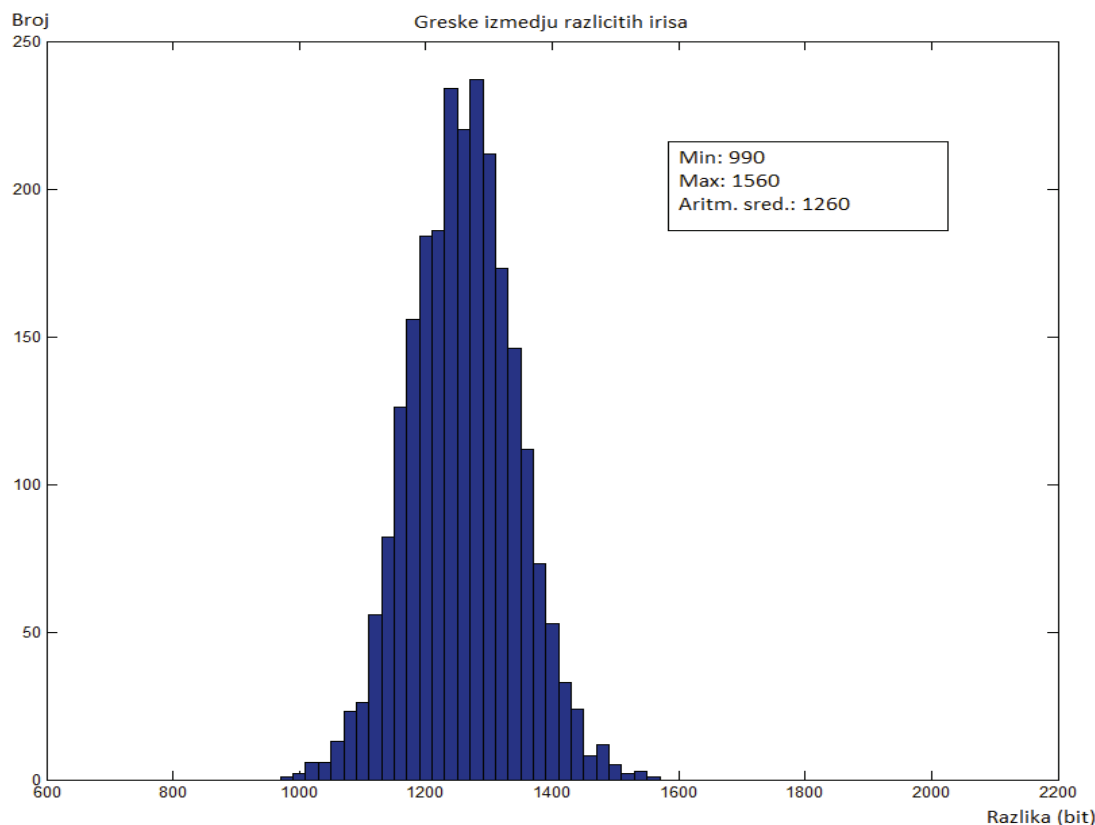
4.5.1. Prikaz eksperimentalnih rezultata

Iris baza podataka koju smo mi koristili, sastoji se od 700 uzoraka irisa od 70 različitih očiju iz „CASIA“ baze podataka, sa po 10 uzoraka od svakog oka. Veličina iris koda koji dobijamo nakon faze kodovanja je 9600 bitova.

Nakon sprovedene rigorozne informacione analize, odlučili smo da koristimo samo segmente sa visokom entropijom. Entropijske vrednosti segmenata date su u podpoglavlju 4.3.2. Sa druge strane, zbog nesavršenosti metoda za automatsku segmentaciju irisa, odlučili smo da koristimo segmente koji nisu podložni greškama. Greške koje su česte, nastaju zbog kapaka i trepavica usled faze za automatsku segmentaciju. Ovakve greške mogu da izazovu velike razlike između dva ista irisa i da tako prevazilaze određeni kapacitet Rid-Solomon koda.

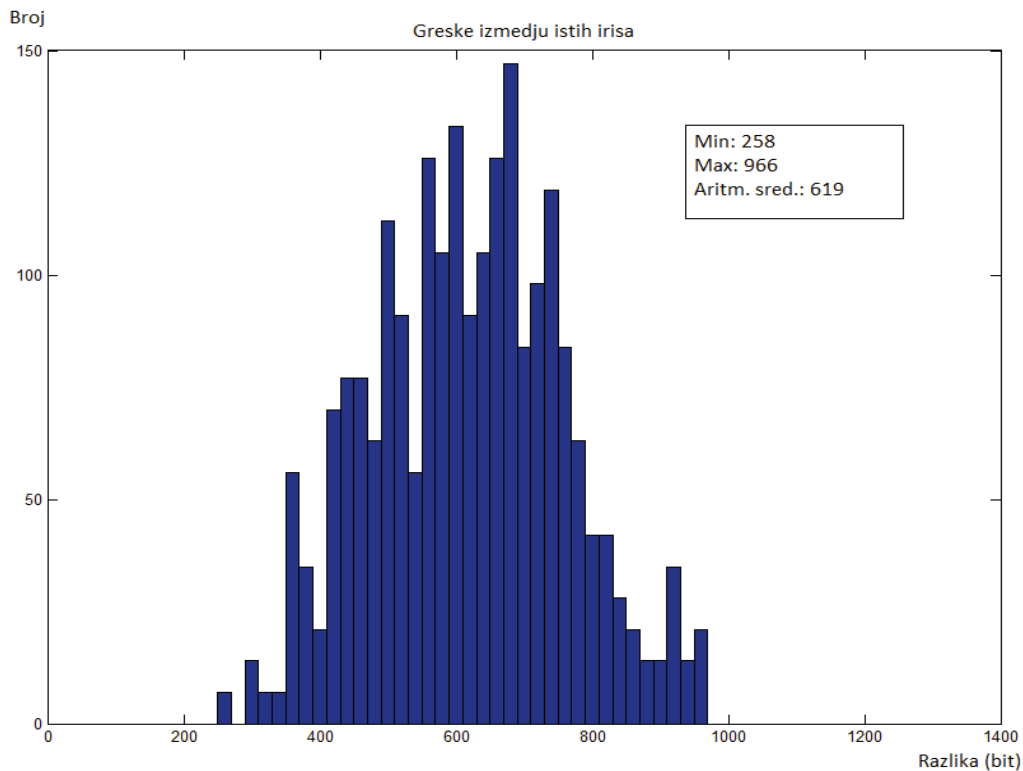
Performanse našeg sistema analizirane su na ukupno 2600 bitova iris koda. Ovi bitovi predstavljaju informaciono najsnažnije bitove, koje smo prepoznali nakon informacione analize.

U prvoj fazi, potrebno je bilo utvrditi maksimalni dozvoljeni kapacitet RS koda za ispravljanje grešaka. U cilju rešavanja ovog problema, na sledećem grafiku su prikazane razlike između dva različita iris koda sa segmentima po 2600 bitova. Maksimalna razlika je 1560 bitova (60 %), minimalna razlika 990 bitova (38%), a aritmetička sredina je 1260 bitova (48.4%) (Slika 47).



Slika 47. Broj grešaka između različitih irisa

U drugom koraku smo utvrdili kolike su razlike između istih uzoraka irisa. Maksimalna razlika je 966 bitova (37%), minimalna razlika je 258 (10%), a aritmetička sredina je 619 bitova (23%).



Slika 48. Broj grešaka između istih irisa

Sada na osnovu ovih rezultata možemo da projektujemo kapacitet RS koda. Kapacitet RS koda ne sme da bude veći od minimalne razlike između različitih irisa (**Slika 47**). Maksimalni kapacitet koda je ograničen na 37%. Kapacitet RS koda, definisan je izrazom:

$$t = \frac{n - k}{2} \quad (101)$$

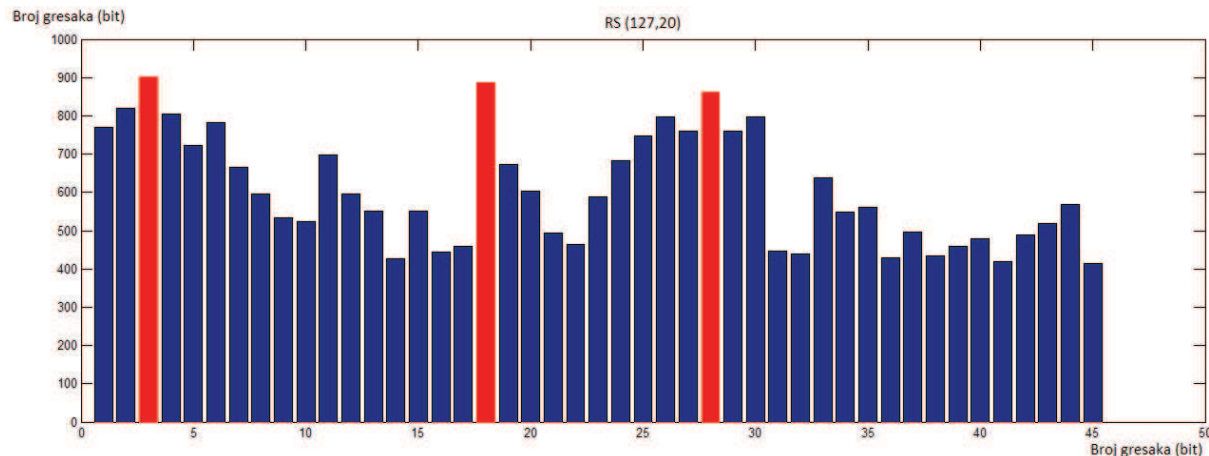
Da bismo postigli traženi kapacitet koda, u startu smo odabrali vrednost parametra $n = 2^7$, a vrednosti parametara $k = \{20, 15, 10\}$. U sledećoj tabeli dati su teorijski kapaciteti kodova za izabrane parametre.

Tabela 13. Teorijski kapacitet RS koda

n	k	t
127	20	53.5
127	15	56
127	10	58.5

Teorijske kapacitete koji su dati u tabeli, nije moguće postići zbog neuniformne distribucije grešaka u iris kodu. Da bi ovaj kapacitet bio postignut, očekivan je jednak broj grešaka u svakom bloku n , što je kod irisa nemoguće.

U nastavku pokušaćemo da na osnovu velikog broja uzoraka i analizom greške po blokovima, utvrdimo stvarni kapacitet RS kodera koji važi za iris kod. Sa druge strane, želimo da potvrdimo stacionarnost greške u iris kodu.



