

Analiza kretanja korisnika temeljem anonimiziranih lokacijskih podataka telekomunikacijske mreže

Žaja, Antonia

Master's thesis / Diplomski rad

2024

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Electrical Engineering and Computing / Sveučilište u Zagrebu, Fakultet elektrotehnike i računarstva**

Permanent link / Trajna poveznica: <https://urn.nsk.hr/urn:nbn:hr:168:851132>

Rights / Prava: [In copyright](#) / [Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2025-03-29**



Repository / Repozitorij:

[FER Repository - University of Zagreb Faculty of Electrical Engineering and Computing repository](#)



SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

DIPLOMSKI RAD br. 432

**ANALIZA KRETANJA KORISNIKA TEMELJEM
ANONIMIZIRANIH LOKACIJSKIH PODATAKA
TELEKOMUNIKACIJSKE MREŽE**

Antonia Žaja

Zagreb, lipanj 2024.

SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

DIPLOMSKI RAD br. 432

**ANALIZA KRETANJA KORISNIKA TEMELJEM
ANONIMIZIRANIH LOKACIJSKIH PODATAKA
TELEKOMUNIKACIJSKE MREŽE**

Antonia Žaja

Zagreb, lipanj 2024.

DIPLOMSKI ZADATAK br. 432

Pristupnica: **Antonia Žaja (0016135912)**
Studij: Računarstvo
Profil: Programsko inženjerstvo i informacijski sustavi
Mentor: izv. prof. dr. sc. Marin Vuković

Zadatak: **Analiza kretanja korisnika temeljem anonimiziranih lokacijskih podataka telekomunikacijske mreže**

Opis zadatka:

U današnje vrijeme lokacija je postala jedan od najosjetljivijih osobnih podataka korisnika. Poznavanje lokacije korisnika omogućava dublje razumijevanje njihovih navika, interesa i ponašanja. Nadalje, analiza lokacijskih podataka omogućuje pružanje personaliziranih usluga i sadržaja, ciljanje oglasa, planiranje rute, ali isto tako postavlja izazove u pogledu zaštite privatnosti korisnika i etičke uporabe tih podataka. Stoga je važno razvijati napredne metode anonimizacije i sigurnosne mehanizme kako bi se osigurala zaštita korisničke privatnosti. Vaš je zadatak istražiti različite metode anonimizacije podataka te razviti tehnike analize kretanja korisnika temeljene na prikupljenim, anonimiziranim, lokacijskim podacima. Cilj je pružiti različitim vrstama usluga korisne uvide u ponašanje korisnika kako bi iste prilagodili, ali i poboljšali korisničko iskustvo, istovremeno štiteći privatnost korisničkih podataka sukladno važećim propisima o zaštiti podataka.

Rok za predaju rada: 28. lipnja 2024.

Zahvaljujem se svome mentoru izv.prof.dr.sc. Marinu Vukoviću na stručnom vodstvu i podršci pruženoj tijekom izrade ovog diplomskog rada. Posebno hvala mojim roditeljima na bezuvjetnoj ljubavi, podršci i neprestanoj vjeri u mene. Hvala Vam što ste uvijek bili uz mene i pružali mi sve što je bilo potrebno da ostvarim svoje ciljeve. Zahvaljujem se baki Ruži na ljubavi i podršci koju mi je uvijek pružala. Veliko hvala mom ujaku Damiru za pomoć u svakom trenutku mog školovanja. Također, zahvaljujem se ostalim članovima obitelji i prijateljima na razumijevanju, pomoći i motivaciji koja me poticala na napredak. Bez Vas i Vaše podrške, ovaj put ne bi bio isti. Hvala Vam svima od srca.

Sadržaj

Uvod	1
1. Primjeri povreda privatnosti lokacijskih podataka	3
2. Postojeća rješenja	5
3. Zapisi pojedinosti telefonskog poziva	8
3.1. Format CDR zapisa	8
4. Analiza kretanja korisnika	11
4.1. Priprema podataka za analizu kretanja korisnika	11
4.2. Definiranje lokacije korisnika	11
4.3. Posjećenost lokacije u vremenskom rasponu	11
4.4. Posjećenost korisnika na određenoj lokaciji	13
4.5. Posjećenost određene lokacije s obzirom na dobnu skupinu korisnika	14
4.6. Posjećenost lokacija s obzirom na lokaciju stanovanja	16
4.7. Posjećenost lokacija s obzirom na lokaciju posla	17
4.8. Posjećenost lokacija s obzirom na spol	18
4.9. Posjećenost lokacija s obzirom na broj upućenih poziva	19
4.10. Posjećenost lokacija s obzirom na broj primljenih poziva	20
4.11. Brzina kretanja	21
5. Potencijalni problemi implementacije transformacija u raznovrsnim uslugama	23
5.1. Posjećenost lokacije u vremenskom rasponu	23
5.2. Posjećenost korisnika na određenoj lokaciji	23
5.3. Posjećenost određene lokacije s obzirom na dobnu skupinu korisnika	23
5.4. Posjećenost lokacija s obzirom na lokaciju stanovanja	24

5.5.	Posjećenost lokacija s obzirom na lokaciju posla	24
5.6.	Posjećenost lokacija s obzirom na spol.....	24
5.7.	Posjećenost lokacija s obzirom na broj upućenih poziva.....	25
5.8.	Posjećenost lokacija s obzirom na broj primljenih poziva.....	25
5.9.	Brzina kretanja.....	25
5.10.	Potencijalni problemi kod općenite analize CDR zapisa	26
6.	Usluge.....	28
6.1.	Demografija	29
6.2.	Marketing	30
6.3.	Zdravstvo	31
6.4.	Turizam.....	32
6.5.	Obrazovanje	33
6.6.	Infrastruktura	34
6.7.	Sigurnost.....	35
6.8.	Financije.....	35
7.	Primjena anonimiziranih transformacija lokacijskih podataka u uslugama različite namjene	37
7.1.	Povezanost usluga s određenom vrstom transformacije lokacijskih podataka.....	37
8.	Izrada profila korisnika na temelju lokacijskih podataka	49
8.1.	Izrada korisničkih profila na proširenom skupu korisničkih podataka.....	50
	Zaključak	52
	Literatura.....	54
	Sažetak	56
	Summary	57

Uvod

Mobilne mreže generiraju ogromne količine podataka o korisnicima, ali ti podaci se često koriste u sirovoj formi te se rijetko u potpunosti iskorištavaju. Međutim, temeljita analiza ovih podataka može otkriti vrijedne uvide u ponašanje i navike korisnika, posebno kada se *Call Detail Record* (CDR) zapisi prošire dodatnim informacijama kao što su podaci o pretplati i ostali relevantni atributi. Korištenjem ovih proširenih podataka možemo značajno unaprijediti razumijevanje korisničkih potreba i optimizirati usluge. Grupiranje korisnika prema njihovim navikama u kretanju i komunikaciji omogućuje stvaranje personaliziranih ponuda i usluga, dok se istovremeno postiže anonimizacija podataka. Anonimizirani podaci štite privatnost korisnika, ali i dalje pružaju vrijedne uvide koji se mogu koristiti u različitim industrijama, kao što su marketing, zdravstvo, turizam, obrazovanje, infrastruktura, sigurnost i financije. U marketingu, na primjer, detaljni podaci o korisnicima omogućuju precizno ciljanje specifičnih grupa s relevantnim ponudama. U zdravstvu, podaci o kretanju mogu pomoći u praćenju širenja bolesti i planiranju intervencija, dok u turizmu analiza kretanja može pružiti uvid u turističke tokove i optimizirati ponude za posjetitelje.

Lokacijski podaci korisnika spadaju u kategoriju osobnih podataka te ih je, kao takve, potrebno zaštititi kako bi se spriječile eventualne povrede privatnosti. Neki od poznatijih primjera povreda privatnosti na temelju lokacijskih podataka istaknuti su u prvom poglavlju. S obzirom na brz tehnološki razvoj te posljedično tome sve veću popularnost navedene teme, drugo poglavlje prezentira neka od postojećih istraživanja. Glavni cilj ovog istraživanja jest analiza kretanja korisnika putem zapisa CDR-a čiji je format detaljno opisan u trećem poglavlju. Četvrto poglavlje provodi analize rezultata temeljenih na zapisima CDR-a koje pružaju mogućnost razumijevanja ponašanja korisnika, oblikovanja njihovih kretanja te interakcija s okolinom. Također prikazuje i proces pripreme podataka za provođenje analiza kretanja, koji predstavlja nužni korak. Broj različitih analiza ovisi o veličini podataka dostupnih u zapisu CDR-a, odnosno što je veći broj atributa u zapisu to je veći broj informacija moguće saznati o korisniku. Osim toga, analiza ovakve vrste podataka omogućuje identifikaciju ključnih obrazaca ponašanja, kao što su učestalost putovanja između određenih lokacija, najčešće rute kretanja te promjene u

vremenskim aktivnostima koje predstavljaju vrijednu informaciju uslugama. Peto poglavlje sažeto prikazuje potencijalne probleme implementacije svake pojedine transformacije u sustave te također ističe i općenite probleme analize lokacijskih podataka pohranjenih u CDR podacima. S obzirom na to da je raspon usluga, u kojima bi rezultati transformacija opisanih u četvrtome poglavlju bile od ključne vrijednosti, širok u šestome poglavlju definirana je kategorizacija usluga te su navedene specificirane usluge svake od kategorija. Sedmo poglavlje sadrži glavni cilj rada, a odnosi se na povezivanje transformacija s određenim vrstama usluga za koje navedene transformacije imaju značaj jer za različite vrste usluga različiti podaci o korisnicima tih usluga imaju smisla. Korištenjem anonimiziranih transformacija usluge različitih vrsta „otvorile“ bi se prema svojim korisnicima. Važno je istaknuti da se radi o anonimiziranim podacima kako bi se osigurala privatnost korisnika. Time se ostvaruje dvostruka korist koja se očitava u tome da je privatnost korisnika zaštićena, a integracijom ovakvih informacija usluge imaju mogućnost unaprijediti iskustvo korisnika te na taj način pružati dodatnu vrijednost i poboljšati svoj rad. Pronalaženje ravnoteže između koristi koje pružaju ovi podaci i zaštite privatnosti ključno je za napredak tehnologije i društva u cjelini, posebno u sve digitalnijem okruženju. Osmo poglavlje fokusira se na kreiranje profila korisnika na temelju poznatih korisničkih podataka te podataka koji predstavljaju rezultate prethodno navedenih transformacija. Također, u prethodno navedenom poglavlju definirana je mogućnost proširivanja dostupnih podataka u zapisu CDR-a, odnosno prikazan je proces proširivanja te njegovi glavni koraci. Cilj je istaknuti mogućnost prilagodbe te povezivanja različitih vrsta zapisa CDR-a pojedinoj usluzi.

1. Primjeri povreda privatnosti lokacijskih podataka

U ovom poglavlju istaknuta je važnost podizanja svijesti o temi povreda privatnosti lokacijskih podataka korisnika s obzirom na to da većina aplikacija prikuplja određene informacije o korisnicima, a da toga korisnici nisu svjesni. Osim što često korisnici nisu informirani o tome koje podatke aplikacije o njima posjeduje, često takvi podaci omogućuju aplikacijama otkrivanje puno više informacija o njihovom ponašanju i preferencijama nego što bi to oni pretpostavili.