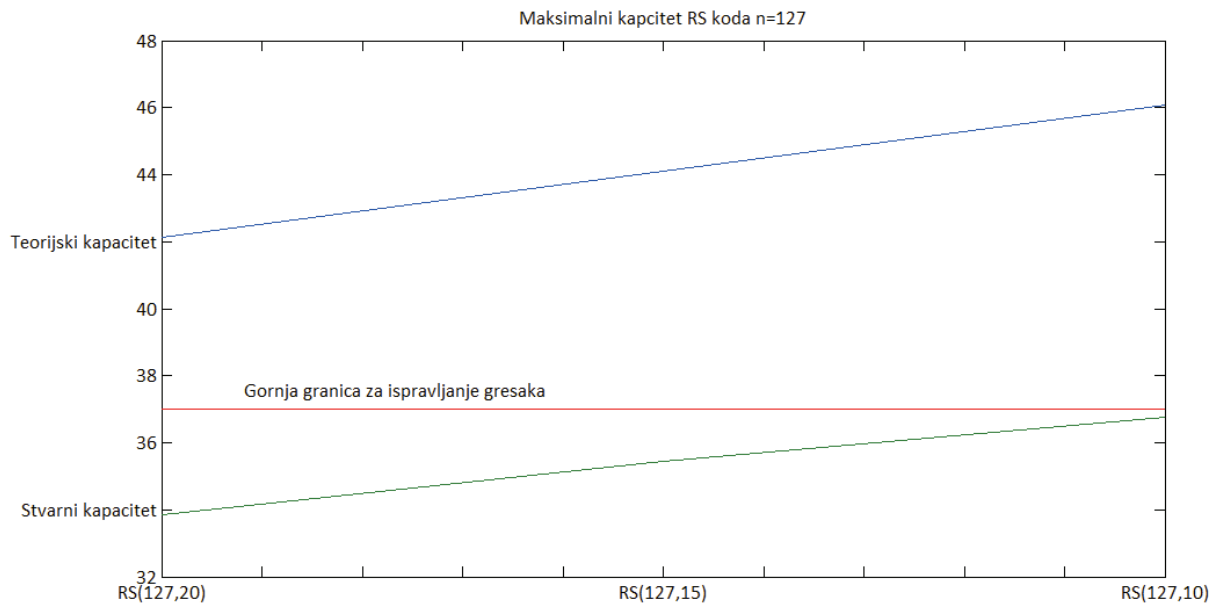
Slika 49. Kapacitet greške za RS(127,20)

Slika 49., predstavlja komparativnu analizu između 10 uzoraka irisa od istog oka u pokušaju da se regeneriše ključ K . Za svaki uzorak irisa, generisan je pseudokod koji u sebi sadrži ključ koji je isti za sve uzorke. Utvrđeno je da koder RS(127,20) nije mogao da regeneriše ključ na ovom testnom uzorku, samo u 3 slučaja kada je greška između irisa bila veća od 860 bitova. Očekivani kapacitet navedenog koda za ispravljanje grešaka po jednom bloku je 53.5, obzirom da ukupno ima 20 blokova, stvarni kapacitet koda je 43 greške, a to znači da je u proseku po jednom bloku mogao da ispravi do 43 greške.

Teorijski kapacitet koda RS(127,20) smanjen je za 20%, što predstavlja stvarni kapacitet. Ovaj kod između bilo koja dva irisa, može maksimalno da ispravi 33% nastalih grešaka. Na osnovu eksperimentalnih rezultata, u tabeli ispod smo prikazali stvarni kapacitet RS kodova za izabrane parametre.

Tabela 14. Stvarni kapacitet RS koda

n	k	t
127	20	43
127	15	45
127	10	47



Slika 50. Maksimalni kapacitet RS koda

Slika 50., prikazuje stvarni i teorijski kapacitet RS koda u procentima. Ako bolje pogledamo, videćemo da teorijski kapacitet prelazi dozvoljenu granicu za ispravljanje grešaka (37%). Međutim, nakon sprovedene analize u cilju karakterizacije grešaka koje su prisutne u iris kodu, utvrđen je „novi“ ili stvarni kapacitet RS koda koji se nalazi ispod dozvoljene granice za ispravljanje grešaka.

Maksimalni kapacitet koda ili gornja granica za ispravljanje grešaka između dva iris koda je 37%. Kapacitet RS koda koji se odnosi na iris podatak, sada možemo da definišemo izrazom:

$$t = \frac{n - k}{2} - 20\% \quad (102)$$

Nakon utvrđenog kapaciteta kodova za ispravljanje grešaka, sproveli smo evaluaciju nad sistemom za generisanje kriptoloških ključeva.

Tabela 15. Performanse sistema kada je $n = 127$

RS(n,k) kod	Veličina iris koda (bit)	Veličina ključa K (bit)	Broj blokova k	Lažno odbacivanje (FRR) %	Lažno prihvatanje (FAR) %
RS(127,20)	2540	400	20	4,75	0
RS(127,15)	2540	300	20	2,21	0
RS(127,10)	2540	200	20	1,26	0

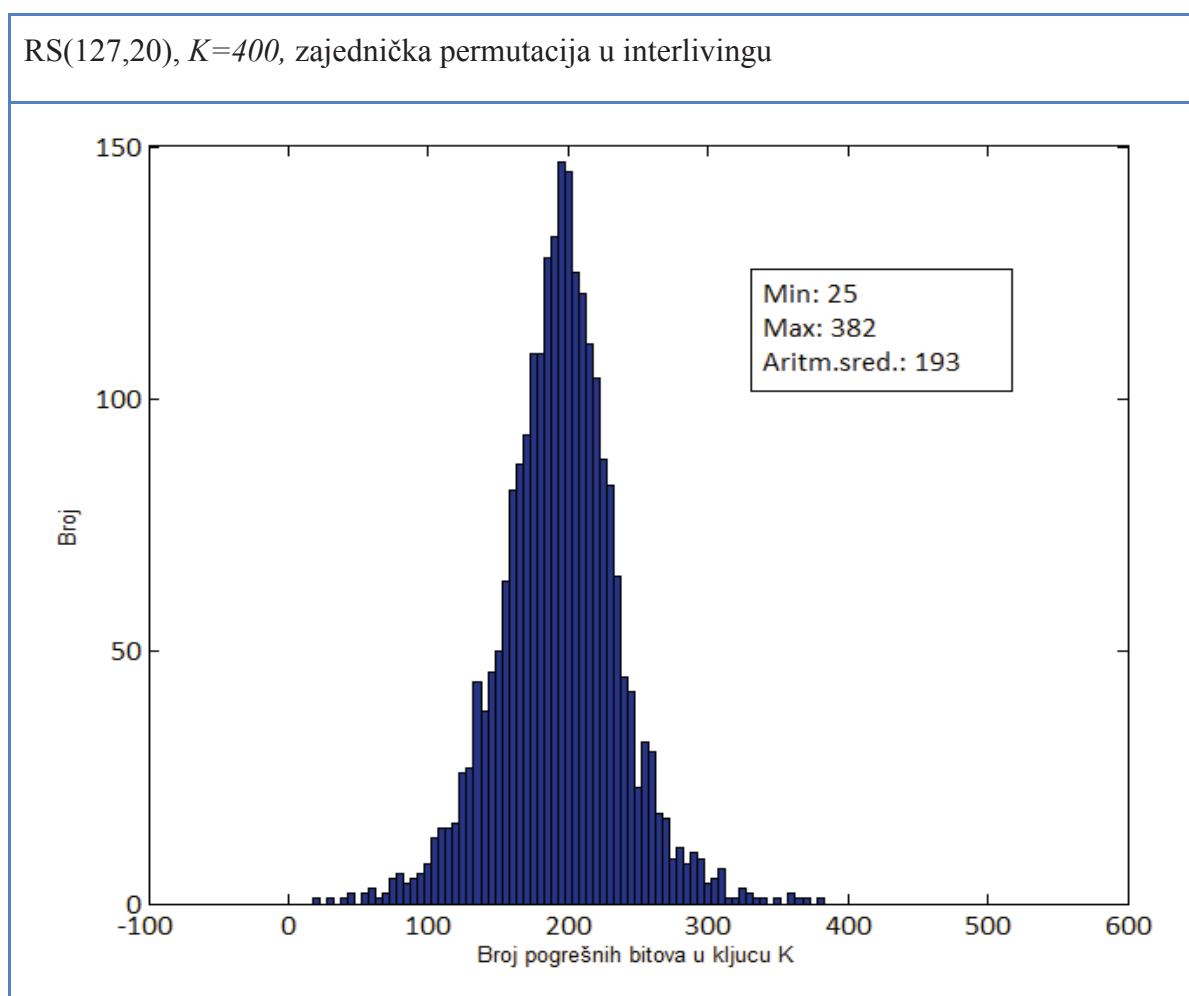
Tabela 15., sadrži rezultate koji predstavljaju odnos između izabranog koda, veličine iris koda i veličine ključa.

Kao što je prikazano, dizajn koda RS(127,20) predstavlja najpogodniji kod. Preko ovog koda moguće je regenerisati ključ od 400 bitova sa stopom lažnog odbacivanja do 4.75%. Za sve kodove stopa lažnog prihvatanja je 0%. Naš mehanizam kodovanja je efikasan za sve iris kodove kod kojih je stopa greške ispod 37%.

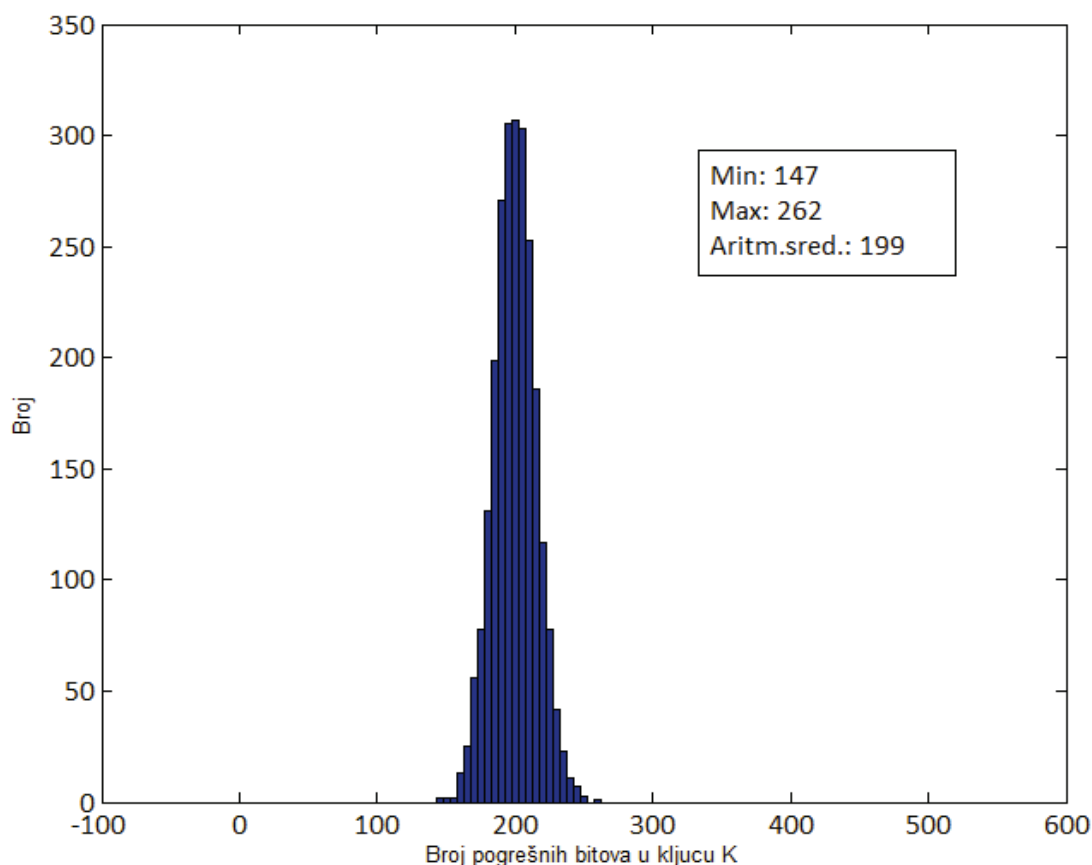
Dobre performanse sa RS kodovima postignute su primenom slučajnog intrlivinga. Pored zasluge za dobre performanse, interliveri igraju vrlo važnu ulogu u smanjenju uzajamne informacije između različitih irisa. Na ovaj način se povećava udaljenost ili smanjuje stopa poklapanja između dva različita irisa. Ovaj rezultat se pozitivno odražava na stopu lažnog prihvatanja (FAR) čija je vrednost 0%.

U nastavku, prikazaćemo komparativnu analizu za 70 različitih irisa (2415 poređenja) od različitih očiju sa zajedničkom permutacijom u intrlivingu, kao i sa zasebnom permutacijom u interlivingu za sve tri konstrukcije RS kodova.

Tabela 16. Broj pogrešnih bitova u ključu u slučaju iste i različite permutacije



RS(127,20), $K=400$, zasebna permutacija u interlivingu



U **Tabela 16.** jasno se vidi uloga interlivinga. Na prvom histogramu, prikazan je minimalni i maksimalni broj grešaka u ključu u međusobnoj komparaciji 70 različitih irisa. Minimalna greška je 25 bitova, maksimalna 382 bita i aritmetička sredina je 193 bita. Na drugom histogramu, imali smo zasebne permutacije. Minimalna greška je bila 147 bitova, maksimalna greška 262 bita i aritmetička sredina 199 bitova.

4.5.2. Diskusija, kritike i poređenje sa postojećim rezultatima

Tabela 17., sadrži zbir rezultata, kao i poređenja, u oblasti za generisanje ključeva na osnovu biometrijskih podataka. Autori [1], postigli su dužinu ključa od 140 bitova sa stopom lažnog odbacivanja FRR od 0,47% i stopom lažnog prihvatanja FAR od 0%, [26] su postigli dužinu ključa od 186 bitova sa stopom lažnog odbacivanja FRR od 0,76% i stopom lažnog prihvatanja FAR 0,096%. Dobijeni rezultati u ova dva rada su mnogo bolji od rezultata u drugim radovima, kod kojih je stopa lažnog prihvatanja i odbacivanja u fazi verifikacije mnogo veća, a generisani biometrijski ključevi značajno manjih dužina, ne mogu da odgovore na zahteve savremenih kriptografskih mehanizama.

Tabela 17. Poređenje sa postojećim rezultatima u ovoj oblasti

Biometrija	Autor	Dužina ključa (bit)	Stopa lažnog odbacivanja (FRR)	Stopa lažnog prihvatanja (FAR)
Glas	[15]	46	20%	-
Svojeručni potpis	[16]	40	28%	1,2%
Otisak prsta	[19]	-	-	-
	[18]	69	30%	-
	[21]	49 i 85	5,4% - 9,9%	3,2% - 2,5%
Lice	[17]	-	0,93%	
	[2]	58	3,5%	0%
Iris	[1]	140	0,47%	0%
	[26]	186	0,76%	0,096%
	- 2013	400	4,75%	0%
EKG i Glas	[22]	240	1.27%	10.62%

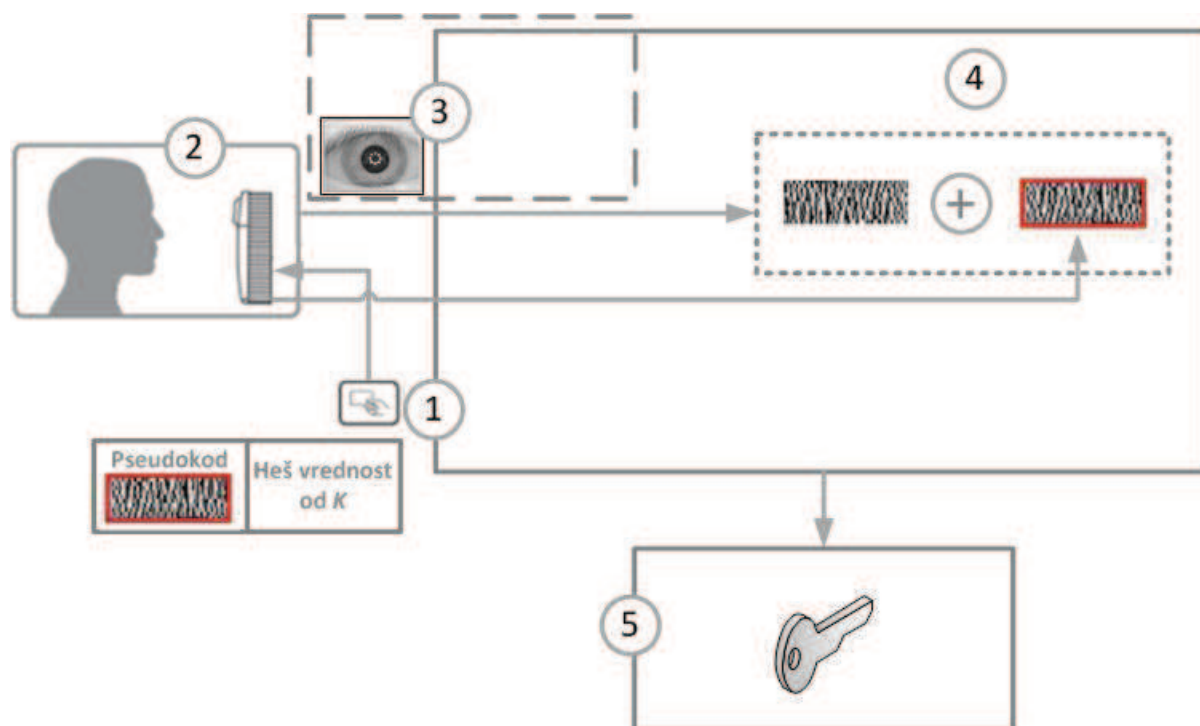
Pregledom do sada postignutih rezultata, možemo da kažemo da smo uspeli da projektujemo sistem koji je u stanju da generiše do sada najveći ključ veličine do 400 bitova, sa stopom lažnog odbacivanja FRR od 4,7% i stopom lažnog prihvatanja FAR od 0%.

Smatramo da je ostalo još dovoljno prostora koji može da se iskoristi za poboljšanje performansi našeg sistema, sa aspekta smanjenja stope lažnog odbacivanja i mogućeg povećanja veličine regenerisanog ključa.

Daugman je u svom radu postigao najbolju vrednost parametra FRR od 0,47%, što predstavlja rezultat na zavidnom nivou, ali sa tri puta manjim ključem. Od velikog značaja za ovakvu vrednost parametra FRR, imala je funkcija pomeraja za 8 bitova ulevo i 8 bitova udesno. Zapravo, od pomeraja do pomeraja ključ je od početka regenerisan, zatim računata je heš vrednost i upoređena sa heš vrednošću koja se nalazi na hardverskom *tokenu* kao što je *smart* kartica. U našem sistemu smo izbegli funkciju pomeraja, zbog želje da sačuvamo vremenske performanse sistema na prihvatljivom nivou, a sa druge strane, odabirom korisnih segmenata, u potpunosti smo izbegli greške koje nastaju usled automatske segmentacije u fazi kodovanja biometrijskog izvora irisa, što je doprinelo sigurno boljim performansama kompletnog sistema.

4.5.3. Oblast primene predloženog rešenja

Osnovna primena predloženog i razvijenog rešenja je u domenu kriptografskih potreba. Dobijeni ključevi mogu biti korišćeni u kriptografske svrhe, kao i kod sistema za sigurnu autentifikaciju korisnika na bazi heš vrednosti ključa K . Arhitektura predloženog sistema zasniva se na tri bezbednosna faktora, a to su, korisnička lozinka (PIN), biometrijski uzorak i *token* (*smart* kartica, *USB*,...). Arhitektura predloženog biometrijskog sistema omogućava proveru identiteta i generisanje ključeva u „*off-line*“ režimu rada. Na ovaj način se izbegava centralizovana baza podataka, što doprinosi stvaranju pozitivnih utisaka za socijalno prihvatanje sistema. Na sledećoj slici (**Slika 51**), prikazana je jedna univerzalna šema za implementaciju predloženog rešenja u „*off-line*“ režimu rada.