Prvi primjer je slučaj u kojem je aplikacija Strava 2018. godine uvela značajku toplinske karte kako bi anonimno prikupila podatke o aktivnostima korisnika na otvorenom, poput trčanja i vožnje bicikla. Ova značajka je osmišljena kako bi pomogla korisnicima identificirati zajedničke rute i lokacije za trening te omogućila povezivanje s drugim ljubiteljima sporta. Ubrzo je otkriveno da se ova toplinska karta može iskoristiti za otkrivanje informacija o tajnim američkim vojnim bazama diljem svijeta tako što je provedena studija na Državnom sveučilištu Sjeverne Karoline koja je prezentirala mogućnost anonimizacije podataka pohranjenih u toplinskoj karti. Istraživači su na temelju prikupljenih podataka analizama slika identificirali početne odnosno zaustavne točke te uspjeli predvidjeti početnu adresu korisnika s visokom točnošću. Rezultati ove studije, kao i sljedeće navedeni primjeri, ističu važnost podizanja svijesti o privatnosti podataka i mogućnosti zaštite istih. [1]

Sljedeći slučaj je primjer Google-a koji je poznat po tome da transparentno traži dozvolu za korištenje informacija o lokaciji, posebno u aplikacijama kao što je Google Maps, kako bi pružio korisnicima preciznu navigaciju. S obzirom na to da postoji rizik u dugotrajnom pohranjivanju korisnikovog kretanja, uveden je alat koji omogućuje korisnicima "pauziranje" tog snimanja putem Google Maps aplikacije pod postavkama za povijest lokacija. Prema provedenim istraživanjima te unatoč tvrdnjama Google-a da pauziranje sprječava pohranu lokacija, dokazano je drugačije. Naime, rezultati pokazuju da u slučajevima kada je povijest lokacija pauzirana, određene aplikacije automatski pohranjuju određene podatke o lokacijama, kao što su primjerice gdje i kada je korisnik otvorio Google Maps. Osim toga, određena korisnikova pretraživanja koja nemaju veze s lokacijom, također

mogu pohraniti točne koordinate lokacije korisnika. Ova saznanja prikazuju povrede privatnosti korisnika te uvelike utječu na korištenje navedenih usluga. [2]

Provedeno je istraživanje koje se bavi analiziranjem pitanja privatnosti u aplikacijama za upoznavanje. Istraživači su analizirali kako aplikacije koje se temelje na lokaciji, prikupljaju i koriste podatke o lokaciji korisnika te su istražili potencijalne rizike i sigurnosne prijetnje koje proizlaze iz prikupljanja tih podataka. Glavni nedostaci provedenog istraživanja priloženi su:

- Prikupljanje i dijeljenje osjetljivih podataka o lokaciji korisnika, što može dovesti do otkrivanja lokacije trećim stranama.
- Nedostatak jasnih pravila privatnosti u aplikacijama, što otežava korisnicima razumijevanje i kontrolu nad svojim podacima.

Zaključno, istraživanje ističe važnost osvještavanja korisnika o sigurnosnim rizicima te poziva na razvoj i implementaciju boljih programskih rješenja uz osiguravanje zaštite privatnosti. [3]

Također postoji istraživanje koje je analiziralo Uber mobilnu aplikaciju. Navedeno istraživanje je ustanovilo da tvrtka koristi precizne tehnike praćenja geolokacija te je otkriveno kako mobilna aplikacija prati lokaciju korisnika čak i nakon završetka vožnje. Posebno interesantno je otkriće da aplikacija prati geolokacije i u situacijama kada korisnik ne koristi Uber uslugu što predstavlja veliki udar na privatnost korisnika te usluge. [4]

2. Postojeća rješenja

Ovo poglavlje donosi sažet prikaz rješenja koja već postoje, a čija tematika je slična tematici ovog istraživačkog rada.

Postoji istraživanje koje istražuje metodologije i tehnike koje omogućuju korisnicima anonimno korištenje usluga temeljenih na lokaciji, čime se štiti njihova privatnost. Autori istražuju načine koji korisnicima omogućuju manipulaciju prostornih i temporalnih podataka te prikrivanje njihove stvarne lokacije i vremena aktivnosti na različitim mjestima. Prostorno prikrivanje uključuje različite tehnike poput dodavanja lažnih i izmišljenih lokacija, promjene geografskog područja, korištenje proxy servera. S druge strane, temporalno prikrivanje odnosi se na sakrivanje vremena kada su korisnici bili aktivni na određenim lokacijama, što se može postići kroz odgode u prijenosu podataka ili korištenjem alata za planiranje aktivnosti. Cilj istraživanja je pružanje alata i tehnika korisnicima za korištenje usluga bez kompromitiranja njihove privatnosti. [5]

Jedno od postojećih rješenja na temu ovog istraživačkog rada nudi i radna skupina IETF Geopriv koja se bavi se pitanjima sigurnosti i privatnosti prijenosa lokacijskih informacija vanjskim uslugama kao i pohranjivanjem istih na lokacijskim poslužiteljima. Njihove glavne zadaće su dizajn protokola i API-a koji omogućuju razmjenu lokacijskih informacija poslužiteljima na povjerljivi način uz očuvanje integriteta. Poslužitelji imaju mogućnost smanjiti rezoluciju podataka kao i transformirati podatke u određene tipove formata i tako osigurati vanjskim uslugama pristup tim podacima. U radu se također istražuje problem privatnosti u kontekstu prikupljanja i objavljivanja podataka o lokaciji korisnika u zatvorenim prostorima, s obzirom na sve veću popularnost usluga temeljenih na lokaciji i društvenog umrežavanja. Predložen je novi okvir koji koristi pet tehnika anonimizacije (k-anonimnost, različitost, t-blizina, anonimnost i prisutnost) kako bi zaštitio privatnost korisnika prilikom objavljivanja podataka o lokaciji. Ovaj okvir omogućuje pružatelju usluge lokacije da korisnicima pruži online usluge, dok istovremeno objavljuje anonimizirane podatke koji čuvaju privatnost. Provedena je analiza praktične izvedivosti ovog okvira kroz simulacije i primjenu na stvarnim skupovima podataka, pri čemu su rezultati pokazali da objavljeni skupovi podataka o lokaciji uspješno štite identitet i informacije o lokaciji korisnika dok istovremeno pružaju korisne informacije

trećim stranama. Nakon što se podaci anonimiziraju, primjenjuju se različite tehnike objavljivanja kako bi se zaštitila privatnost korisnika, uključujući (α, k) -anonimnost, l -raznolikost, t -blizinu i δ -prisutnost. Ovaj proces uključuje generaliziranje vektora korisnika, određivanje semantičkih informacija, pronalaženje korisnika u blizini, primjenu dodatnih tehnika anonimizacije te finaliziranje zaštićenog područja korisnika prije objavljivanja. [6]

Ideja HERE platforme slična je ideji ovog rada jer pruža širok spektar usluga temeljenih na lokaciji, uključujući precizno određivanje položaja, usluge navigacije, kartografiju, informacije o prometu i prilagođavanje usluga korisnikovoj lokaciji. Ključni cilj platforme je omogućiti korisnicima da iskoriste prednosti tehnologije lokacija u različite svrhe, kao što su pronalaženje najbržeg puta do odredišta, pronalaženje obližnjih usluga i atrakcija te praćenje prometa u stvarnom vremenu. Kako bi zaštitila privatnost korisnika, HERE platforma koristi niz tehnika anonimizacije, uključujući k -anonimnost, raznolikost, blizinu, anonimnost i prisutnost. Ove tehnike osiguravaju da podaci o lokaciji korisnika mogu biti dijeljeni s trećim stranama bez otkrivanja njihovog identiteta ili točne lokacije, čime se podržava sigurno i odgovorno korištenje podataka o lokaciji u različitim kontekstima i aplikacijama. [7]

Također postoji istraživanje koje se bavi problemom objavljivanja skupnih podataka u uslugama temeljenih na lokaciji s ciljem očuvanja privatnosti korisnika. Kako tradicionalne tehnike zaštite privatnosti nisu uvijek primjenjive ili bi mogle rezultirati gubitkom informacija, istraživački članak "Towards Privacy Preserving Publishing of Set-Valued Data on Location-based Services" predstavlja novi pristup koji omogućuje objavljivanje podataka uz očuvanje privatnosti u hibridnom oblaku. U fazi objavljivanja podataka koristi se tehnika particioniranja podataka nazvana prošireno kvazi-identifikatorsko particioniranje, koja razdvaja pojmove zapisa koji sudjeluju u identificiranju kombinacija te sprječava poslužitelja u oblaku da poveže zapis s rijetkim kombinacijama izraza. Slijedi faza upita za podatke u kojoj se koristi interaktivna diferencijalna strategija privatnosti kako bi se osiguralo da statistički upiti ne ugrožavaju privatnost korisnika. Kroz opsežne eksperimente na stvarnim skupovima podataka, dokazana je valjanost i praktičnost predložene sheme. [8]

Sljedeći rad bavi sličnom tematikom, a on istražuje različite tehnike anonimizacije podataka koje se koriste za očuvanje privatnosti korisnika u uslugama temeljenim na lokaciji. Autori su članku u "Preserving User Privacy in LBS: A Survey on Anonymization Techniques" analizirali nekoliko tehnika, uključujući k-anonimnost, generalizaciju, supresiju, agregaciju te diferencijalnu privatnost. Zaključak istraživanja ističe da nema univerzalnog pristupa koji bi odgovarao svim scenarijima te da je odabir tehnike anonimizacije ovisan o specifičnim zahtjevima aplikacije i razini privatnosti koja se traži. Nedostaci nekih tehnika uključuju gubitak korisnih informacija ili smanjenje korisne vrijednosti podataka, dok su prednosti uključivale zaštitu privatnosti korisnika i pružanje anonimnih podataka za daljnju analizu. [9]

U ovom poglavlju je važno istaknuti i OpenDP aplikaciju koja objedinjuje alate i resurse za anonimizaciju podataka u različitim kontekstima, među kojima su i lokacijski podaci. Ovaj projekt pruža podršku u razvoju otvorenih i transparentnih tehnologija kako bi se omogućilo sigurno dijeljenje podataka dok se istovremeno štiti privatnost korisnika. [10]

Još jedan rad bavi se temom zaštite privatnosti lokacijskih podataka te pruža rješenja za zaštitu privatnosti korisnika. U radu "Privacy-preserving location-based services for smart cities: Taxonomy and research challenges" istraživači su analizirali različite tehnike i metode za očuvanje privatnosti korisnika u pametnim gradovima, poput enkripcije, diferencijalne privatnosti i upravljanja pristankom. Zaključak članka sličan je kao i u navedenim rješenjima, a to je da ističe važnost pronalaženja ravnoteže između funkcionalnosti usluga temeljenih na lokaciji i zaštite privatnosti kao i pozivanje na daljnja istraživanja i razvoj novih tehnika i pristupa koji će omogućiti sigurno korištenje lokacijskih podataka u pametnim gradovima. [11]

3. Zapisi pojedinosti telefonskog poziva

U ovome je poglavlju definiran osnovni ulazni skup podataka na kojem se temelji ovaj rad.

U telekomunikacijskoj mreži, *Call Data Records* (CDR) odnosi se na zapise detalja telefonskog poziva. Ovaj skup podataka pruža širok raspon informacija o uspostavljenim pozivima. Distribucija zapisa CDR-a općenito je automatski proces koji se odvija unutar sustava telekomunikacijskih operatera ili pružatelja usluga. Ovi podaci od velike su važnosti za telekomunikacijske operatere iz razloga što se na temelju ovakve vrste podataka provodi naplata usluga, praćenje korištenja usluga, otkrivanje prijevara, mrežna integracija i planiranje, nadzor mreže, analiza prometa i dr.

Također, na temelju zapisa CDR-a moguće je provesti različite analize kao što su analiza duljine poziva, analiza prosječne potrošnje korisnika, analiza prometa po vremenu, analiza lokacije na kojoj je uspostavljen i završen poziv, analiza kvalitete usluga, analiza prevencije prevara i dr. Rezultati navedenih analiza omogućuju otkrivanje osobnih podataka korisnika, stoga je važno da je sam proces distribucije, kao i pohrana zapisa CDR-a zaštićena adekvatnim sigurnosnim mehanizmima. [12]

3.1. Format CDR zapisa

U ovome je potpoglavlju prikazan općeniti format zapisa CDR-a.

Format CDR-a ovisi o krajnjoj usluzi koja iste koristi, a općenito zapisi na kojima će biti provedena analiza bit će formata prikazanog u tablici u nastavku.

Tablica 3.1.: Prikaz formata zapisa CDR-a

PermutedMSISDN	Jedinstveni broj korisnika u mobilnoj mreži
StartTime	Vrijeme početka poziva
ID	Jedinstveni identifikator zapisa
FileSegmentID	Jedinstveni identifikator segmenta unutar kojeg su pohranjeni svi potrebni podaci o mrežnoj komunikaciji

RecordTypeID	Identifikator tipa zapisa (1=MOC, 2 = MTC, 3 = CF, 8 = SMS MO, 9 = SMS MT)
IMSI	Jedinstveni međunarodni identitet mobilnog korisnika koji je pohranjen u SIM kartici
IMEISv	IMEI Software version - Broj koji se sastoji od 16 znamenki, od kojih prvih 14 sadrže informacije o uređaju, preostale 2 sadrže informacije o verziji softvera koji mobilni uređaj koristi
PermutedCallingNumber	Jedinstveni broj koji započinje usmjeravanje poziva
PermutedCalledNumber	Jedinstveni broj koji je pozvan od strane usmjeritelja poziva
PermutedOriginalCalledNumber	Jedinstveni originalni broj koji je pozvan od strane usmjeritelja poziva
CauseForTermination	Identifikator uzroka prestanka poziva
CauseForForwarding	Identifikator uzroka preusmjeravanja
SeizureTime	Vrijeme početka poziva
StopTime	Vrijeme završetka poziva
Duration	Vrijeme trajanja poziva
FirstMCC	Jedinstvena oznaka države na početku uspostavljanja poziva
FirstMNC	Jedinstvena oznaka mobilne mreže na početku uspostavljanja poziva
FirstLAC	Jedinstveni broj lokacije određenog područje na početku uspostavljanja poziva

FirstCID	Jedinstveni broj ćelije bazne stanice primopredajnika na početku uspostavljanja poziva
LastMCC	Jedinstvena oznaka države na završetku poziva
LastMNC	Jedinstvena oznaka mobilne mreže na završetku poziva
LastLAC	Jedinstveni broj lokacije određenog područja na završetku poziva
LastCID	Jedinstveni broj ćelije bazne primopredajnika na završetku poziva

Prethodno navedeni atributi omogućuju provedbu složenijih analiza koje su opisane u sljedećem poglavlju. Iako je u ovome radu naglasak na analizi kretanja korisnika te primjeni rezultata te analize u različitim uslugama, cilj ovog rada je također prezentirati koje su sve informacije o korisniku poznate na temelju njegove lokacije što je navedeno u osmom poglavlju.