Slika 51. Jedna šema primene sistema

Na slici (**Slika 51**) su označene faze u sistemu sa brojevima od 1 do 5. U prvoj (1) fazi korisnik sistema prilaže smart karticu koja na sebi sadrži pseudokod i heš vrednost od zaključanog ključa u pseudokodu. Nakon unosa ispravne korisničke lozinke, podaci sa smart kartice se prosleđuju sistemu za regenerisanje ključa (4). U drugoj (2) fazi, očitava se i prosleđuje korisnikova originalna biometrija irisa preko biometrijskih skenera. U trećoj (3) fazi primenjuje se sistem za generisanje iris koda (biometrijskog obrasca). Zatim, generisani iris kod se prosleđuje sistemu za regenerisanje ključa. U četvrtoj (4) fazi, prelazi se u fazu regenerisanja ključa sa našim novopredloženim rešenjem u disertaciji. Uspešnost ove faze potvrđuje se sa identičnom heš vrednosti koja je dobijena sa smart kartice, a rezultat ove faze je biometrijski ključ koji se može koristiti u razne kriptografske svrhe.

5. Zaključak

U ovom poglavlju sumiraćemo naš rad na razvoju jedne klase sistema za generisanje kriptoloških ključeva na osnovu biometrijskih podataka. Takođe, navešćemo polja u kojim se daljim istraživanjima i razvojem mogu ostvariti dodatni rezultati i doprinosi. Inicijalno postavljeni ciljevi biće rezimirani kroz uporedno navođenje ostvarenih rezultata.

5.1. Sumarni ciljevi istraživanja

Na početku istraživanja, u ovom radu, imali smo utisak da su dostignuti maksimumi kod postojećih sistema za generisanje ključeva na osnovu biometrije, kao i da su dostignuti određeni nivoi stabilnosti i bezbednosti, koja je neophodna za njihovo korišćenje sa aspekta društvene prihvatljivosti. Međutim, nakon određenih testiranja aktuelnih rešenja, javila se sumnja da je teorijski maksimum ipak iznad postignutih praktičnih rezultata.

Daljim radom, potvrđene su sumnje i postavljeni su čvrsti ciljevi istraživanja da se podigne granica efikasnosti blizu teorijske granice. Nakon uvida u rezultate teorijsko-informacione analize, postavljeni su novi teorijski okviri sistema za sintezu kriptoloških ključeva, na osnovu biometrijskih podataka.

Osnovna motivacija ovog istraživanja, bila je potreba da se omogući razvoj profesionalne klase sistema za generisanje ključeva velikih dužina, tako da zadovolji potrebe savremenih kripto sistema, korišćenjem postojećih komponenta za kodovanje biometrijskih izvora koji obuhvataju kompletan proces, od izbora biometrije, preko slike do biometrijskih obrazaca.

U skladu sa tim, kao osnovni cilj postavljen je razvoj sopstvene klase sistema za generisanje kriptoloških ključeva na osnovu biometrije irisa. Za uspešno ostvarivanje ovog cilja bilo je neophodno sprovesti razne kontrole, kao i uraditi rigorozne informacione analize nad svim komponentama sistema u skladu sa kriptografskim pravilima i principima koje je bilo bitno ispoštovati.

Kompleksnost savremenih kripto sistema, kao i potreba za kriptološkim ključevima većih dužina, učinile su rad na ovoj tezi izuzetno složenim i zahtevnim.

5.2. Ostvareni rezultati i doprinosi

U ovom radu je predstavljeno praktično realizovano sopstveno rešenje za generisanje kriptoloških ključeva, na osnovu biometrijskih podataka sa sopstvenim algoritmima. Osnovni doprinos ovog rada na polju kriptografske bezbednosti čini sledeće:

1. izvršena je rigorozna informaciona analiza biometrije irisa u cilju sinteze optimizovanog sistema za ekstrakciju kriptoloških ključeva visokog kvaliteta;

2. izvršena je optimizacija parametara odgovarajuće transformacione funkcije, u cilju dobijanja što veće entropije i što niže uzajamne informacije u transformacionom domenu;
3. estimacijom lokalne entropije i uzajamne informacije, identifikovani su segmenti irisa koji su najpogodniji za ove potrebe;
4. postavljeni su okviri za sintezu sistema za ekstrakciju čisto slučajnih nizova iz iris biometrije, bez narušavanja njenih osvedočenih autentifikacionih svojstava;
5. dizajnirana je nova generička šema sistema za generisanje kriptoloških ključeva. Naša šema kombinuje kod za ispravljanje grešaka i tehniku interlivinga;
6. postignuta je veća dužina ključa od 400 bitova (što je znatno iznad aktuelnih rezultata u ovoj oblasti), sa stopom lažnog prihvatanja od 0% i stopom lažnog odbacivanja od 4,75%. Generička šema se čvrsto oslanja na dobijene okvire, nakon detaljno sprovedene teorijsko-informacione analize. Raniji pokušaji sinteze ovakvih sistema u ovoj oblasti su imali uglavnom neprihvatljive rezultate za sisteme ovakve namene ili su generisali relativno kratke ključeve koji su bili neupotrebljivi za kriptografske algoritme;
7. smatramo da je ovaj rad jedinstven zbog svoje specifičnosti koja se ogleda u gore već navedenoj teorijsko-informacionoj analizi, preko koje smo usvojili zajedničku jedinicu mere (Šenenov bit) koja je osnova za sintezu svakog kripto sistema, što nije bio slučaj do sada;
8. na osnovu mesta problema (između biometrije i kriptografije), definisali smo novu oblast pod nazivom „biometrijska kriptografija“ koja se bavi velikim problemom, kao što je uspostavljanje veze između biometrije i kriptografije na takav način, da je moguće zadovoljiti tačnost koja je zahtevana od strane kriptografije i pored varijabilnosti koja je prisutna kod biometrijskih podataka.

5.3. Predlog daljeg rada

Rešenje predstavljeno u ovom radu u potpunosti je odgovorilo na primarno postavljenu zadatak. Zadatak je bio razviti novu klasu sistema za generisanje kriptoloških ključeva na osnovu biometrije irisa. Rešenje je zasnovano na upotrebi sopstvenog sistema za generisanje kriptoloških ključeva sa ciljem potpune kontrole nad kvalitetom generisanih ključeva i procesima za generisanje i regenerisanje.

U toku rada su uočene sledeće mogućnosti i pravci daljeg razvoja:

1. prototip za predloženo rešenje;
2. bimodalni biometrijski sistema;
3. biometrijske PKI infrastrukture.

Primarna ideja u daljem radu je fokus na razvoj prototipa koji će omogućiti potpunu implementaciju našeg predloženog rešenja. Za potrebe rada na ovoj tezi i projekta kod Ministarstva za nauku i tehnološki razvoj Republike Srbije, šifra: TR-32054, nabavljena je profesionalna oprema za skeniranje oba irisa. Nabavljen je model kamere LG iris 4000 (**Slika**

52.), treća generacija kamere i broj jedan kod biometrijskih sistema za prepoznavanje na osnovu irisa. Nudi potpunu integraciju sa svim aplikacijama i fleksibilnost. Intuitivni vizuelni korisnički interfejs omogućava korisnicima da se brzo pozicioniraju za upis ili prepoznavanje na osnovu slike oba oka istovremeno.

Kada je zahtevano više faktora u autentifikaciji. Iris LG 4000 je izuzetno fleksibilan sa ugrađenim čitačem smart kartica vodećih svetskih proizvođača. Višefaktorska autentifikacija je podržana preko numeričke tastature za unos lozinke ili PIN koda u zavisnosti od modela uređaja.



Slika 52. LG Iris 4000 skener irisa

Druga ideja za dalji rad je razvoj bimodalnih biometrijskih sistema. Bimodalni sistemi kombinuju dva biometrijska signala. Ovo podrazumeva kombinovanje različitih biometrija. Na primer, kombinovanje EKG signala i govora, irisa i govora ili EKG signala i irisa.



Slika 53. Bimodalni biometrijski sistem

Takođe, i u ovom slučaju se koriste transformacione funkcije za ekstrakciju čiste informacije iz navedenih biometrija. Rešavanjem apstraktnih pojmova u domenu biometrijskih podataka, kao što je uzajamna informacija iz oblasti teorijsko-informacione analize, bićemo u

moгуćnosti da razvijemo algoritme za estimaciju uzajamnih informacija između dve ili više biometrija. Ekstraktovana informacija koristiće se kao osnova za sintezu sistema za generisanje još kvalitetnijih kriptoloških ključeva većih dužina u odnosu na teorijski postavljene okvire.

Pod razvojem biometrijske PKI infrastrukture, podrazumevamo razvoj novog sistema sa javnim ključevima koji će biti ekvivalentan postojećoj PKI infrastrukturi. Ovaj sistem je direktno zasnovan na biometrijskim podacima. Zbog zahtevanih većih dužina ključeva, razvoj ovog rešenja zavisiće od uspeha razvoja bimodalnih sistema. Ova ideja trenutno postoji samo na apstraktnom nivou. Prvi zadatak je pronalazak ekvivalentne šeme sa svim funkcionalnim komponentama koje postoje u tradicionalnoj šemi PKI infrastrukture.

Ideja o biometrijskoj PKI infrastrukturi ima veliki potencijal u današnjem informacionom društvu. Postoji više problema sa tradicionalnom PKI infrastrukturom u elektronskim servisima koji obezbeđuju tajnost, integritet i neporecivost. Glavni nedostatak je zahtev u poverenje u treću stranu od poverenja, koja je zadužena za izdavanje i potpisivanje privatnih ključeva. Postavlja se onda pitanje koliko možemo da verujemo jednom ovakvom sistemu, ukoliko se radi o primeni u institucijama od velikog značaja (vojska, diplomatija, policija)? Drugi problem je robusna infrastruktura koja onemogućava uzajamnu autentifikaciju u elektronskim servisima. Na primer, najveći prodavac knjiga u svetu je Amazon. Elektronski servis za plaćanje kod Amazona podržava samo jednosmernu autentifikaciju u smeru Amazon korisnik. Na ovaj način, korisnik može da se uveri u legitimitet Amazon elektronskog servisa, ali Amazon ne vrši proveru da li se radi o legitimnom vlasniku kreditne kartice. Možemo da zaključimo da je, uglavnom, krajnji korisnik nezaštićen. Ovakvih slučajeva ima mnoštvo i u bankarskim elektronskim servisima za Internet plaćanja.

Razvoj ovog sistema omogućava jednostavnije generisanje i distribuciju kriptoloških ključeva, s druge strane, bilo bi omogućeno da svaki korisnik ili učesnik neke *online* transakcije bude maksimalno bezbedan, tako što svako za sebe u komunikaciji predstavlja sopstveno sertifikaciono telo koje ima mogućnost da generiše i verifikuje sopstveni sertifikat, kao i da opozove sertifikat u slučaju da je bezbednost narušena.

Dodatak A

U dodatku A predstavljeni su listinzi programskih kodova koji su napisani u Matlab-u za potrebe ove disertacije.

```
function [ decoded brojac p H ] = aproksimacija2( irisKod )
%APROKSIMACIJA 2 REDA
p = [.25 .25 .25 .25];
s = [1 2 3 4];
r = huffmandict(s,p);
irisKod = irisKod';
%I segment
irisKod = reshape (irisKod,1,4800);
decoded = huffmandeco(irisKod,r);
[brojac p] = hist(decoded,4);
p = brojac/2400;
H = -sum(p.*log2(p + (p == 0)));
%entropija po jednom bitu
H = H/2;
hist(decoded,4)
end
```

Listing 1. Šenonov model aproksimacije II reda

```
function [ decoded brojac p H ] = aproksimacija3( irisKod )
%APROKSIMACIJA 3 REDA
p = [.125 .125 .125 .125 .125 .125 .125 .125];
s = [1 2 3 4 5 6 7 8];
r = huffmandict(s,p);
irisKod = irisKod';
irisKod = reshape (irisKod,1,4800);
decoded = huffmandeco(irisKod,r);
[brojac p] = hist(decoded,8);
p = brojac/1600;
H = -sum(p.*log2(p + (p == 0)));
%entropija po jednom bitu
H = H/3;
hist(decoded,8)
end
```

Listing 2. Šenonov model aproksimacije III reda

```
function [ decoded brojac p H ] = aproksimacija4( irisKod )
    %APROKSIMACIJA 4 REDA
    p = [.0625 .0625 .0625 .0625 .0625 .0625 .0625 .0625
         .0625 .0625 .0625 .0625 .0625 .0625 .0625 .0625];
    s = [1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16];
    r = huffmandict(s,p);
    irisKod = irisKod';
    irisKod = reshape (irisKod,1,4800);
    decoded = huffmandeco(irisKod,r);
    [brojac p] = hist(decoded,16);
    p = brojac/1200;
    H = -sum(p.*log2(p + (p == 0)));
    %entropija po jednom bitu
    H = H/4;
    %hist(decoded,16)
end
```

Listing 3. Šenonov model aproksimacije IV reda

```
function [ decoded brojac p H ] = aproksimacija5( irisKod )
    %APROKSIMACIJA 5 REDA
    p = [.03125 .03125 .03125 .03125 .03125 .03125 .03125 .03125 .03125 .03125 .03125 .03125
         .03125 .03125 .03125 .03125 .03125 .03125 .03125 .03125 .03125 .03125 .03125 .03125
         .03125 .03125 .03125 .03125 .03125 .03125 .03125 .03125 .03125];
    s = [1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32];
    r = huffmandict(s,p);
    irisKod = irisKod';
    irisKod = reshape (irisKod,1,4800);
    decoded = huffmandeco(irisKod,r);
    [brojac p] = hist(decoded,32);
    p = brojac/960;
    H = -sum(p.*log2(p + (p == 0)));
    %entropija po jednom bitu
    H = H/5;
    hist(decoded,32)
end
```

Listing 4. Šenonov model aproksimacije V reda

```
function M = uzajamna_informacija(X,Y)
X = double(X);
Y = double(Y);

X_norm = X - min(X(:)) + 1;
Y_norm = Y - min(Y(:)) + 1;

matAB(:,1) = X_norm(:);
matAB(:,2) = Y_norm(:);
h = accumarray(matAB+1, 1); % združeni histogram

hn = h./sum(h(:)); % normalizovani združeni histogram
y_marg=sum(hn,1); % marginalne verovatnoće
x_marg=sum(hn,2); % marginalne verovatnoće

Hy = - sum(y_marg.*log2(y_marg + (y_marg == 0))); % Entropija od X
Hx = - sum(x_marg.*log2(x_marg + (x_marg == 0))); % Entropija od Y

arg_xy2 = hn.*(log2(hn+(hn==0)));
h_xy = sum(-arg_xy2(:)); % združena entropija H(X,Y)
M = Hx + Hy - h_xy; % uzajamna informacija I(X;Y)
```

Listing 5. Uzajamna informacija između dva irisa

```
function [ kljuc, pseudo_kod ] = RS_127_10_200_FP_GK( iris, pin )
    iris = prviSegment(iris);
    iris = iris';
    irisA = iris(1:60,:);
    irisB = iris(174:320,:);
    irisC = iris(405:480,:);
    iris = [reshape(irisA,1,540) reshape(irisB,1,1323) reshape(irisC,1,684)];
    % interleaver
    iris = randintrlv(iris, pin);
    % RS koder
    n = 127; k = 10;
    coder = fec.rsenc(n,k);
    % random kljuc K
    k = randint(1, 200, [0 1]);
    kljuc = k;
    k = vec2mat(k,10)';
    kodovani_kljuc = encode(coder,k);
    kodovani_kljuc = reshape(kodovani_kljuc,1,2540);
    pseudo_kod = bitxor(kodovani_kljuc, iris(1:2540));
end

function [ regenerisani_kljuc, greska ] = RS_127_10_200_FP_RK( i, pseudokod, pin, kljuc )
    n = 127; k = 10;
    coder = fec.rsenc(n,k);
    iris = firstPartW(i);
    iris = iris';
    irisA = iris(1:60,:);
    irisB = iris(174:320,:);
    irisC = iris(405:480,:);
    iris = [reshape(irisA,1,540) reshape(irisB,1,1323) reshape(irisC,1,684)];
    % interleaver
    iris = randintrlv(iris, pin);
    % RS dekode
    decoder = fec.rsdec(coder);
    codeKey = bitxor(pseudokod, iris(1:2540));
    codeKey = codeKey(1:2540)';
    codeKey = vec2mat(codeKey,127)';
    % regenerisanje ključa
    regenerisani_kljuc = decode(decoder,codeKey);
    regenerisani_kljuc = reshape(regenerisani_kljuc,1,200);
    regenerisani_kljuc = regenerisani_kljuc(1:200);
    greska = sum(sum(bitxor(regenerisani_kljuc, kljuc)));
end
```

Listing 6. Generisanje ključa veličine 200 bitova

```
function [ kljuc, pseudokod ] = RS_127_15_300_FP_GK( i, pin )
    iris = prviSegment(i);
    iris = iris';
    irisA = iris(1:60,:);
    irisB = iris(174:320,:);
    irisC = iris(405:480,:);
    iris = [reshape(irisA,1,540) reshape(irisB,1,1323) reshape(irisC,1,684)];

    % interleaver
    iris = randintrlv(iris, pin);

    % RS koder
    n = 127; k = 15;
    coder = fec.rsenc(n,k);

    % random ključ
    k = randint(1, 300, [0 1]);
    kljuc = k;
    k = vec2mat(k,15)';
    kodovani_kljuc = encode(coder,k);
    kodovani_kljuc = reshape(kodovani_kljuc,1,2540);
    pseudokod = bitxor(kodovani_kljuc, iris(1:2540));
end

function [ regenerisani_kljuc, greska ] = RS_127_15_300_FP_unlock( i, pseudokod, pin, kljuc )
    n = 127; k = 15;
    coder = fec.rsenc(n,k);
    iris = prviSegment(i);
    iris = iris';
    irisA = iris(1:60,:);
    irisB = iris(174:320,:);
    irisC = iris(405:480,:);
    iris = [reshape(irisA,1,540) reshape(irisB,1,1323) reshape(irisC,1,684)];%2540

    % interleaver
    iris = randintrlv(iris, pin);

    % RS dekoder
    decoder = fec.rsdec(coder);
    kodovani_kljuc = bitxor(pseudokod, iris(1:2540));
    kodovani_kljuc = kodovani_kljuc(1:2540)';
    kodovani_kljuc = vec2mat(kodovani_kljuc,127)';

    % regenerisanje ključa
    regenerisani_kljuc = decode(decoder,kodovani_kljuc);
    regenerisani_kljuc = reshape(regenerisani_kljuc,1,300);
    regenerisani_kljuc = regenerisani_kljuc(1:300);
    greska = sum(sum(bitxor(regenerisani_kljuc, kljuc)));
end
```

Listing 7. Generisanje ključa veličine 300 bitova

```
function [ kljuc, pseudokod ] = RS_127_20_400_FP_GK( i, pin )
    iris = prviSegment(i);
    iris = iris';
    irisA = iris(1:60,:);
    irisB = iris(174:320,:);
    irisC = iris(405:480,:);
    iris = [reshape(irisA,1,540) reshape(irisB,1,1323) reshape(irisC,1,694)];

    % interleaver
    iris = randintrlv(iris, pin);

    % RS koder
    n = 127; k = 20;
    coder = fec.rsenc(n,k);

    % generisanje ključa
    k = randint(1, 400, [0 1]);
    kljuc = k;
    k = vec2mat(k,20)';
    kodovani_kljuc = encode(coder,k);
    kodovani_kljuc = reshape(kodovani_kljuc,1,2540);
    pseudokod = bitxor(kodovani_kljuc, iris(1:2540));
end

function [ regenerisani_kljuc, greska ] = RS_127_20_400_FP_RK( i, pseudokod, pin, kljuc )
    n = 127; k = 20;
    coder = fec.rsenc(n,k);
    iris = firstPartW(i);
    iris = iris';
    irisA = iris(1:60,:);
    irisB = iris(174:320,:);
    irisC = iris(405:480,:);
    iris = [reshape(irisA,1,540) reshape(irisB,1,1323) reshape(irisC,1,694)];

    % interleaver
    iris = randintrlv(iris, pin);

    % RS dekoder
    decoder = fec.rsdec(coder);
    kodovani_kljuc = bitxor(pseudokod, iris(1:2540));
    kodovani_kljuc = kodovani_kljuc(1:2540)';
    kodovani_kljuc = vec2mat(kodovani_kljuc,127)';

    % regenerisanje ključa
    regenerisani_kljuc = decode(decoder,kodovani_kljuc);
    regenerisani_kljuc = reshape(regenerisani_kljuc,1,400);
    regenerisani_kljuc = regenerisani_kljuc(1:400);
    greska = sum(sum(bitxor(regenerisani_kljuc, kljuc)));
end
```

Listing 8. Generisanje ključa veličine 400 bitova

Skraćenice

BK – biometrijska kriptografija (eng. *biometric cryptography*)