4. Analiza kretanja korisnika

Ovo poglavlje prezentira složene analize kretanja korisnika koje se zasnivaju na njihovim lokacijama. Također u ovome je poglavlju ukratko prikazana ideja pripreme podataka iz početnog ulaznog skupa za analizu kretanja korisnika.

4.1. Priprema podataka za analizu kretanja korisnika

Čišćenje podataka, odnosno priprema istih nužan je korak u ovoj analizi zbog unaprjeđivanja performansi to jest optimalnosti izvršavanja pojedinih analiza.

Početni skup podataka obrađen je tako da su uklonjeni „bezvrijedni“ podaci, što je i očekivano s obzirom na to da su podaci stvarni. Podaci koji se smatraju „bezvrijednima“ odnose se na slučajeve u kojima je početna i završna lokacija ista te one u kojima je jedna od dviju lokacija jednaka *null* vrijednosti.

4.2. Definiranje lokacije korisnika

Lokaciju je moguće definirati na temelju zapisa CDR-a tako što kombinacija atributa FirstLAC i FirstCID čine početnu, a kombinacija LastLAC i LastCID završnu lokaciju. Lokacija korisnika također spada u osobne podatke tako da bilo kakvo korištenje ovih podataka bez prethodne anonimizacije istih označava kršenje GDPR-a.

Analiza kretanja označava transformaciju podataka, odnosno prilagođavanje podataka s obzirom na vrstu usluge za koju su ti podaci predviđeni.

U nastavku će biti definirane moguće transformacije podataka na temelju zapisa CDR-a.

4.3. Posjećenost lokacije u vremenskom rasponu

Navedena transformacija označava proces grupiranja posjećenih lokacija na određenom geografskom području kroz određeni vremenski raspon. Vremenski raspon može se definirati kao vremenski raspon u danu koji je podijeljen na sate u danu, kao vremenski raspon u mjesecu koji je podijeljen na dane u mjesecu, kao vremenski raspon u godini koji je podijeljen na mjesece u godini i dr.

Navedena analiza omogućuje davanje odgovora na sljedeća pitanja:

- Koja je najposjećenija lokacije u razdoblju od x do y?

- Koja je najmanje posjećena lokacije u razdoblju od x do y?
- Koji je najposjećenija lokacija u x razdoblju?

Odgovori na ova pitanja omogućuju pružateljima usluge poboljšanje i proširenje njihove usluge ovisno o vrsti iste.

U nastavku je priložen pseudokod koji izvršava ovu transformaciju.

FunkcijaIzračunajPostotakPosjećenostiLokacija(zapisiCDR, gornjaGranicaVremenskogRaspona, donjaGranicaVremenskogRaspona, listaLokacija):

```
//Inicijalizacija prazne strukture za praćenje posjećenosti svake
lokacije
posjećenostLokacija = {}

// Postavljanje broja posjeta na svaku lokaciju na početku
Za svaku lokaciju u listiLokacija:
    Postavi brojPosjetaNaLokaciju za lokaciju na 0

// Iteriranje kroz zapise CDR-a i brojanje posjeta za svaku
lokaciju
Za svaki zapis u zapisiCDR:
    Ako vrijeme zapisa nije unutar definiranog vremenskog
    raspona, preskoči zapis i prijeđi na sljedeći
    Ako vrijeme zapisa je unutar definiranog vremenskog raspona
        Odredi lokaciju zapisanu u CDR-u
        Ako je lokacija prisutna u listiLokacija:
            Povećaj brojPosjetaNaLokaciju za tu lokaciju za 1

// Računanje postotka posjećenosti za svaku lokaciju
Za svaku lokaciju u listiLokacija:
    Izračunaj postotak posjećenosti lokacije u odnosu na ukupan
    broj posjeta u tom vremenskom rasponu
    Spremi postotak posjećenosti lokacije u posjećenostLokacija
Vrati posjećenostLokacija
```

Kod 4.1. Prikaz pseudokoda za transformaciju posjećenosti lokacije u određenom vremenskom rasponu

4.4. Posjećenost korisnika na određenoj lokaciji

Ova vrsta transformacije provodi detaljnu analizu posjećenosti korisnika na određenoj lokaciji, što omogućuje dublje razumijevanje obrasca kretanja i aktivnosti korisnika na toj lokaciji. Kroz ovu analizu, istražuju se različiti aspekti posjećenosti, uključujući ukupno vrijeme provedeno na lokaciji, broj različitih korisnika koji su je posjetili, učestalost posjeta tijekom određenog vremenskog razdoblja te varijacije u posjećenosti ovisno o danu ili vremenu dana. Navedena analiza pruža odgovor na sljedeća pitanja:

- Koji je postotak posjećenosti određene lokacije?
- Koje je prosječno vrijeme zadržavanja korisnika na određenoj lokaciji?
- Koji je broj različitih korisnika koji su posjetili određenu lokaciju?
- Koji je postotak posjećenosti određene lokacije u određenom vremenskom rasponu u danu?
- Koji je postotak posjećenosti korisnika kojima je lokacija stanovanja blizu određene lokacije?
- Koji je postotak posjećenosti korisnika kojima je lokacija posla blizu određene lokacije?

U nastavku je priložen pseudokod koji predstavlja jednu od mogućih inačica ove transformacije:

```
Funkcija DetaljnaAnalizaPosjećenostiNaLokaciji(zapisiCDR, lokacija):  
  
    Inicijalizacija praznih struktura za praćenje posjećenosti:  
    ukupnoVrijemeNaLokaciji = 0  
    brojRazličitihKorisnika = 0  
    učestalostPosjeta = {}  
    posjećenostPoDanuIVremenu = {}  
  
    // Iteriranje kroz zapise CDR-a i analiza posjećenosti korisnika na  
    lokaciji  
  
    Za svaki zapis u zapisiCDR:  
  
        Ako je zapis za odabranu lokaciju:  
  
            Povećaj ukupnoVrijemeNaLokaciji za vrijeme provedeno na  
            lokaciji u zapisu
```

Ako korisnik nije prethodno zabilježen:

Povećaj brojRazličitihKorisnika za 1

Ažuriraj učestalostPosjeta za odabrano vremensko razdoblje

Ažuriraj posjećenostPoDanuIVremenu na osnovu dana i vremena posjeta

Vrati ukupnoVrijemeNaLokaciji, brojRazličitihKorisnika, učestalostPosjeta, posjećenostPoDanuIVremenu

Kod 4.2. Prikaz pseudokoda za transformaciju posjećenosti lokacije na određenoj lokaciji

4.5. Posjećenost određene lokacije s obzirom na dobnu skupinu korisnika

Posjećenost određene lokacije s obzirom na dobnu skupinu korisnika je vrsta analize koja se usredotočuje na istraživanje posjećenosti određene lokacije s obzirom na dobnu skupinu korisnika. Cilj je razumjeti kako se posjeti lokaciji razlikuju među različitim dobnim skupinama, pružajući dublji uvid u preferencije i obrasce ponašanja istih.

Analiza posjećenosti lokacije prema dobnim skupinama omogućuje identifikaciju koje dobne skupine češće posjećuju određene lokacije, kao i koje aktivnosti preferiraju dok su tamo. Na primjer, moglo bi se otkriti da mlađi ljudi češće posjećuju određeni park radi druženja, dok starije osobe više vremena provode u obližnjim trgovinama.

Osim toga, analiza može pružiti uvide u trajanje posjeta i učestalost ponovnih posjeta lokaciji među različitim dobima. Možda će biti moguće primijetiti da mladi ljudi često provode kraće, ali češće posjete određenoj lokaciji, dok starije osobe preferiraju dulje posjete ali rjeđe dolaske.

Prethodno navedena analiza pruža odgovore na sljedeća moguća pitanja:

- Koje su najposjećenije lokacije tijekom radnog vremena, a koje izvan radnog vremena za različite dobne skupine?
- Postoje li razlike u učestalosti posjeta određenim lokacijama među različitim dobima?

- Koje su aktivnosti popularne kod određenih dobrih skupina na određenim lokacijama (npr. parkovi, trgovački centri, kafići, restorani)?
- Kako se razlikuje posjećenost određenih lokacija između radnih dana i vikenda za različite dobne skupine?
- Koje su lokacije najčešće posjećene tijekom radnog vremena mladih ljudi u usporedbi s umirovljenicima?

Sljedeći pseudokod izvršava navedenu transformaciju.

Funkcija AnalizaPosjećenostiPoDobnojSkupini(zapisiCDR, lokacija):

```

    Inicijalizacija praznih struktura za praćenje posjećenosti po
    dobnim skupinama:

    posjećenostPoDobuDana = {}

    trajanjePosjetaPoDobnojSkupini = {}

    učestalostPonovnihPosjetaPoDobnojSkupini = {}

    // Iteriranje kroz zapise CDR-a i analiza posjećenosti lokacije po
    dobnim skupinama

    Za svaki zapis u zapisiCDR:

        Odredi dobnu skupinu korisnika na osnovu dostupnih podataka

        Ako dobna skupina nije prethodno zabilježena, inicijaliziraj
        je u strukturi

        Ažuriraj posjećenostPoDobuDana za odabranu dobnu skupinu i
        dan posjeta

        Ažuriraj trajanjePosjetaPoDobnojSkupini za odabranu dobnu
        skupinu

        Ažuriraj učestalostPonovnihPosjetaPoDobnojSkupini za odabranu
        dobnu skupinu

    VratiposjećenostPoDobuDana, trajanjePosjetaPoDobnojSkupini,
    učestalostPonovnihPosjetaPoDobnojSkupini

```

Kod 4.3. Prikaz pseudokoda za transformaciju posjećenosti lokacije na određenoj lokaciji s obzirom na dobnu skupinu

4.6. Posjećenost lokacija s obzirom na lokaciju stanovanja

Ova analiza istražuje posjećenost različitih lokacija u odnosu na lokaciju stanovanja korisnika. Cilj je ustanoviti kako se preferencije korisnika razlikuju ovisno o njihovom mjestu prebivališta te kako udaljenost lokacije stanovanja od određene lokacije utječe na učestalost posjeta.

Prateći posjete korisnika na različitim lokacijama u odnosu na lokaciju njihovog stanovanja, moguće je identificirati koje su lokacije popularne među lokalnom populacijom, a koje privlače posjetitelje iz drugih dijelova grada ili čak izvan grada.

Ova vrsta analize pruža odgovore na sljedeća moguća pitanja:

- Koje su najčešće posjećene lokacije korisnika koji stanuju u x lokaciji?
- Koje su glavne destinacije koje stanovnici određene lokacije najčešće posjećuju?
- Koliko često stanovnici određene lokacije posjećuju određene lokacije?
- Koje su aktivnosti preferirane za vikend među stanovnicima određene lokacije?

Pseudokod za navedenu transformaciju je sljedeći:

Funkcija AnalizaPosjećenostiPoLokacijiStanovanja(zapisiCDR, lokacije):

```
Inicijalizacija praznih struktura za praćenje posjećenosti s
obzirom na lokaciju stanovanja
učestalostPosjetaPoLokacijiSObziromNaLokacijuStanovanja = {}
// Iteriranje kroz zapise CDR-a i analiza posjećenosti lokacija u
odnosu na lokaciju stanovanja
Za svaki zapis u zapisiCDR:
    Ako je vrijeme posjeta između 19 sati navečer i 7 sati
    ujutro:
        Lokacija se bilježi kao lokacija stanovanja
        Ako lokacija stanovanja nije prethodno zabilježena,
        inicijaliziraj je u strukturi
    Ako vrijeme posjeta nije između 19 sati navečer i 7 sati
    ujutro:
```

```
Ažuriraj
učestalostPosjetaPoLokacijiSObziromNaLokacijuStanovanja
za posjećenu lokaciju
```

```
Vrati učestalostPosjetaPoLokacijiSObziromNaLokacijuStanovanja
```

Kod 4.4. Prikaz pseudokoda za transformaciju posjećenosti lokacije na određenoj lokaciji s obzirom na lokaciju stanovanja

4.7. Posjećenost lokacija s obzirom na lokaciju posla

Analiza posjećenosti lokacija u odnosu na lokaciju posla pruža uvid u obrasce kretanja i aktivnosti korisnika temeljenih na njihovom radnom mjestu. U ovom dijelu istraživanja istražuju se kako lokacija posla utječe na preferencije korisnika u odabiru lokacija za rekreaciju, kupovinu, prehranu te druge svakodnevne aktivnosti. Ova analiza omogućuje otkrivanje povezanosti između radnog okruženja i životnog stila korisnika, otvarajući prostor za prilagodbu usluga i sadržaja kako bi se bolje zadovoljile njihove potrebe i preferencije.