FRR – stopa lažnog odbacivanja (eng. *False Reject Rate*)

FAR – stopa lažnog prihvatanja (eng. *False Accept Rate*)

ECC – kodovi za detektovanje i ispravljanje grešaka (eng. *Error Correcting Code*)

RS – Rid-Solomonov kod za ispravljanje grešaka (eng. *Reed-Solomon*)

H – entropija (eng. *entropy*)

I – informacija (eng. *information*)

DOF – broj stepeni slobode (eng. *Degree Of Freedom*)

CA – sertifikaciono telo (eng. *Certification Authority*)

BF – broj izračunavanja metodom potpune pretrage ključa (eng. *Brute Force*)

Bibliografija

- [1] *Combining Crypto with Biometrics Effectively*. **Feng Hao, Ross Anderson, John Daugman**. s.l. : IEEE Computer Society, 2006. IEEE TRANSACTIONS ON COMPUTERS. str. 1081-1088.
- [2] *Face Biometrics with Renewable Templates*. **M. van der Veen, T. Kevenaar, G.-J. Schrijen, T. H. Akkermans, and Fei Zuo**. San Jose : s.n., 2006. Proceedings of SPIE, Volume 6072: Security, Steganography, and Watermarking of Multimedia Contents VIII, Edward J. Delp III, Ping Wah Wong, Editors, 60720J.
- [3] *On Enabling Secure Applications Through Off-line Biometric Identification*. **David, G.I., Yair Frankel, Brian J. Matt**. Oakland, CA : Security and Privacy, 1998. Proceedings. 1998 IEEE Symposium on, 1998. str. 148.
- [4] *A fuzzy commitment scheme*. **Ari Juels, Martin Wattenberg**. New York : ACM, 1999. CCS '99 Proceedings of the 6th ACM conference on Computer and communications security. str. 28-36.
- [5] *Enhancing security and privacy in biometrics-based authentication systems*. **N. K. Ratha, J. H. Connell, R. M. Bolle**. 3, s.l. : IBM SYSTEMS JOURNAL, 2001, T. 40. 0018-8670.
- [6] *Feature-Based Image Metamorphosis*. **T. Beier, S. Neely**. New York : s.n., 1992. str. 35–42.
- [7] *Cancelable Iris Biometric*. **Jinyu Zuo, Nalini K. Ratha, and Jonathan H. Connell**. s.l. : ICPR 2008. 19th International Conference, 2008. In proceeding of: Pattern Recognition.
- [8] *MATLAB Source Code for a Biometric Identification System Based on Iris Patterns*. **Libor Masek, Peter Kovesi**. 2003.
- [9] *Cancelable Iris Biometrics and Using Error Correcting Codes to Reduce*. **Sanjay Kanade, Dijana Petrovska-Delacretaz, Bernadette Dorizzi**. Miami, FL : Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on, 2009. pp. 120-127. 978-1-4244-3991-1.
- [10] *Tokenless Cancelable Biometrics Scheme for Protecting IrisCodes*. **Osama Ouda, Norimichi Tsumura, Toshiya Nakaguchi**. Istanbul : 2010 International Conference on Pattern Recognition, 2010. Pattern Recognition (ICPR), 2010 20th International Conference on. str. 882- 885.
- [11] *Fuzzy Extractors: How to Generate Strong Keys from Biometrics and Other Noisy Data*. **Y. Dodis, L. Reyzin, A. Smith**. s.l. : Eurocrypt 2004, 2008. Eurocrypt 2004. str. 523-540.

- [12] *Reusable Cryptographic Fuzzy Extractors*. **Xavier Boyen**, Voltage Security, Palo Alto. New York : ACM Press, 2004. In 11th ACM Conference on Computer and Communications Security (CCS 2004). str. 82-91.
- [13] *Biometric Fuzzy Extractor Scheme*. **F. Hernandez Alvarez, L. Hernandez Encinas, C. Sanchez Avila**. Las Vegas : WORLDCOMP'09, 2009. The 2009 International Conference on Security and Management (SAM'09). str. 563-569.
- [14] *Password Hardening Based on Keystroke Dynamics*. **F. Monrose, M.K. Reiter, R. Wetzel**. 1999. Proc. Sixth ACM Conf. Computer and Comm. Security (CCCS).
- [15] *Cryptographic Key Generation from Voice*. **F. Monrose, M.K. Reiter, Q. Li, S. Wetzel**. 2001. Proc. 2001 IEEE Symp. Security and Privacy, May 2001.
- [16] *Private Key Generation from On-Line Handwritten Signatures*. **F. Hao, C.W. Chan**. 2002, Information Management & Computer Security, str. 159-164.
- [17] *Computation of Cryptographic Keys from Face Biometrics*. **A. Goh, D.C. L. Ngo**. 2003. Proc. Int'l Federation for Information Processing 2003. str. 1-13.
- [18] *Secure Smart Card-Based Fingerprint Authentication*. **T.C. Clancy, N. Kiyavash, D.J. Lin**. 2003. Proc. 2003 ACM SIGMM Workshop Biometrics Methods and Application (WBMA).
- [19] *Biometric Encryption*. **C. Soutar, D. Roberge, A. Stoianov, R. Gilroy, and B.V.K. Vijaya Kumar**. 1999. ICSA Guide to Cryptography, McGraw-Hill.
- [20] *Virtual PIN: Biometric Encryption Using Coding Theory*. *BIOSIG: Biometric and Electronic Signatures*. **U. Martini, S. Beinlich**. Darmstadt : s.n., 2003. the 1st Conference on Biometrics and Electronic Signatures of the GI Working Group BIOSIG. str. 91-99.
- [21] *Practical Biometric Authentication with Template Protection*. **P. Tuyls, A. H. M. Akkermans, T. A. M. Kevenaar, G.-J. Schrijen, A. M. Bazen, and R. N.J. Veldhuis**. Newyork : s.n., 2005. 5th International Conference, AVBPA. str. 20-22.
- [22] *Bimodal Biometric System for Cryptographic Key Generation Using Wavelet*. **H. A. Garcia-Baleon, V. Alarcon-Aquino, O. Starostenko, J. F. Ramirez-Cruz**. 2009. 2009 Mexican International Conference on Computer Science. str. 186-196.
- [23] *The Importance of Being Random: Statistical Principles of Iris Recognition*. **Daugman, J.** 2003, Pattern Recognition, str. 279-291.
- [24] *On the Relation of Error Correction and Cryptography to an Off Line Biometrics Based Identification Scheme*. **G.I. Davida, Y. Frankel, B.J. Matt, and R. Peralta**. 1999. Proc. Workshop Coding and Cryptography.
- [25] *Hadamard Matrix and Their Applications*. **Agaian, S.S.** s.l. : Springer Verlag, 1985.

- [26] *Three Factor Scheme for Biometric-Based Cryptographic Key Regeneration Using Iris*. **Sanjay K, Danielle C, Emine K, Dijana P.D, and Bernadette D.** France : TELECOM & Management SudParisEvry, 2008.
- [27] *Multimedia content protection via biometrics-based encryption*. **Umut Uludug and Anil K. Jain.** East Lansing : Computer Science and Engineering Department, Michigan State University, 2003. str. 237-240.
- [28] *Biometric authentication for web-based course examinations*. **Brent Auernheimer, Max J. Tsai.** Fresno : s.n., 2005. Proceedings of the 38th Hawaii International Conference on System Sciences - 2005.
- [29] *A Novel Digital Content Protection System Based on Iris Biometric*. **Meihua Wang, Kefeng Fan, Qingqi Pei, Shuguang Wang, Shan Cao.** 2007. Fourth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD 2007).
- [30] *Key Management Using Biometrics*. **Haiyong Chen, Hongwei Sun, Kwok-Yan Lam.** Beijing : 2007 IEEE, 2007. str. 321-326.
- [31] *PKI Key Generation Based on Iris Features*. **Yazhuo Gong, Kaifa Deng, Pengfei Shi.** Shanghai : s.n., 2008. 2008 International Conference on Computer Science and Software Engineering. str. 166-169.
- [32] *Bio-Chaotic Stream Cipher-Based Iris Image Encryption*. **Abdullah Sharaf Alghamdi, Hanif Ullah, Maqsood Mahmud, Muhammad Khurram Khan.** Riyadh : s.n., 2009. 2009 International Conference on Computational Science and Engineering. str. 739-744.
- [33] *A mathematical theory of communication*. **Shannon, C. E.** 1948, Bell System Tehnical Journal, str. 379-423.
- [34] *Communication Theory of Secrecy*. **Shannon, C. E.** 1949, Bell system Tehnical Journal, str. 656-715.
- [35] *A new Outlook on Shannon Information Measures*. **Raymond W, Yueng.** 1991. IEEE Transactions on IT. T. 37.
- [36] **Shannon, Claude E.** *Collected Papers*. s.l. : Wiley-IEEE Press; 1 edition (January 5, 1993), 1993.
- [37] *A system for automated iris recognition*. **R. Wildes, J. Asmuth, G. Green, S. Hsu, R. Kolczynski, J. Matey, S. McBride.** Sarasota : s.n., 1994. Proceedings IEEE Workshop on Applications of Computer Vision. str. 121-128.
- [38] *Image representation using 2D gabor wavelets*. **Lee., T.** IEEE Transactions of Pattern Analysis and Machine Intelligence. str. 1996.

- [39] **Masek, Libor.** Recognition of Human Iris Patterns for Biometric Identification. *Iris Recognition*. [Na mreži] 30 11 2003. <http://www.csse.uwa.edu.au/~pk/studentprojects/libor/>.
- [40] **Reed-Solomon Code Performance for M-ary Modulation over AWGN Channel. Saurabh Mahajan, Gurpadam Singh.** 2011, International Journal of Engineering Science and Technology (IJEST), str. 3739-3745.
- [41] *Digital Image Processing Using MATLAB. Gonzalez, R.C., R.E. Woods, S.L. Eddins.* New Jersey, Prentice Hall : s.n., 2003.
- [42] **Tomko., G.J.** *Method and apparatus for securely handling data in a database of biometrics and associated data.* 5790668 U.S., 1998.
- [43] **Tiberg., M.** *A Method and a System for Biometric Identification or Verification.* 0202147-5 Swedish, 2002.
- [44] **Tan, Tieniu.** Center for Biometrics and Security Research. [Na mreži] [Citirano: 15 10 2008.] <http://biometrics.idealtest.org/dbDetailForUser.do?id=4>.
- [45] **Stoianov., G.J. Tomko and A.** *Method and apparatus for securely handling a personal.* 5712912 U.S., 1998.
- [46] **R.M. Bolle, J.H.Connel, and N.K.Ratha.** *System and method for distorting a biometric for transactions with enhanced security and privacy.* 6,836,554 U.S., 2004.
- [47] *Cancelable biometric filters for face recognition. M. Savvides, B.V.K.Vijaya Kumar and P.K.Khosla.* Cambridge : s.n., 2004. Proceedings of the 17th International Conference on Pattern Recognition (ICPR'04). str. 922–925.
- [48] **G.J. Tomko, C. Soutar, and G.J. Schmidt.** *Fingerprint controlled public key cryptographic system.* 5541994 U.S., 1996.
- [49] —. *Biometric controlled key generation.* 5680460 U.S., 1997.
- [50] **C. Soutar, D. Roberge, A.V. Stoianov, R. Gilroy, and B. V. K. Vijaya Kumar.** *Method for secure key management using a biometric.* 6219794 U.S., 2001.
- [51] **Bodo., A.** *Method for producing a digital signature with aid of a biometric feature.* DE 42 43 908 A1 Germany, 1994.
- [52] **Bjorn., V.** *Cryptographic key generation using biometric data.* 6035398 U.S., 2000.
- [53] **Berlekamp, E.R., and P. Tong.** *Improved Interleavers for Algebraic Block Codes.* 4559625 U. S., 17 12 1985.
- [54] *Vulnerabilities in Biometric Encryption Systems. Adler., A.* 2005. Audio- and video-based Biometric Person Authentication (AVBPA). str. 1100–1109.

- [55] *Visual Cryptography*. **Moni Naor and Adi Shamir**. s.l. : EUROCRYPT , 1994.
- [56] *Using Voice to Generate Cryptographic Keys*. **F. Monrose, M.K. Reiter, Q. (Peter) Li , and S. Wetzel**. Crete : s.n., 2001. A Speaker Odyssey. The Speech Recognition Workshop.
- [57] *Using Threshold Attribute-Based Encryption for Practical Biometric-Based Access Control*. **D. Nali, C. Adams, and A. Miri**. 2005. International Journal of Network Security, Vol.1, No.3. str. 173–182.
- [58] *Toward speech-generated cryptographic keys on resource constrained devices*. **F. Monrose, M. K. Reiter, Q. Li, D. P. Lopresti, and C. Shih**. 2002. Proc. 11th USENIX Security Symp. str. 283–296.
- [59] *Teaching interactive cryptography: the case for CrypTool*. **Milan Milosavljević, Saša Adamović, Irina Branović, Dejan Živković**. s.l. : ICEST, 2011. Communication and Energy Systems and Technologies - ICEST. str. 46.
- [60] *Symmetric hash functions for fingerprint minutiae*. **S. Tulyakov, F. Farooq, and V. Govindaraju**. 2005. Lecture Notes in Computer Science, vol. 3687. str. 30-38.
- [61] **Q. Li, Y. Sutcu, and N. Memon**., Secure sketch for biometric templates. *Advances in Cryptology – ASIACRYPT 2006*. 2006, str. 99-113.
- [62] *Secure private key generation using a fingerprint*. **C. Soutar and G.J. Tomko**. 1996. In CardTech/SecurTech Conference Proceedings. str. 245-252.
- [63] *Secure fuzzy vault based fingerprint verification system*. **S. Yang and I. Verbauwhede**. 2004. In Thirty-Eighth Asilomar Conference on Signals, Systems, and Computers (2004). str. 577–581.
- [64] *Robust and Secure Biometrics: Some Application Examples*. **T. Kevenaar, G.J. Schrijen, A. Akkermans, M. Damstra, P. Tuyls, and M. van der Veen**. Rome : s.n., 2006. Information Security Solutions Europe (ISSE) Conference.
- [65] *Reliable biometric authentication with privacy protection*. **E. Verbitskiy, P. Tuyls, D. Denteneer, and J.-P. Linnartz**. Veldhoven : s.n., 2003. In Proc. of the 24th Symp. on Inf. Theory in the Benelux. str. 125–132.
- [66] *Privacy protecting biometric authentication systems: an overview*. **P. Tuyls, E. Verbitskiy, J. Goseling, D. Denteneer**. Vienna : s.n. XII European Signal Processing Conference (EUSIPCO 2004, Vienna,Austria). str. 1397–1400.
- [67] **T. Connie, A. Teoh, M. Goh, and D. Ngo**. PalmHashing: a novel approach for cancelable biometrics. *Information Processing Letters* 93. 2005, str. 1–5.

- [68] *On enabling secure applications through off-line biometric identification.* **G.I. Davida, Y. Frankel, and B.J. Matt.** Oakland : s.n., 1998. Proc. of the IEEE 1998 Symp. on Security and Privacy. str. 148–157.
- [69] *Novel biometric digital signature for Internet based applications.* **P.K. Janbandhu and M.Y. Siyal.** 2001, Information Management and Computer Security, str. 205–212.
- [70] *New shielding functions to enhance privacy and prevent misuse of biometric templates.* **J.-P. Linnartz and P. Tuyls.** 2003. In Proc. of the 4th Int. Conf. on Audio and Video Based Biometric Person Authentication. str. 393–402.
- [71] *MultiHashing, human authentication featuring biometrics data and tokenised random number: a case study FVC2004.* **D. Maio and L. Nanni.** 2005, NeuroComputing, vol. 69, str. 242-249.
- [72] *Iris Biometric Cryptography for Identity Document.* **Sim Hiew Moi, Nazeema Binti Abdul Rahim, Puteh Saad, Pang Li Sim, Zalmiyah Zakaria.** 2009. 2009 International Conference of Soft Computing and Pattern Recognition. str. 736-741.
- [73] *Human authentication featuring signatures and tokenised random number.* **L. Nanni and A. Lumini.** 2006, NeuroComputing , vol.69, no.7-9, str. 858-861.
- [74] *How to Protect Biometric Templates.* **Y. Sutcu, Q. Li, and N. Memon.** San Jose : SPIE Conf. on Security, Steganography and Watermarking of Multimedia Contents IX, January 2007, San Jose, CA.Proceedings of SPIE, v. 6505. Editors: Edward J. Delp III, Ping Wah Wong, 2007.
- [75] *Fuzzy Vault for Fingerprints.* **U. Uludag, S. Pankanti, A. K. Jain.** 2005, AVBPA 2005 : audio- and video-based biometric person authentication (Hilton Rye Town NY, 20-22 July 2005).Springer, str. 310-319.
- [76] *Fuzzy identity based encryption.* **A. Sahai and B. Waters.** 2005. Proceedings of EUROCRYPT'05 on Advances in Cryptology, LNCS 3494. str. 457–473.
- [77] *Finding the Original Point Set Hidden among Chaff.* **E.-C. Chang, R. Shen, and F. W. Teo.** Taipei : s.n., 2006. Proceedings of the 2006 ACM Symposium on Information, computer and communications security. ASIACCS'06 March 21-24. str. 182–188.
- [78] *Error-Correction Coding for Digital Communications.* **Clark, George C., Jr., and J. Bibb Cain.** 1981, Plenum Press.
- [79] **Aiden A. Bruen.** *Cryptography, information theory, and error-correction.* San Diego : WILEY-INTERSCIENCE , 2005.
- [80] *Capacity and examples of template protecting biometric authentication systems.* **P. Tuyls and J. Goseling.** 2004. Biometric Authentication Workshop. str. 158-170.

- [81] *Cancelable key-based fingerprint templates*. **R. Ang, R. Safavi-Naini, and L. McAven.** 2005. ACISP: Australasian conference on information security and privacy No10, Brisbane, Lecture Notes in Computer Science, vol. 3574. str. 242-252.
- [82] *Biometrics: A Tool for Information Security*. **A. K. Jain, A. Ross, and S. Pankanti.** 2006. IEEE transactions on information forensics and security, vol. 1, No. 2, June 2006. str. 125-143.
- [83] *Biometrics based Asymmetric Cryptosystem Design Using Modified Fuzzy Vault Scheme*. **A. Nagar, S. Chaudhury.** 2006. 18th International Conference on Pattern Recognition (ICPR'06). str. 537-540.
- [84] *Biometric perils and patches*. **R.M.Bolle, J.H.Connel, and N.K.Ratha.** 2002, Pattern Recognition, str. 2727-2738.
- [85] *Biometric Hash: High-Confidence Face Recognition*. **D. C. L. Ngo, A. B. J. Teoh, and A. Goh.** 2006. IEEE Transactions on circuits and systems for video technology, vol. 16, No. 6, June 2006. str. 771-775.
- [86] **Ann Cavoukian, Alex Stoianov.** *Biometric Encryption: A Positive-Sum Technology that Achieves Strong Authentication, Security AND Privacy*. Toronto : Information and Privacy Commissioner of Ontario, 2007.
- [87] *Biometric Encryption using image processing*. **C. Soutar, D. Roberge, A.V. Stoianov, R. Gilroy, and B. V. K. Vijaya Kumar.** 1998. Proc.SPIE, Optical Security and Counterfeit Deterrence Techniques II, vol. 3314. str. 178-188.
- [88] *Biometric Encryption - Enrollment and Verification Procedures*. **C. Soutar, D. Roberge, A.V. Stoianov, R. Gilroy, and B. V. K. Vijaya Kumar.** 1998. SPIE, Optical Pattern Recognition IX, v. 3386. str. 24 - 35.
- [89] **C. Soutar, D. Roberge, A. Stoianov, R. Gilroy and B.V.K. Vijaya Kumar.** *Biometric Encryption*. s.l. : ICSA Guide to Cryptography, McGraw-Hill, 1999.
- [90] *Biometric Data Hiding: A 3 Factor Authentication Approach to Verify Identity with a Single Image Using Steganography, Encryption and Matching*. **Neha Agrawal and Marios Savvides.** Pittsburgh : 2009 IEEE, 2009. IEEE. str. 85-92.
- [91] *Biometric Cryptosystems: Issues and Challenges*. **U. Uludag, S. Pankanti, S. Prabhakar, and A. K. Jain.** Proceedings of the IEEE, v. 92, no. 6, June 2004. str. 948-960.
- [92] **A. Teoh , D. Ngo, and A. Goh.** Biohashing: two factor authentication featuring fingerprint data and tokenised random number. *Pattern Recognition* 37. 2004, str. 2245-2255.