Navedena vrsta transformacije daje odgovor na sljedeća moguća pitanja:

- Koje su najčešće posjećene lokacije korisnika koji rade na x lokaciji?
- Koje su najčešće posjećene lokacije koje korisnici posjećuju nakon radnog vremena?
- Koje su najpopularnije lokacije za vrijeme pauze među korisnicima koji rade na određenim radnim mjestima?
- Postoje li specifične aktivnosti koje korisnici preferiraju nakon radnog vremena, ovisno o njihovoj lokaciji posla?
- Koje su najčešće destinacije za vikend izlete među korisnicima koji rade na određenim lokacijama?

U nastavku je priložen pseudokod koji izvršava navedenu transformaciju:

```
Funkcija AnalizaPosjećenostiPoLokacijiPosla(zapisiCDR, lokacije):
    Inicijalizacija praznih struktura za praćenje posjećenosti s
    obzirom na lokaciju posla
    učestalostPosjetaPoLokacijiSObziromNaLokacijuPosla = {}
    // Iteriranje kroz zapise CDR-a i analiza posjećenosti lokacija u
    odnosu na lokaciju posla
    Za svaki zapis u zapisiCDR:
```


Ako je vrijeme posjeta između 8 sati ujutro i 18 sati navečer:

```
Lokacija se bilježi kao lokacija posla
```

```
Ako lokacija posla nije prethodno zabilježena,  
inicijaliziraj je u strukturi
```

Ako vrijeme posjeta nije između 8 sati ujutro i 18 sati navečer:

```
Ažuriraj
```

```
učestalostPosjetaPoLokacijiSObziromNaLokacijuPosla za  
posjećenu lokaciju
```

```
Vrati učestalostPosjetaPoLokacijiSObziromNaLokacijuPosla
```

Kod 4.5. Prikaz pseudokoda za transformaciju posjećenosti lokacije na određenoj lokaciji s obzirom na lokaciju posla

4.8. Posjećenost lokacija s obzirom na spol

Transformacija posjećenih lokacija s obzirom na spol korisnika predstavlja važan segment analize koji istražuje kako se preferencije i obrasci kretanja razlikuju između muškaraca i žena. Ovaj aspekt istraživanja omogućuje otkrivanje razlika u aktivnostima i interesima muškaraca i žena te kako njihov spol može utjecati na izbor lokacija za razne svakodnevne aktivnosti. Analizom posjećenosti lokacija s obzirom na spol, istražuju se razlike u preferencijama za trgovine, restorane, kulturne institucije te druge vrste lokacija, otvarajući prostor za prilagodbu usluga i sadržaja kako bi se bolje zadovoljile potrebe i interesi muškaraca i žena.

Ova vrsta analize pruža odgovor na sljedeća moguća pitanja:

- Koje su lokacije kojima se žene najmanje kreću?
- Koje su lokacije kojima se muškarci najmanje kreću?
- Koje su lokacije kojima se žene najviše kreću?
- Koje su lokacije kojima se muškarci najviše kreću?

Funkcija AnalizaPosjećenihLokacijaPoSpolu(zapisiCDR, lokacije, spolKorisnika):

```
Inicijalizacija praznih struktura za praćenje posjećenosti po  
lokacijama i spolu korisnika:
```

```
posjećenostPoLokacijiPoSpolu = {}
```

```
// Iteriranje kroz zapise CDR-a i analiza posjećenosti lokacija po  
spolu korisnika
```

```
Za svaki zapis u zapisiCDR:
```

```
    Odredi spol korisnika na osnovu dostupnih podataka
```

```
    Ako spol nije prethodno zabilježen, inicijaliziraj ga u strukturi
```

```
    Ažuriraj posjećenostPoLokacijiPoSpolu za odabranu lokaciju i spol  
    korisnika
```

```
Vrati posjećenostPoLokacijiPoSpolu
```

Kod 4.6. Prikaz pseudokoda za transformaciju posjećenosti lokacije na određenoj lokaciji s obzirom na spol

4.9. Posjećenost lokacija s obzirom na broj upućenih poziva

Ova vrsta analize proučava vezu između intenziteta komunikacije putem telefonskih poziva i posjećenosti različitih lokacija. Analiza se temelji na pretpostavci da je broj poziva koji korisnici obavljaju s neke lokacije povezan s njihovim stvarnim prisustvom na toj lokaciji. Dakle, visoki broj poziva s određene lokacije sugerira čestu posjećenost, dok nizak broj poziva može ukazivati na manju frekvenciju posjeta.

Ova analiza omogućuje dublje razumijevanje navika i obrasca kretanja korisnika. Primjerice, moguće je identificirati popularne lokacije na temelju broja poziva koji se obavlja s tih lokacija. Također, moguće je istražiti kako se ti uzorci razlikuju između različitih vremenskih perioda, dana u tjednu ili čak sezona.

U konačnici, analiza posjećenosti lokacija s obzirom na broj upućenih poziva pruža uvide u povezanost između komunikacijskih uzoraka i stvarnih aktivnosti korisnika na različitim lokacijama. To omogućuje poslovnim subjektima da bolje razumiju svoju ciljanu publiku i prilagode svoje usluge kako bi bolje zadovoljili njihove potrebe i preferencije.

Ova vrsta analize pruža odgovor na sljedeća moguća pitanja:

- Koja je najpopularnija popularna lokacija među ljudima koji su primili određeni broj poziva?

- Koja je najmanje popularna lokacija među ljudima koji su primili određeni broj poziva?
- Koja je najpopularnija popularna lokacija među ljudima koji su primili određeni broj poziva u određenom vremenskom intervalu?
- Koja je najmanje popularna lokacija među ljudima koji su primili određeni broj poziva u određenom vremenskom intervalu?

4.10. Posjećenost lokacija s obzirom na broj primljenih poziva

Analiza povezanosti posjećenosti lokacija s brojem primljenih poziva istražuje kako komunikacijski obrasci putem telefonskih poziva odražavaju frekvenciju posjeta različitim lokacijama. Pretpostavka je da veći broj primljenih poziva s neke lokacije implicira češće prisustvo korisnika na toj lokaciji, dok manji broj poziva sugerira manju frekvenciju posjeta.

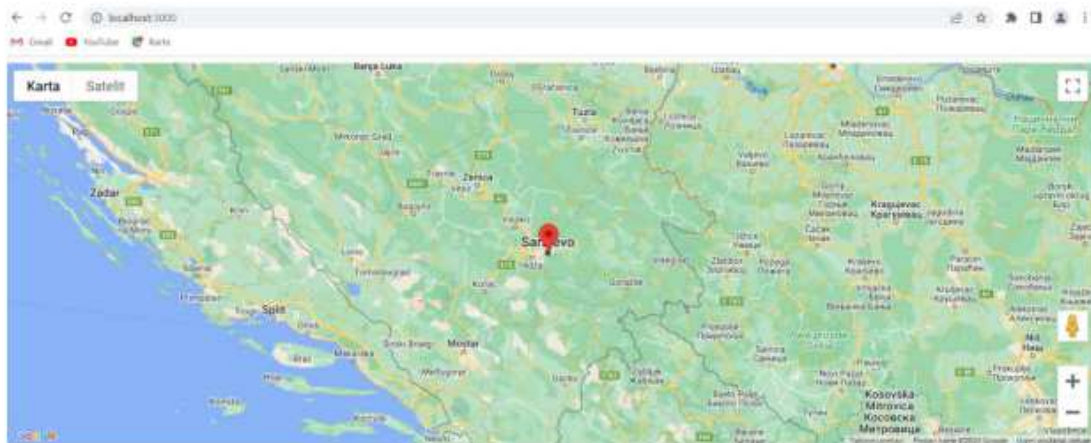
Ovaj uvid omogućuje bolje razumijevanje kretanja korisnika i njihovih navika. Na primjer, moguće je identificirati popularne destinacije na temelju broja primljenih poziva s tih lokacija, te istražiti kako se ovi obrasci mijenjaju tijekom različitih vremenskih perioda ili dana u tjednu.

Ova analiza može odgovoriti na pitanja poput:

- Koja je najpopularnija lokacija među korisnicima koji su primili određeni broj poziva?
- Koja je najmanje popularna lokacija među korisnicima s određenim brojem primljenih poziva?
- Kako se mijenja popularnost lokacija tijekom različitih vremenskih intervala ili sezona?
- Koji su trendovi posjećenosti lokacija u odnosu na intenzitet primljenih poziva?

4.11. Brzina kretanja

Ova vrsta analize utvrđuje brzinu kretanja korisnika koja predstavlja vrlo važnu informaciju za određene vrste usluga. Sama zamisao ove ideje je odrediti brzinu na temelju stvarne udaljenosti između početne lokacije, točke uspostavljanja poziva i završne lokacije, točke završavanja poziva te trajanja telefonskog poziva. Kako bi se odredila točna početna i završna lokacija korisnika potrebno je transformirati navedene lokacije u geografsku širinu i dužinu. S obzirom na veliki ulazni skup podataka korišten je Api, odnosno Geolocation Api. Proces je započeo generiranjem Api ključa, a transformacija i reprezentacija ostvarena je s pomoću http zahtjeva. U nastavku je prikazana transformacija lokacije iz zapisa CDR-a u geografsku širinu i dužinu.



Slika 4.11.: Transformacija atributa zapisa CDR u geografsku širinu i dužinu

U analizi je korištena Matrica udaljenosti Apija s pomoću koje je moguće odrediti udaljenost između dviju lokacija, odnosno reprezentirati udaljenost kao prevaljeni put koji će biti korišten u izračunu brzine kretanja korisnika. Kod za generiranje matrice udaljenosti prikazan je u kodu u nastavku.

```
import requests

origin = "43.105191,17.717586"

destination = "44.105191,18.717586"

url = "https://maps.googleapis.com/maps/api/distancematrix/json?origins={origin}&destinations={destination}&units=imperial&key=AIzaSyAaAEfulGvMPln-4nPVesBvbg_tXedfeds"
```

```

payload={}

headers = {}

response = requests.request("GET", url, headers=headers, data=payload)
print(response.text)

```

Kod 4.11.1.:Kod za generiranje matrice udaljenosti

Rezultat odnosno tijelo matrice udaljenosti prikazan je u nastavku.

```

{ "destination_addresses" : [ "Cesta bez naziva, Bosnia and Herzegovina"
],
"origin_addresses" : [ "4P49+38P, Tasovčići, Bosnia and Herzegovina" ],
"rows" : [ { "elements" : [ {
"distance" : { "text" : "132 mi", "value" : 212179 },
"duration" : { "text" : "3 hours 35 mins", "value" : 12896 },
"status" : "OK" } ] } ],
"status" : "OK" }

```

Kod 4.11.2.:Matrica udaljenosti

Brzinu kretanja je moguće izračunati s pomoću sljedeće formule:

$$v = \frac{\textit{udaljenost}}{\textit{trajanje telefonskog poziva}}$$

Formula 4.11.:Formula za izračun brzine kretanja

Ova vrsta transformacije u kombinaciji s prethodno navedenim transformacijama omogućuju davanje odgovora na pitanja kao što su:

- Koja je prosječna brzina kretanja korisnika čija je najčešće posjećena lokacija x?
- Koja je prosječna brzina kretanja osoba muškog spola?
- Koja je prosječna brzina kretanja osoba ženskog spola?
- Koja je prosječna brzina kretanja osoba mlađe životne dobi?
- Koja je prosječna brzina kretanja osoba starije životne dobi?
- Koja je prosječna brzina kretanja korisnika određene lokacije stanovanja?
- Koja je prosječna brzina kretanja korisnika određene lokacije posla?