- [93] *Application-specific biometric templates*. **M. Braithwaite, U.C. von Seelen, J. Cambier, J. Daugman, R. Glass, R. Moore, and I.Scott**. Tarrytown, NY : s.n., 2002. IEEE Workshop on Automatic Identification Advanced Technologies. str. 167-171.
- [94] **F. Bavaud J.C. Chappelier, J. Kohlas**. *An Introduction to Information Theory*. s.l. : UniFr, 2005.
- [95] **A. Lumini and L. Nanni**. An improved BioHashing for human authentication. *Pattern Recognition*, vol.40, no.3. 2006, str. 1057-1065.
- [96] *An advanced multi-modal method for human a featuring biometrics data and tokenised random numbers*. **L. Nanni and A. Lumini**. 2006, NeuroComputing, vol. 69, no.13, str. 1706-1710.
- [97] *A schema for cryptographic keys generation using hybrid biometrics*. **K. Vorderhobli, C. Pattinson, and H. Donelan**. Liverpool : s.n., 2006. In: 7th annual postgraduate symposium: The convergence of telecommunications, networking and broadcasting. str. 26-27.
- [98] *A Novel Biometric Key Scheme*. **Q. Li, X. Niu, and S. Sun**. (Chinese Journal of Electronics2005).
- [99] *A Biometric Identity Based Signature Scheme*. **A. Burnett, F. Byrne, T. Dowling, and A. Duffy**. New York : s.n., 2005. Applied Cryptography and Network Security Conference.
- [100] **Alfred J. Menezes, Paul C. van Oorschot, Scot A. Vanstone**. *Handbook of Applied Cryptography*. New Jersey : CRC Press, 2001.
- [101] **Alexander W. Dent, Chris J. Mitchell**. *User's Guide to Cryptography and Standards, Computer Security Series*. Boston : Artech House, 2005.
- [102] **Stamp, Mark**. *Information Security*. 2nd Edition. New Jersey : John Wiley & Soons, 2011.
- [103] **Schneier, Bruce**. *Applied Cryptography: Protocols, Algorithms, and Source Code in C*. 2nd Edition. Canada : John Wiley & Soons, 1996.
- [104] **Niels Ferguson, Bruce Schneier, Tadayoshi Kohno**. *Cryptography Engineering, Design, Principles and Practical Application*. Canada : John Wiley & Soons, 2010.
- [105] **Fred Piper and Sean Murphy**. *Cryptography: A Very Short Introduction*. London : Oxford University Press, 2002.

Objavljeni radovi u naučnim časopisima i konferencijama

Kao autor ili koautor objavio je više od 29 radova na domaćim i međunarodnim konferencijama i časopisima:

Radovi u međunarodnim časopisima:

1. Saša Adamović, Milan Milosavljević „*Information Analysis of Iris Biometrics for the Needs of Cryptology Key Extraction*“ [Časopis] = SJEE // Serbian Journal of Electrical Engineering. - February 2013. - 1 : T. 10. - str. 1-12. - 10.2298/SJEE1301001A.
2. S. Adamović, M. Šarac, M. Veinović, M. Milosavljević, A. Jevremović, „*An Interactive and Collaborative Approach to Teaching Cryptology*“ [Časopis] = IFETS // Journal of Educational Technology & Society, 1436-4522 (u procesu objavljivanja)
3. M. Šarac, T. Radojevic, N. Stanišić, S. Adamović, D. Radovanovic, „*Safety of e-business Applications in Serbia: Applied Knowledge Based on SSL Traffic*“ [Časopis] = JIBC // Journal of Internet Banking and Commerce. - December 2012, vol. 17, no.3.
4. D. Radovanovic, T. Radojević, S. Adamović, M. Šarac, „*Information technology audit process and cobit*“ [Časopis] = MI // Metalurgia International, ISBN 1582-2214, Romania, 2013.

Udžbenici:

1. Mladen Veinović, Saša Adamović „*Kriptologija I – Osnove za analizu i sintezu šifarskih sistema*“ [Knjiga]. - Beograd : Univerzitet Singidunum, 2013. - 1 : T. 1 : str. 247. - 978-86-7912-469-2.

Saopštenja sa domaćeg-međunarodnog skupa – konferencije:

1. Adamović S., Veinović M., „Analiza zaštite podataka u bazama podataka na lokalnom i serverskom nivou“, 16. TELFOR, Beograd, ID=1026, 2008.
2. Mladen V., Adamović S, „Uticaj kriptografskih algoritama na bazu podataka“, Singidunum revija, Vol.5 No 2, Beograd, 2009, pp.98-105, UDK: 004.383.2.056, ISSN: 1820-8819.
3. Adamović S., Veinović M., Šarac M., „Uticaj virtualizacije na performanse pri zaštiti podataka u bazama podataka sa kriptografskim algoritmima“, Sinergija 2009, Bijeljina, pp.143-145, 2009.

4. Šarac M., Veinović M., Adamović S., „Komparacija i analiza resursa fizičkog i virtualnog okruženja datacentra”, Sinergija 2009, Bijeljina, pp.131-134, 2009.
5. Saša Adamović, Marko Šarac, Dalibor Radovanović, Miloš Milenković: „*Data protection in database at the local and server level*“, Međunarodna konferencija, Bugarska - Gabrovo – Unitech 09.
6. D. Radovanović, M. Šarac, S. Adamović, M. Milenković: „*Analysis and classification of attacks on the security of smart cards with additional emphasis on security pin code*“, Međunarodna konferencija, Bugarska - Gabrovo – Unitech 09.
7. Radovanović D., Radojević T., Adamović S., Šarac M., „*One concept for IT audit*” UNITECH 2010 - International Scientific Conference, 19 – 20 November 2010, Gabrovo, Bulgaria, ISBN 1313-230X, Vol. I, p.477-482.
8. Šarac M., Radovanović D., Adamović S., „Rešavanje problema uskog grla virtuelnih servera, virtuelnim iSCSI uređajima za skladištenje podataka“, Singidunum revija, Vol. 6, No. 2, Beograd, 2010, ISSN 1820-8819, 193-198
9. Radovanović D., Šarac M., Lučić D., Adamović S., „Analiza metoda za kontrolu i reviziju informacionih sistema“, –, INFOTEH 2010, Mart 17-19, 2010, Jahorina, Bosna, CD Zbornik radova, ISBN-99938-624-2-8, INFOTEH-JAHORINA Vol. 9, Ref. E-III-2, p. 567-570.
10. Adamović S., Šarac M., Radovanović D., Radojević T., „*Secret key extraction over unauthenticated public channels*” UNITECH 2010 - International Scientific Conference, 19 – 20 November 2010, Gabrovo, Bulgaria, ISBN 1313-230X, Vol. I, p.472-476
11. Adamović S., Šarac M., Milenković M., Radovanović D., „Generatori slučajnih sekvenci i njihov uticaj na sigurnost“, INFOTEH 2010, Mart 17-19, 2010, Jahorina, Bosna, CD Zbornik radova, ISBN-99938-624-2-8, Vol. 9, Ref. E-VI-1, p.820-822
12. Milenković M., Adamović S., Šarac M., Radovanović D., „Upravljanje X.509 sertifikatima u PKI sistemu“, ZITEH 10, Beograd, Mart 5-6, 2010, CD Zbornik, ISBN 978-86-90951111-6
13. Radovanović D., Šarac M., Adamović S., „Cobit pristup reviziji informacionih sistema“, XII međunarodni simpozijum Fakulteta organizacionih nauka, SYMORG 2010, Zlatibor, Jun 9 - 12, 2010
14. Šarac M., Radovanović D., Adamović S., „Analiza poslovne vrednosti serverske virtuelizacije“, XII međunarodni simpozijum Fakulteta organizacionih nauka, SYMORG 2010, Zlatibor, Jun 9 - 12, 2010

15. Adamović S., Šarac M., Radovanović D., „Generatori slučajnih brojeva i njihov uticaj na sigurnost“, XII međunarodni simpozijum Fakulteta organizacionih nauka, SYMORG 2010, Zlatibor, Jun 9 - 12, 2010
16. Mihailović V., Šarac M., Adamović S., Radovanović D., „Video konferencijski sistemi i njihova primena u obrazovnom sistemu“, Sinergija 2010, 19 mart 2010, Bijeljina, Bosnia, ISBN 978-99955-26-21-4, p. 122-128
17. Šarac M., Adamović S., Radovanović D., „Analiza sigurnosti bežičnih mreža IEEE 802.11. na teritoriji grada Beograda“, Infoteh 2011, Mart 16-18, Jahorina, Bosnia, ISBN-99938-624-2-8, Vol. 10, Ref. B-III-1, p. 191-194
18. Radovanović D., Šarac M., Adamović S., Lučić D., „*Necessity of IT Service Management and IT Governance*“ – IEEE Conference, MIPRO 2011- DE, May 23-27, 2011, Opatija, Croatia, ISSN 1847-3938, p. 84-87
19. Adamović S., Branović I., Živković D., Tomašević V., Milosavljević M. „*Teaching interactive cryptography: the case for CrypTool*“ – IEEE Conference, ICEST 2011, Jun 29, 2011, Niš, Serbia
20. Stamenković D., Šarac M., Adamović S., Radovanović D., „Energetska efikasnost u primeni virtualnog datacentra i upotreba ekoloških računara u edukativnim ustanovama“, Sinergija 2012
21. Šarac M., Veinović M., Adamović S., Radovanović D., Jevremović A. „Analiza sigurnosti ssl saobraćaja u bežičnim računarskim mrežama“, Infoteh 2012, pp 760-765 Jahorina, Bosnia, ISBN-99938-624-2-8, Vol. 11, Ref. B-III-1
22. Veinović M., Šarac M., Adamović S., „Bezbednosni problemi u zaštićenim bežičnim komunikacijama“, maj-mesec matematike M3, SANU 2012, Beograd, Srbija
23. Adamović S., Milosavljević M., „Teorijsko informaciona analiza iris biometrije za potrebe ekstrakcije kriptoloških ključeva“, 56 konferencija ETRAN 2012, Zlatibor 11-14 juna 2012. Godine
24. Radovanović D., Lučić D., Adamović S., Šarac M., „*What is IT Governance?*“, UNITECH 2012 - International Scientific Conference, 16 – 17 November 2012, Gabrovo, Bulgaria