5. Potencijalni problemi implementacije transformacija u raznovrsnim uslugama

Cilj ovog poglavlja je identificirati i opisati potencijalne probleme i izazove koji se mogu pojaviti prilikom implementacije transformacija temeljenih na zapisima CDR-a u raznovrsnim uslugama. U nastavku će za svaku transformaciju, opisanu u prethodnom poglavlju, biti priloženi potencijalni problemi odnosno nedostaci implementacija rezultata.

5.1. Posjećenost lokacije u vremenskom rasponu

Potencijalni problemi:

- **Vrijeme i datum:** Pogreške u vremenskom zapisu mogu dovesti do netočnih podataka o posjećenosti.
- **Učestalost ažuriranja:** Ako se podaci ne ažuriraju redovito, analize mogu biti zastarjele.
- **Ponašanje korisnika:** Različite aktivnosti korisnika u različitim vremenskim razdobljima (npr. radni dan vs. vikend) mogu otežati tumačenje podataka.

5.2. Posjećenost korisnika na određenoj lokaciji

Potencijalni problemi:

- **Preciznost lokacije:** Podaci koji se odnose na lokacije baznih stanica mogu biti netočni u gustim urbanim područjima ili unutar zgrada.
- **Mrežno pokrivanje:** Slaba pokrivenost mreže može rezultirati nedostatkom podataka za određena područja.
- **Razlikovanje posjetitelja:** Razlikovanje između stalnih posjetitelja i prolaznih korisnika može biti izazovno.

5.3. Posjećenost određene lokacije s obzirom na dobnu skupinu korisnika

Potencijalni problemi:

- **Točnost podataka o dobi:** Ako korisnički podaci nisu točni ili ažurirani, analiza može biti netočna.
- **Dobne preferencije:** Različite dobne skupine mogu imati različite navike korištenja mobitela, što može utjecati na rezultate.

5.4. Posjećenost lokacija s obzirom na lokaciju stanovanja

Potencijalni problemi:

- **Migracija:** Ljudi se mogu seliti, što može uzrokovati netočnosti u podacima o lokaciji stanovanja.
- **Preciznost adrese:** Ako adrese nisu precizno zabilježene, analiza može biti netočna, primjerice kod promjene naziva ulica.

5.5. Posjećenost lokacija s obzirom na lokaciju posla

Potencijalni problemi:

- **Promjene posla:** Česte promjene u zaposlenju korisnika mogu rezultirati netočnim podacima.
- **Različiti radni sati:** Fleksibilno radno vrijeme može otežati analizu standardnih uzoraka posjećenosti.
- **Geografska raspodjela:** U velikim urbanim sredinama, koncentracija radnih mjesta može dovesti do problema s točnošću podataka.

5.6. Posjećenost lokacija s obzirom na spol

Potencijalni problemi:

- **Točnost podataka o spolu:** Ako podaci nisu točni ili su nepotpuni, analize mogu biti netočne.
- **Preferencije:** Različite preferencije između spolova mogu utjecati na korištenje mobilnih uređaja i prikupljanje podataka.

5.7. Posjećenost lokacija s obzirom na broj upućenih poziva

Potencijalni problemi:

- **Ponašanje korisnika:** Korisnici koji često koriste aplikacije za razmjenu poruka umjesto telefonskih poziva mogu biti podzastupljeni.
- **Varijabilnost:** Broj poziva može varirati ovisno o sezoni (npr. blagdani), što može utjecati na rezultate.

5.8. Posjećenost lokacija s obzirom na broj primljenih poziva

Potencijalni problemi:

- **Filtriranje:** Teškoće u razlikovanju važnih od nevažnih poziva.
- **Ponašanje korisnika:** Navike korisnika vezane za primanje poziva mogu varirati, utječući na točnost analize.

5.9. Brzina kretanja

Potencijalni problemi:

- **Preciznost koordinata baznih stanica:** Netočnosti u podacima koji se odnose na lokacije baznih stanica mogu rezultirati netočnim procjenama brzine kretanja.
- **Prometne gužve:** Prometni uvjeti mogu utjecati na brzinu kretanja i otežati interpretaciju podataka.
- **Različiti načini prijevoza:** Razlike u brzinama između pješaka, biciklista i vozača mogu komplicirati analizu.

Implementacija ovih transformacija zahtijeva pažljivo razmatranje i rješavanje navedenih problema kako bi se osigurala točnost i pouzdanost rezultata analize.

5.10. Potencijalni problemi kod općenite analize CDR zapisa

U ovom poglavlju su istaknuti i općenito mogući problemi prilikom implementacije rezultata transformacija, a oni su navedeni u nastavku.

Tablica 5.1.: Prikaz mogućih problema prilikom općenite analize CDR podataka

Privatnost korisnika	Zapisi CDR-a sadrže osjetljive informacije o korisnicima, te je nužan korak anonimizacija istih, no i dalje postoji mogućnost deanonimizacije.
Točnost podataka	Točnost CDR podataka može varirati, te utjecati na preciznost transformacija. Primjerice, signali mobitela mogu biti slabiji u ruralnim područjima, što može rezultirati netočnim podacima o lokaciji. Također, prisutnost lažnih ili netočnih zapisa može negativno utjecati na rezultate analize.
Skalabilnost	Obrada velikih količina CDR podataka zahtijeva značajne računalne resurse i može biti izazovna s aspekta skalabilnosti.
Integracija s postojećim programskim rješenjima	Integracija novih transformacija s postojećim sustavima i uslugama može biti složena jer Postojeći sustavi možda neće biti dizajnirani za rad s vrstom podataka i analitika koje transformacije generiraju, što može zahtijevati značajne promjene u infrastrukturi.

Pravni izazovi	Korištenje CDR podataka regulirano je zakonskim propisima koji variraju ovisno o zemlji. Pridržavanje tih propisa ključno je kako bi se izbjegle pravne posljedice. Također, potrebno je osigurati transparentnost i pristanak korisnika za prikupljanje i analizu njihovih podataka.
Kvaliteta i ažuriranje podataka	Za točne analize, podaci moraju biti visoke kvalitete i redovito ažurirani. Ako podaci nisu ažurirani ili su nepotpuni, rezultati analize mogu biti netočni ili neadekvatni za donošenje odluka.

Implementacija transformacija koje se temelje na CDR podacima omogućava značajno unapređenje raznih usluga, ali zahtijeva pažljivo planiranje i upravljanje kako bi se prevladali navedeni izazovi i osigurali točni, relevantni i korisni rezultati.

6. Usluge

U ovome poglavlju su definirane kategorije usluga kojima bi korištenje rezultata analiza iz prethodnog poglavlja bilo značajno. Usluge su podijeljene u sljedeće kategorije:

- demografija
- marketing
- zdravstvo
- turizam
- obrazovanje
- infrastruktura
- sigurnost
- financije.

U svakoj od navedenih kategorija definirane su usluge koje pripadaju toj kategoriji te su prikazane u nastavku.

Tablica 6.1.: Kategorizacija usluga

DEMOGRAFIJA	<ul style="list-style-type: none">• migracije stanovništva• demografske promjene
MARKETING	<ul style="list-style-type: none">• društvene mreže• turističke usluge• usluge za pronalaženje posla• usluge naplate i bankarstva• usluge za istraživanje tržišta i analitika• transportne usluge
ZDRAVSTVO	<ul style="list-style-type: none">• usluge za vježbanje i zdravlje• usluge za upravljanje energijom i ekološke inicijative• usluge za praćenje epidemija u zdravstvu

TURIZAM	<ul style="list-style-type: none"> • usluge za planiranje parkinga • turističke usluge
OBRAZOVANJE	<ul style="list-style-type: none"> • usluge za edukaciju
INFRASTRUKTURA	<ul style="list-style-type: none"> • strateško planiranje i infrastruktura • (re)organizacija koncesija i licenci • transportne usluge • usluge za pronalaženje izgubljenih predmeta • usluge vremenske prognoze • usluge za planiranje parkinga
SIGURNOST	<ul style="list-style-type: none"> • usluge za hitne slučajeve • usluge za pronalaženje izgubljenih predmeta
FINANCIJE	<ul style="list-style-type: none"> • usluge naplate i bankarstva • usluge za istraživanje tržišta i analitika • usluge za pronalaženje posla • usluge osiguranja

6.1. Demografija

Za demografske procese, kao što su migracija stanovništva i demografske promjene, analiza anonimiziranih lokacijskih podataka pružala bi korisne uvide u razumijevanje kretanja korisnika te njihovih trendova bez narušavanja njihove privatnosti. Primjerice, analiza migracija stanovništva može uključivati identifikaciju i praćenje kretanja pojedinaca između različitih lokacija. Anonimizirani podaci iz zapisa CDR-a mogli bi pružiti informacije o lokacijama na kojima su korisnici bili prisutni, omogućujući identifikaciju uzoraka migracija bez otkrivanja identiteta pojedinaca. Korištenjem transformacija poput broja različitih korisnika koji su

posjetili određenu lokaciju ili prosječnog vremena zadržavanja na lokaciji, moguće bi bilo dobiti uvid u intenzitet migracija između različitih područja.

Slično tome, demografske promjene mogle bi se pratiti putem anonimiziranih transformacija podataka iz zapisa CDR-a. Na primjer, analiza posjećenosti određene lokacije u određenom vremenskom rasponu u danu može pružiti uvid u dnevne obrasce kretanja stanovništva unutar područja. Kombiniranjem tih podataka s demografskim informacijama o populaciji omogućila bi se identifikacija trendova i promjena u ponašanju stanovništva, poput sezonskih varijacija ili promjena u urbanizaciji.

Korištenje anonimiziranih transformacija podataka iz zapisa CDR-a, bilo bi moguće izvesti korisne analize demografskih trendova bez ugrožavanja privatnosti pojedinaca, pružajući tako vrijedne uvide u kretanje i evoluciju populacije.

6.2. Marketing

Rezultati analize anonimiziranih lokacijskih podataka imaju ključnu ulogu u suvremenom poslovanju te bi ova vrsta analize omogućavala dublje razumijevanje potreba i ponašanja korisnika u različitim sektorima.

U sektoru marketinga, društvene mreže su postale nezaobilazan alat za komunikaciju s korisnicima i promociju proizvoda ili usluga. Analiza podataka bi pružala detaljan uvid u ponašanje korisnika, njihove interese i preferencije. Na temelju ovih podataka, marketinški stručnjaci mogli bi kreirati personalizirane kampanje oglašavanja, ciljajući specifične demografske skupine i prilagođavajući svoje poruke kako bi bile relevantne i privlačne.

Kroz analizu anonimiziranih podataka iz zapisa CDR-a moguće je pratiti kretanje turista, identificirati popularne destinacije i sezonske trendove. Ovi podaci su od izuzetne važnosti za turistički sektor jer bi omogućavali prilagodbu marketinških strategija, poboljšanje ponude i usluga te bolje zadovoljavanje potreba turista.

Usluge dostave također bi koristile analizu podataka iz zapisa CDR-a kako bi optimizirale svoje procese dostave, smanjile troškove i poboljšale korisničko

iskustvo. Praćenje vremena dostave, rute dostave i učestalost isporuka omogućilo bi tvrtkama za dostavu pružanje brže i efikasnije usluge svojim korisnicima.

U sektoru zapošljavanja, analiza podataka pomogla bi u identifikaciji trendova na tržištu rada, pronalaženju potencijalnih kandidata i prilagodbi strategija regrutiranja. Poznavanje potražnje za određenim vještinama ili zanimanjima omogućilo bi tvrtkama da bolje planiraju svoje regrutne aktivnosti i privuku najkvalitetnije kandidate.

Naplata i bankarstvo također bi koristile analizu podataka kako bi otkrili prijevare, identificirali trendove u potrošnji i prilagodili ponudu financijskih proizvoda. Kroz analizu transakcija i ponašanja korisnika, banke bi mogle pružati personalizirane usluge i proizvode koji bi odgovarali potrebama i željama svojih klijenata.

Uslugama istraživanja tržišta i analitike, analiza podataka pružala bi duboke uvide u ponašanje potrošača, trendove u potražnji i konkurentsku analizu. Ovi uvidi omogućili bi tvrtkama da bolje razumiju svoje tržište, identificiraju konkurentske prednosti i kreiraju strategije rasta i razvoja.

Analiza anonimiziranih lokacijskih podataka u transportnom sektoru omogućilo bi praćenje prometa, optimizaciju rute i poboljšanje učinkovitosti prijevoza. Kroz analizu kretanja vozila i putnika, tvrtke za prijevoz mogle bi pružiti pouzdane usluge i optimizirati svoje operacije.

Korištenjem analize podataka poslovne strategije mogle bi se prilagoditi kako bi bolje odgovarale potrebama korisnika i stvorile konkurentnu prednost na tržištu. Ova analiza omogućila bi tvrtkama da budu agilnije i inovativnije u svom poslovanju, što je ključno za dugoročni uspjeh i održivost na tržištu.

6.3. Zdravstvo

Rezultati analiza lokacijskih podataka, kao što je prethodno navedeno, pružali bi uvide u ponašanje ljudi i dinamiku njihova kretanja te bi imali značajni utjecaj na različite aspekte zdravstvenih usluga. Zdravstvene institucije bi tako mogle poboljšati pružanje zdravstvene skrbi, kao i unaprijediti proces liječenja te brže i učinkovitije reagirati na javnozdravstvene izazove. Primjerice, analizom

anonimiziranih lokacijskih podataka pružala bi se mogućnost praćenja kretanja ljudi u stvarnome vremenu što je ključni aspekt za praćenje epidemija i širenja zaraznih bolesti. Također, ovakav pristup omogućavao bi zdravstvenim institucijama identifikaciju žarišta zaraze, praćenja kretanja potencijalno zaraženih osoba te poduzimanja mjera za sprečavanje daljnjeg širenja bolesti.

U uslugama za vježbanje i praćenje zdravlja anonimizirani lokacijski podaci činili bi ključnu ulogu u pružanju informacija o navikama kretanja korisnika te bi omogućavali prilagodbu primjerice wellness programa i fitness aktivnosti korisničkim preferencijama i potrebama.

Analiza podataka kretanja korisnika u području upravljanja energijom i ekoloških inicijativa također bi pružala uvide u obrasce kretanja ljudi i njihovu prisutnost. Na temelju tih podataka pružatelji usluga za upravljanje energijom mogli bi optimizirati distribuciju energije te poticati korištenje obnovljivih izvora energije.

Kroz korištenje anonimiziranih lokacijskih podataka, zdravstveni sektor mogao bi unaprijediti svoje usluge, bolje razumjeti potrebe korisnika te brže i učinkovitije reagirati na zdravstvene prijetnje i izazove, što bi rezultiralo poboljšanim zdravstvenim ishodima i boljim zadovoljstvom korisnika.

6.4. Turizam

Kroz detaljno praćenje kretanja turista i analizu njihovih preferencija na temelju rezultata anonimiziranih podataka, pružatelji usluga mogli bi prilagoditi svoje ponude i poboljšati kvalitetu usluga korištenjem.

Analiza lokacijskih podataka imala bi značajnu ulogu u uslugama za planiranje parkinga tako da bi omogućavala optimizaciju upravljanja parkiralištima i pružanja boljeg iskustva parkiranja. Također, analiza bi omogućavala identifikaciju slobodnih parkirnih mjesta kao i predviđanje gužvi te optimizaciju ruta do željeno parkirališta.

Turističke usluge koristile bi anonimizirane lokacijske podatke za personalizaciju ponuda i prilagodbu iskustava putnika. Kroz analizu podataka o kretanju turista, pružatelji usluga mogli bi prepoznati popularne turističke destinacije, identificirati

sezonske trendove te prilagoditi marketinške strategije kako bi privukli više posjetitelja.

Korištenjem anonimiziranih lokacijskih podataka, turistički sektor mogao bi unaprijediti svoje usluge, poboljšati iskustva putnika te prilagoditi ponude prema potrebama i preferencijama korisnika. Ovo bi zasigurno rezultiralo boljim turističkim iskustvima, većim zadovoljstvom putnika te poticanjem razvoja turizma.

6.5. Obrazovanje

Korištenje transformacija lokacijskih podataka imao bi potencijal u unaprjeđivanju različitih aspekata obrazovanja, posebno kroz pružanje boljih usluga za edukaciju. Kroz praćenje kretanja učenika/studenata i analizu njihovih aktivnosti, pružatelji usluga za edukaciju mogli bi poboljšati procese učenja, prilagoditi ponude te pružiti personalizirano iskustvo.

Usluge za edukaciju koristile bi anonimizirane lokacijske podatke kako bi razumjele obrasce kretanja učenika/studenata i njihovu prisutnost na različitim lokacijama, poput kampusa, knjižnica ili učionica. Kroz analizu ovih podataka, pružatelji usluga mogli bi identificirati trendove u ponašanju studenata, prepoznati područja koja zahtijevaju dodatnu podršku te prilagoditi raspored predavanja i aktivnosti kako bi maksimizirali učinkovitost učenja.

Osim toga, anonimizirani lokacijski podaci mogli bi se koristiti za optimizaciju logistike i upravljanje resursima u obrazovnim institucijama. Kroz analizu podataka o kretanju učenika/studenata, pružatelji usluga mogli bi bolje planirati raspored predavanja, osigurati optimalno korištenje prostora te pružiti bolje usluge podrške, poput usluga savjetovanja.

Korištenjem anonimiziranih lokacijskih podataka, obrazovni sektor mogao bi unaprijediti svoje usluge, prilagoditi ponude te osigurati bolje iskustvo učenja za sve sudionike obrazovnog procesa. Ovo bi rezultiralo poboljšanim akademskim rezultatima, većim zadovoljstvom studenata te poticanjem kontinuiranog razvoja obrazovanja.

6.6. Infrastruktura

Anonimizirani lokacijski podaci mogli bi imati ključnu ulogu u poboljšanju infrastrukture i pružanju efikasnijih usluga u navedenom sektoru u vidu planiranja, upravljanja i optimiziranja infrastrukturnih resursa.

Strateško planiranje infrastrukture koristilo bi anonimizirane lokacijske podatke kako bi razumjelo obrasce kretanja i prisutnost ljudi u različitim područjima. Kroz analizu ovih podataka, pružatelji usluga mogli bi identificirati područja koja zahtijevaju dodatne infrastrukturne investicije te prilagoditi planove razvoja kako bi bolje odgovarali potrebama korisnika.

Reorganizacija koncesija i licenci koristila bi anonimizirane lokacijske podatke za bolje upravljanje infrastrukturnim projektima i resursima. Kroz analizu podataka o kretanju ljudi i njihovim aktivnostima, pružatelji usluga mogli bi bolje razumjeti potrebe korisnika, identificirati područja s povećanom potražnjom te prilagoditi koncesije i licence kako bi bolje odgovarale trenutnim potrebama.

Transportne usluge mogle bi koristiti navedene podatke za optimizaciju rute i upravljanje prometom. Kroz analizu podataka o kretanju vozila i putnika, pružatelji usluga bi tako mogli identificirati prometne gužve, predviđati potražnju te prilagoditi ponudu kako bi bolje odgovarala potrebama korisnika.

Usluge za pronalaženje izgubljenih predmeta i planiranje parkinga mogle bi koristiti anonimizirane lokacijske podatke kako bi pružile bolje usluge korisnicima. Kroz analizu podataka o kretanju ljudi i njihovim aktivnostima, ovim uslugama bi bilo omogućena identifikaciju lokacija izgubljenih predmeta i slobodnih parkirnih mjesta, što bi rezultiralo poboljšanim korisničkim iskustvom i većim zadovoljstvom korisnika.

Korištenjem anonimiziranih lokacijskih podataka, infrastrukturne usluge mogle bi se prilagoditi potrebama korisnika, optimizirati resurse i procese te osigurati bolje iskustvo korisnicima.

6.7. Sigurnost

U sigurnosnim uslugama, analiza anonimiziranih podataka kretanja korisnika bi također imala ključnu ulogu jer bi kroz detaljno praćenje kretanja i analize aktivnosti korisnika, pružatelji ovih usluga mogli brže reagirati na hitne situacije, kao i optimizirati resurse i pružiti bolju podršku svojim korisnicima.

Uslugama za pronalaženje izgubljenih predmeta, anonimizirani lokacijski podaci olakšavali bi pronalazak izgubljenih predmeta, poput mobitela, novčanika ili ključeva. S pomoću ovih podataka, korisnicima bi se pružale informacije poput posljednje poznate lokacije izgubljenog predmeta, što bi značajno povećavale šansu za pronalazak.

Tako bi usluge za hitne slučajeve mogle koristiti ovu vrstu podataka kako bi brže i efikasnije odgovorile na hitne situaciju te primjerice identificirali lokaciju hitnog događaja, predvidjele najbržu rutu dolaska hitne službe te koordinirale preostale akcije spašavanja.

Osim toga, anonimizirani lokacijski podaci mogli bi se koristiti za analizu sigurnosnih prijetnji i prevenciju kriminala identifikacijom područja s povećanim rizikom te poduzimanjem preventivnih mjera u svrhu smanjivanja sigurnosnih prijetnji te povećavanjem sigurnosti građana.

Zaključno korištenjem anonimiziranih lokacijskih podataka u ovoj kategoriji usluga rezultiralo bi većom sigurnošću građana odnosno poboljšanom kvalitetom života u zajednici.

6.8. Financije

Značajan je utjecaj lokacijskih podataka na sektor financija, te bi analiza istih imala zasigurno utjecaj u vidu poboljšanja usluga, optimiziranja procesa te osiguravanja boljeg iskustva korisnika.

Usluge naplate i bankarstva mogle bi koristiti navedene podatke kako bi mogle bolje razumjeti potrebe korisnika te pružale personalizirane financijske usluge. Također,

poslužitelji ove vrste usluge mogli bi identificirati obrasce potrošnje, predvidjeti financijske potrebe te prilagoditi ponudu kako bi odgovarala potrebama korisnika.

U uslugama za istraživanje tržišta i analitiku korištenje lokacijskih podataka omogućilo bi pružanje uvide u ponašanje potrošača, trendove u potražnji te konkurentsku analizu. Tako bi pružatelji ovih usluga mogli identificirati ključne utjecajne faktore tržišta te prilagoditi strategiju za ostvarivanje konkurentске prednosti na tržištu.

Usluge za pronalaženje posla mogle bi koristiti anonimizirane lokacijske podatke kako bi pružile bolje usluge tražiteljima posla i poslodavcima. Kroz analizu podataka o kretanju ljudi i njihovim aktivnostima, pružatelji usluga mogli bi identificirati potencijalne kandidate za posao, prepoznati trendove na tržištu rada te olakšati pronalazak odgovarajućih poslova i kandidata.

Analiza anonimiziranih lokacijskih podataka imala bi značajan utjecaj na usluge osiguranja, omogućujući osiguravajućim društvima bolje razumijevanje potreba korisnika, identifikaciju rizika te prilagodbu svojih usluga korisničkim potrebama. Osiguravatelji bi također mogli poboljšati procjenu rizika odnosno identificirati područja s povećanim rizikom od šteta te predvidjeti potencijalne štete, personalizirati ponude te osigurati adekvatniju podrške korisnicima.

Osim toga, anonimizirani lokacijski podaci mogli bi se koristiti za poboljšanje sigurnosti i prevenciju šteta tako što bi osiguravatelji mogli identificirati potencijalne rizike te poduzeti preventivne mjere kako bi smanjili vjerojatnost šteta i poboljšali sigurnost korisnika.

Korištenjem anonimiziranih lokacijskih podataka, financijske usluge bi mogle prilagoditi svoje ponude, poboljšati korisničko iskustvo te ostvariti konkurentsku prednost na tržištu. Ovo bi rezultiralo većom efikasnošću, boljim financijskim odlukama te većim zadovoljstvom korisnika.

7. Primjena anonimiziranih transformacija lokacijskih podataka u uslugama različite namjene

U ovom je poglavlju prezentirana povezanost transformacija lokacijskih podataka s uslugama u kojima bi te transformacije imale najviše primjena i značaja. Također nužna je anonimizacija navedenih podataka kako bi se uslugama omogućilo nesmetano korištenje bez kršenja zakona o privatnosti podataka. Na taj način usluge bi mogle na temelju takvih rezultata unaprijediti svoje poslovanje, dok bi korisnici usluga bili zaštićeni od zlouporabe podataka.

7.1. Povezanost usluga s određenom vrstom transformacije lokacijskih podataka

U nastavku su priložene tablice u kojima su prikazane transformacije koji bi se mogle koristiti za usluge određene kategorije.

Tablica 7.1.: Prikaz povezanih usluga marketinga s transformacijama

društvene mreže	Transformacija 1: grupiranje posjećenosti određene lokacije u određenom vremenskom rasponu u danu Transformacija 2: najčešće posjećene lokacije u određenom dijelu grada Transformacija 3: transformacija posjećenih lokacija s obzirom na dob korisnika Transformacija 4: transformacija posjećenih lokacija s obzirom na lokaciju stanovanja Transformacija 5: transformacija posjećenih lokacija s obzirom na lokaciju posla
usluge za pronalaženje posla	Transformacija 1: grupiranje posjećenosti određene lokacije u određenom vremenskom rasponu u danu

	<p>Transformacija 2: najčešće posjećene lokacije u određenom dijelu grada</p> <p>Transformacija 3: transformacija posjećениh lokacija s obzirom na dob korisnika</p> <p>Transformacija 4: transformacija posjećениh lokacija s obzirom na lokaciju stanovanja</p>
transportne usluge	<p>Transformacija 1: grupiranje posjećениh određene lokacije u određenom vremenskom rasponu u danu</p> <p>Transformacija 2: najčešće posjećene lokacije u određenom dijelu grada</p> <p>Transformacija 3: transformacija posjećениh lokacija s obzirom na dob korisnika</p> <p>Transformacija 4: transformacija posjećениh lokacija s obzirom na lokaciju stanovanja</p> <p>Transformacija 5: transformacija posjećениh lokacija s obzirom na lokaciju posla</p> <p>Transformacija 8: transformacija brzine kretanja</p>
usluge za istraživanje tržišta i analitika	<p>Transformacija 1: grupiranje posjećениh određene lokacije u određenom vremenskom rasponu u danu</p> <p>Transformacija 2: najčešće posjećene lokacije u određenom dijelu grada</p> <p>Transformacija 3: transformacija posjećениh lokacija s obzirom na dob korisnika</p> <p>Transformacija 4: transformacija posjećениh lokacija s obzirom na lokaciju stanovanja</p>

	Transformacija 5: transformacija posjećenih lokacija s obzirom na lokaciju posla
turističke usluge	<p>Transformacija 1: grupiranje posjećenosti određene lokacije u određenom vremenskom rasponu u danu</p> <p>Transformacija 2: najčešće posjećene lokacije u određenom dijelu grada</p> <p>Transformacija 3: transformacija posjećenih lokacija s obzirom na dob korisnika</p> <p>Transformacija 4: transformacija posjećenih lokacija s obzirom na lokaciju stanovanja</p> <p>Transformacija 5: transformacija posjećenih lokacija s obzirom na lokaciju posla</p>
usluge naplate i bankarstva	<p>Transformacija 2: najčešće posjećene lokacije u određenom dijelu grada</p> <p>Transformacija 3: transformacija posjećenih lokacija s obzirom na dob korisnika</p> <p>Transformacija 4: transformacija posjećenih lokacija s obzirom na lokaciju stanovanja</p> <p>Transformacija 5: transformacija posjećenih lokacija s obzirom na lokaciju posla</p>

Tablica 7.2.: Prikaz povezanih usluga zdravstva s transformacijama

usluge za vježbanje i zdravlje	<p>Transformacija 1: grupiranje posjećenosti određene lokacije u određenom vremenskom rasponu u danu</p> <p>Transformacija 2: najčešće posjećene lokacije u određenom dijelu grada</p>
--------------------------------	--

	<p>Transformacija 3: transformacija posjećenih lokacija s obzirom na dob korisnika</p> <p>Transformacija 4: transformacija posjećenih lokacija s obzirom na lokaciju stanovanja</p> <p>Transformacija 5: transformacija posjećenih lokacija s obzirom na lokaciju posla</p>
<p>usluge za upravljanje energijom i ekološke inicijative</p>	<p>Transformacija 1: grupiranje posjećenosti određene lokacije u određenom vremenskom rasponu u danu</p> <p>Transformacija 2: najčešće posjećene lokacije u određenom dijelu grada</p> <p>Transformacija 4: transformacija posjećenih lokacija s obzirom na lokaciju stanovanja</p> <p>Transformacija 5: transformacija posjećenih lokacija s obzirom na lokaciju posla</p>
<p>usluge za praćenje epidemija u zdravstvu</p>	<p>Transformacija 3: transformacija posjećenih lokacija s obzirom na dob korisnika</p> <p>Transformacija 4: transformacija posjećenih lokacija s obzirom na lokaciju stanovanja</p>

Tablica 7.3.: Prikaz povezanih usluga turizma s transformacijama

<p>usluge za planiranje parkinga</p>	<p>Transformacija 1: grupiranje posjećenosti određene lokacije u određenom vremenskom rasponu u danu</p> <p>Transformacija 2: najčešće posjećene lokacije u određenom dijelu grada</p> <p>Transformacija 4: transformacija posjećenih lokacija s obzirom na lokaciju stanovanja</p> <p>Transformacija 5: transformacija posjećenih lokacija s obzirom na lokaciju posla</p> <p>Transformacija 8: transformacija brzine kretanja</p>
<p>turističke usluge</p>	<p>Transformacija 1: grupiranje posjećenosti određene lokacije u određenom vremenskom rasponu u danu</p> <p>Transformacija 2: najčešće posjećene lokacije u određenom dijelu grada</p>

Tablica 7.4.: Prikaz povezanih demografskih usluga s transformacijama

<p>migracije stanovništva</p>	<p>Transformacija 1: grupiranje posjećenosti određene lokacije u određenom vremenskom rasponu u danu</p> <p>Transformacija 2: najčešće posjećene lokacije u određenom dijelu grada</p>
-------------------------------	--

	<p>Transformacija 4: transformacija posjećenih lokacija s obzirom na lokaciju stanovanja</p>
<p>demografske promjene</p>	<p>Transformacija 1: grupiranje posjećenosti određene lokacije u određenom vremenskom rasponu u danu</p> <p>Transformacija 2: najčešće posjećene lokacije u određenom dijelu grada</p> <p>Transformacija 4: transformacija posjećenih lokacija s obzirom na lokaciju stanovanja</p> <p>Transformacija 3: transformacija posjećenih lokacija s obzirom na dob korisnika</p>

Tablica 7.5.: Prikaz povezanih obrazovnih usluga s transformacijama

<p>usluge za edukaciju</p>	<p>Transformacija 1: grupiranje posjećenosti određene lokacije u određenom vremenskom rasponu u danu</p> <p>Transformacija 2: najčešće posjećene lokacije u određenom dijelu grada</p> <p>Transformacija 3: transformacija posjećenih lokacija s obzirom na dob korisnika</p>
----------------------------	---

Tablica 7.6.: Prikaz povezanih usluga infrastrukture s transformacijama

<p>strateško planiranje i infrastruktura</p>	<p>Transformacija 1: grupiranje posjećenosti određene lokacije u određenom vremenskom rasponu u danu</p> <p>Transformacija 3: transformacija posjećenih lokacija s obzirom na dob korisnika</p> <p>Transformacija 4: transformacija posjećenih lokacija s obzirom na lokaciju stanovanja</p> <p>Transformacija 5: transformacija posjećenih lokacija s obzirom na lokaciju posla</p>
<p>(Re)organizacija koncesija i licenci</p>	<p>Transformacija 1: grupiranje posjećenosti određene lokacije u određenom vremenskom rasponu u danu</p> <p>Transformacija 2: najčešće posjećene lokacije u određenom dijelu grada</p> <p>Transformacija 3: transformacija posjećenih lokacija s obzirom na dob korisnika</p> <p>Transformacija 4: transformacija posjećenih lokacija s obzirom na lokaciju stanovanja</p> <p>Transformacija 5: transformacija posjećenih lokacija s obzirom na lokaciju posla</p>

<p>transportne usluge</p>	<p>Transformacija 1: grupiranje posjećenosti određene lokacije u određenom vremenskom rasponu u danu</p> <p>Transformacija 2: najčešće posjećene lokacije u određenom dijelu grada</p> <p>Transformacija 3: transformacija posjećenih lokacija s obzirom na dob korisnika</p> <p>Transformacija 4: transformacija posjećenih lokacija s obzirom na lokaciju stanovanja</p> <p>Transformacija 5: transformacija posjećenih lokacija s obzirom na lokaciju posla</p> <p>Transformacija 8: transformacija brzine kretanja</p>
<p>usluge vremenske prognoze</p>	<p>Transformacija 1: grupiranje posjećenosti određene lokacije u određenom vremenskom rasponu u danu</p> <p>Transformacija 2: najčešće posjećene lokacije u određenom dijelu grada</p> <p>Transformacija 4: transformacija posjećenih lokacija s obzirom na lokaciju stanovanja</p> <p>Transformacija 5: transformacija posjećenih lokacija s obzirom na lokaciju posla</p> <p>Transformacija 8: transformacija brzine kretanja</p>

<p>usluge za planiranje parkinga</p>	<p>Transformacija 1: grupiranje posjećenosti određene lokacije u određenom vremenskom rasponu u danu</p> <p>Transformacija 2: najčešće posjećene lokacije u određenom dijelu grada</p> <p>Transformacija 4: transformacija posjećenih lokacija s obzirom na lokaciju stanovanja</p> <p>Transformacija 5: transformacija posjećenih lokacija s obzirom na lokaciju posla</p>
--------------------------------------	---

Tablica 7.7.: Prikaz povezanih sigurnosnih usluga s transformacijama

<p>usluge za hitne slučajeve</p>	<p>Transformacija 1: grupiranje posjećenosti određene lokacije u određenom vremenskom rasponu u danu</p> <p>Transformacija 2: najčešće posjećene lokacije u određenom dijelu grada</p> <p>Transformacija 4: transformacija posjećenih lokacija s obzirom na lokaciju stanovanja</p> <p>Transformacija 5: transformacija posjećenih lokacija s obzirom na lokaciju posla</p> <p>Transformacija 8: transformacija brzine kretanja</p>
----------------------------------	---

<p>usluge za pronalaženje izgubljenih predmeta</p>	<p>Transformacija 1: grupiranje posjećenosti određene lokacije u određenom vremenskom rasponu u danu</p> <p>Transformacija 2: najčešće posjećene lokacije u određenom dijelu grada</p>
--	--

Tablica 7.8.: Prikaz povezanih financijskih usluga s transformacijama

<p>usluge naplate i bankarstva</p>	<p>Transformacija 1: grupiranje posjećenosti određene lokacije u određenom vremenskom rasponu u danu</p> <p>Transformacija 2: najčešće posjećene lokacije u određenom dijelu grada</p> <p>Transformacija 3: transformacija posjećenih lokacija s obzirom na dob korisnika</p> <p>Transformacija 4: transformacija posjećenih lokacija s obzirom na lokaciju stanovanja</p> <p>Transformacija 5: transformacija posjećenih lokacija s obzirom na lokaciju posla</p>
<p>usluge za istraživanje tržišta i analitika</p>	<p>Transformacija 1: grupiranje posjećenosti određene lokacije u određenom vremenskom rasponu u danu</p> <p>Transformacija 2: najčešće posjećene lokacije u određenom dijelu grada</p> <p>Transformacija 3: transformacija posjećenih lokacija s obzirom na dob</p>

	<p>korisnika</p> <p>Transformacija 4: transformacija posjećenih lokacija s obzirom na lokaciju stanovanja</p> <p>Transformacija 5: transformacija posjećenih lokacija s obzirom na lokaciju posla</p>
<p>usluge za pronalaženje posla</p>	<p>Transformacija 3: transformacija posjećenih lokacija s obzirom na dob korisnika</p> <p>Transformacija 4: transformacija posjećenih lokacija s obzirom na lokaciju stanovanja</p> <p>Transformacija 5: transformacija posjećenih lokacija s obzirom na lokaciju posla</p>
<p>usluge osiguranja</p>	<p>Transformacija 1: grupiranje posjećenosti određene lokacije u određenom vremenskom rasponu u danu</p> <p>Transformacija 2: najčešće posjećene lokacije u određenom dijelu grada</p> <p>Transformacija 3: transformacija posjećenih lokacija s obzirom na dob korisnika</p> <p>Transformacija 4: transformacija posjećenih lokacija s obzirom na lokaciju stanovanja</p> <p>Transformacija 5: transformacija posjećenih lokacija s obzirom na lokaciju posla</p>

	Transformacija 8: transformacija brzine kretanja
--	---

Sama povezanost usluga i transformacija ovisi i krajnjim ciljevima i namjerama usluge koja ih koristi, prethodno je definirana njihova općenita povezanost.

8. Izrada profila korisnika na temelju lokacijskih podataka

U ovome poglavlju stavljen je naglasak na procesu profiliranja korisnika na temelju prikupljenih podataka iz zapisa CDR-a te provedenih transformacijama.

Prvi korak u kreiranju korisničkog profila je prikupljanje podataka. Podaci na temelju kojih će biti izrađeni korisnički profili biti će korisnički podaci te podaci koji se nalaze u zapisima CDR-a. Zapisi CDR-a, kao što je i navedeno, predstavljaju ulazni skup podataka na temelju kojih se zasniva ovaj cjelokupni istraživački rad.

Korisnički podaci koje teleoperatori posjeduju o svojim korisnicima te koji će biti iskorišteni u profiliranju korisnika su sljedeći:

- naziv fizičke/pravne osobe
- adresa boravišta na temelju adrese dostave računa
- email adresa
- kontakt telefona
- OIB
- datum rođenja
- spol
- broj osobne iskaznice
- serijski broj SIM kartice
- pretplatnički broj
- model uređaja
- tarifni model
- statistika korištenja pojedinih usluga
- statistika računa.

Rezultati transformacija opisanih u poglavlju 4. omogućuju proširivanje skupa poznatih podataka o korisnicima sa sljedećim podacima:

- lokacija posla
- prosječno vrijeme provedeno na poslu
- najčešće posjećena lokacija
- prosječno vrijeme provedeno na najčešćoj lokaciji

- prosječno vrijeme provedeno na lokaciji stanovanja
- najčešće posjećene lokacije primarnog sektora djelatnosti te prosječno vrijeme zadržavanja
- najčešće posjećene lokacije sekundarnog sektora djelatnosti te prosječno vrijeme zadržavanja
- najčešće posjećene lokacije tercijarnog sektora djelatnosti te prosječno vrijeme zadržavanja
- najčešće korištena vrsta prijevoza
- prosječna brzina kretanja
- prosječna udaljenost posjećenih lokacija od mjesta stanovanja
- prosječna udaljenost posjećenih lokacija od mjesta posla.

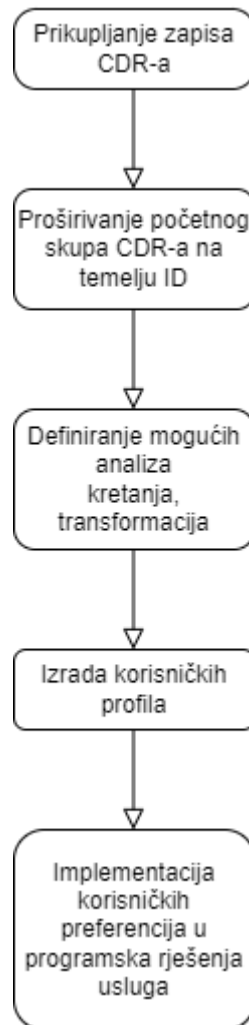
Navedene podatke moguće je dobiti provođenjem transformacija nad ulaznim skupom zapisa CDR-a. Navedeni format zapisa CDR-a moguće je proširiti s različitim vrstama podataka te tako povećati broj dostupnih podataka o korisnicima.

8.1.Izrada korisničkih profila na proširenom skupu korisničkih podataka

U ovome potpoglavlju istaknuta je mogućnost proširivanja skupa ulaznih podataka te isto tako proširivanje poznatih informacija o korisnicima. Početni skup zapisa CDR-a moguće je proširiti određenim dodatnim skupovima zapisa CDR-a koji sadrže dodatni skup informacija, drugačiji od početnog. Važno je istaknuti činjenicu da je u tom procesu proširivanja nužno definirati podatak s pomoću kojeg će ta dva skupa biti povezani, to se naravno najčešće odnosi na jedinstveni identifikator korisnika. S obzirom na vrstu usluge u kojoj će podaci biti korišteni, dodatni skup treba prilagoditi vrsti usluge. Na primjer, ovisno o vrsti usluge mogli bi proširiti skup s informacijama o platnim listama, povijesti plaćanja, saldu računa, najčešćem obliku komunikacije, IP i MAC adresa uređaja i dr. Nužno je naglasiti kako upravo ovakav prošireni skup podataka na raspolaganju imaju teleoperatori. U ovom istraživačkom radu prošireni skup se nije koristio, ali je važno istaknuti koje sve dodatne informacije posjeduju teleoperatori.

Analiza kretanja korisnika na proširenom skupu podataka provodi se kao i na početnom te je shodno tomu omogućeno dodavanje novih vrsta transformacija.

Također, proširivanje početnog skupa može dovesti do proširivanja vrsta usluga, odnosno dodavanja novih vrsta kategorija usluga te izrađivanja korisničkih profila. Zaključno, što je skup podataka veći to je veći broj transformacija te veći broj usluga u kojima su te transformacije primjenjive. Cjelokupni proces proširivanja početnog skupa podataka bit će obuhvaćen na sljedećem grafičkom prikazu.



Slika 8.1. Grafički prikaz procesa proširivanja zapisa CDR-a

Zaključno, cilj je ostvariti „otvaranje“ teleoperatera prema uslugama te na taj način omogućiti uslugama unaprjeđivanje, poboljšanje rada te prilagođavanje preferencijama korisnika. S druge strane, s obzirom na to da se radi o osjetljivim podacima nužno je osigurati njihovu anonimizaciju, kako ne bi došlo do ugrožavanja privatnosti i kršenja GDPR-a. Analize kretanja korisnika odnosno transformacije predstavljaju metode anonimizacije te su navedene dovoljne kako bi se sačuvala privatnost zato što u većini slučajeva nije od interesa promatrati jednog korisnika nego grupu korisnika.

Zaključak

Ovaj rad ističe važnost podizanja svijesti o zaštiti lokacijskih podataka, koji u današnje vrijeme predstavljaju jednu od najvrjednijih informacija jer je na temelju njih moguće saznati dodatne informacije o korisniku, što je u ovom radu i napravljeno kroz definiranje analiza kretanja korisnika. Početni ulazni skup na kojem se temelji cjelokupno istraživanje su zapisi CDR-a koji sadrže detalje uspostavljenog poziva. Skup podataka koji su pohranjene u zapisu, ovisi o razini ustanove koje generira iste. S pomoću atributa iz zapisa CDR-a definirana je početna i završna lokacija korisnika. Početnu lokaciju čini kombinacija atributa FirstLAC i FirstCID, a završnu kombinacija LastLAC i LastCID. Na temelju prethodnih lokacija, definirane su analize kretanja korisnika koje se odnose na posjećenost lokacije u vremenskom rasponu, posjećenost korisnika na određenoj lokaciji, posjećenost lokacije s obzirom na dob, posjećenost s obzirom na lokaciju stanovanja i lokaciju posla, kao i posjećenost s obzirom na broj upućenih i primljenih poziva. Također je definirana i transformacija koja određuje brzinu kretanja korisnika s pomoću matrice udaljenosti i Google Maps Apija. U navedenoj analizi brzina je definirana kao put, koji izračuna matrica udaljenosti, kroz vrijeme, koje predstavlja trajanje poziva. Ovakav pristup ima brojne potencijalne probleme u vidu zaštite privatnosti korisnika, ali i integraciji rezultata prethodno navedenih transformacija u postojeće sustave. Također, jedan od ključnih potencijalnih problema predstavlja i ažurnost i točnost podataka, stoga je važno prije analiza pripremiti podatke i ukloniti one „bezvrijedne“. „Bezvrijedni“ podaci se odnose na slučajeve u kojima je početna lokacija jednaka završnoj, odnosno na slučaj u kojem je jedna od lokacija *null* vrijednost. Rezultati provedenih analiza otkrivaju širok skup informacija o korisnicima koji je primjenjiv u različitim vrstama usluga. Neke od usluga u kojima su primjenjivi su financije, infrastruktura, demografija, zdravstvo i dr. Korištenje ovakvog tipa podataka omogućuje uslugama poboljšanje njihova rada te prilagođavanje rada preferencijama korisnika. Istovremeno, ovakvo korištenje zahtjeva i zaštitu privatnosti korisničkih podataka, što se i ostvaruje provođenjem analiza kretanja korisnika. Cjelokupna ideja rada je povezivanja transformacija i usluga jer različite analize kretanja korisnika interesantne su različitim vrstama usluga. Samo povezivanje provedeno je logički, s obzirom na područje djelovanja svake usluge. Što je veći broj transformacija, to je veći broj usluga u kojima su rezultati istih

primjenjivi. Za svaku uslugu određena informacija ima određenu vrijednost. Provođenje transformacija povećava broj poznatih informacija o korisniku, odnosno rezultira izrađivanjem korisničkih profila koji se povezuju s uslugama kojima takvi podaci predstavljaju vrijednost. Zapise CDR-a moguće je proširiti i tako dodatno povećati broj poznatih informacija o korisniku. Prošireni skup potrebno je povezati s početnim skupom na temelju zajedničkog atributa. Istaknuto je kako teleoperatori posjeduju mnogobrojne proširene skupove o svojim korisnicima kao što su platne liste, povijest plaćanja, MAC adrese uređaja i dr. te na taj način mogu provesti još dublju analizu kretanja korisnika.

Literatura

- [1] Fan, W., Lin, Z. Evaluating the Potential Use of Crowdsourced Bicycle Data in North Carolina. Istraživački rad. University of North Carolina at Charlotte, 2018.
- [2] Brad Deflin, "Google's \$57M GDPR Fine: Analyzing Key Takeaways for Privacy Compliance," CPO Magazine, <https://www.cpomagazine.com/data-privacy/googles-57m-gdpr-fine-analyzing-key-takeaways-for-privacy-compliance/>, pristupljeno 28. lipnja 2024.
- [3] Hoang, N. P., Asano, Y., Yoshikawa, M. Your neighbors are my spies: Location and other privacy concerns in GLBT-focused location-based dating applications. International Conference on Advanced Communication Technology. PyeongChang, Korea, (2017)
- [4] Hayes, D. R., Snow, C., Altuwayjiri, S. Geolocation Tracking and Privacy Issues Associated with the Uber Mobile Application. Istraživački rad. Pace University New York, 2017.
- [5] Gruteser, M., Grunwald, D. Anonymous Usage of Location-Based Services Through Spatial and Temporal Cloaking. MobiSys. New York, United States, (2003), 21-42
- [6] Alishahi, M., Fathalizadeh, A., Moghtadaiee, V. On the privacy protection of indoor location dataset using anonymization. Open University of the Netherlands, University of Amsterdam, Shahid Beheshti University, 2022.
- [7] HERE Platform, HERE Technologies, <https://platform.here.com/portal/>, pristupljeno 28. lipnja 2024.
- [8] Zhang, H., Zhou, Z., Ye, L., Du, X. Towards Privacy Preserving Publishing of Set-Valued Data on Location-based Services. IEEE Transactions on Cloud Computing, 6, 1(2018), 316-329
- [9] Jaing, H., Li, J., Zhao, P., Zeng, F., Xiao, Z., Iyengar, A. Location Privacy-preserving Mechanisms in Location-based Services: A Comprehensive Survey. ACMJournals. 54, 1(2021), 1-36
- [10] OpenDP Platform, OpenDP, <https://opendp.org>, pristupljeno 28. lipnja 2024.

- [11] Gupta, A. K., Shanker, U., Location Privacy Predervation for Location Based Service Applications: Taxonomies, Issues and Future Research Directions. Springer Link. 134(2024), 1617-1639
- [12] Adapt IT Telecoms, "Analytics in Telecoms White Paper," Adapt IT, <https://telecoms.adaptit.tech/resources/analytics-white-paper>, pristupljeno 28. lipnja 2024.

Sažetak

Naslov: Analiza kretanja korisnika temeljem anonimiziranih lokacijskih podataka telekomunikacijske mreže

Sažetak: Ovaj rad istražuje važnost zaštite lokacijskih podataka korisnika u uslugama različitih vrsta, s ciljem dobivanja uvida u obrasce kretanja i ponašanja korisnika. Analizom zapisa CDR-a definirane su analize kretanja korisnika odnosno transformacije. Rezultati navedenih transformacija predstavljaju dodatne informacije o korisnicima na temelju kojih je moguća izrada korisnički profila. Korisničke profile moguće je filtrirati s obzirom na usluge kojima su takvi podaci od interesa i koristi. Što je veći broj poznatih atributa o korisnicima, to je veći skup primjene istih u većem broju različitih usluga. Korištenjem ovakvog tipa informacije, usluge različitih vrsta imaju mogućnost poboljšanja rada te prilagodbe rada prema preferencijama svojih korisnika, istovremeno osiguravajući zaštitu privatnosti, što je u današnjem vremenu od iznimne važnosti.

Ključne riječi: zapis CDR-a, usluge, analize kretanja korisnika, zaštita privatnosti korisnika, anonimizacija, telekomunikacijska mreža, korisnički profil

Summary

Title: User movement analysis based on anonymized location data from telecommunications networks

Summary: This thesis defines the importance of protecting user's location data in different types of services, with the aim of gaining insight into user movement and behavior patterns. The analysis of the CDR record defines the analysis of the user's movement, i.e. the transformation. The results of the mentioned transformations represent additional information about users, based on which it is possible to create a user profile. It is possible to filter user profiles about the services for which such data is of interest and use. The greater the number of known attributes about users, the greater the set of applications of the same in a greater number of different services. By using this type of information, services of various types could improve their work and adjust their work according to the preferences of their users, while at the same time ensuring privacy protection, which is extremely important nowadays.

Keywords: CDR record, services, user movement analysis, user privacy protection, anonymization, telecommunications network, user profile