

Prikupljanje podataka iz mnoštva uporabom mobilnih platformi

Jukić, Luka

Undergraduate thesis / Završni rad

2023

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **Josip Juraj Strossmayer University of Osijek, Faculty of Electrical Engineering, Computer Science and Information Technology Osijek / Sveučilište Josipa Jurja Strossmayera u Osijeku, Fakultet elektrotehnike, računarstva i informacijskih tehnologija Osijek**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:200:358709>

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

Download date / Datum preuzimanja: **2024-12-27**

Repository / Repozitorij:

[Faculty of Electrical Engineering, Computer Science and Information Technology Osijek](#)



**SVEUČILIŠTE JOSIPA JURJA STROSSMAYERA U OSIJEKU
FAKULTET ELEKTROTEHNIKE, RAČUNARSTVA I
INFORMACIJSKIH TEHNOLOGIJA OSIJEK**

Sveučilišni studij

**PRIKUPLJANJE PODATAKA IZ MNOŠTVA
UPORABOM MOBILNIH PLATFORMI**

Završni rad

Luka Jukić

Osijek, 2023.

**FERIT**FAKULTET ELEKTROTEHNIKE, RAČUNARSTVA
I INFORMACIJSKIH TEHNOLOGIJA **OSIJEK****Obrazac Z1P - Obrazac za ocjenu završnog rada na preddiplomskom sveučilišnom studiju**

Osijek, 28.08.2023.

Odboru za završne i diplomske ispite

Prijedlog ocjene završnog rada na preddiplomskom sveučilišnom studiju

Ime i prezime Pristupnika:	Luka Jukić
Studij, smjer:	Programsko inženjerstvo
Mat. br. Pristupnika, godina upisa:	R4506, 27.07.2020.
OIB Pristupnika:	35834854621
Mentor:	doc. dr. sc. Bruno Zorić
Sumentor:	,
Sumentor iz tvrtke:	
Naslov završnog rada:	Prikupljanje podataka iz mnoštva uporabom mobilnih platformi
Znanstvena grana rada:	Informacijski sustavi (zn. polje računarstvo)
Zadatak završnog rad:	U teorijskom dijelu rada potrebno je opisati podatke iz mnoštva (engl. crowdsourced data), njihovu primjenu te načine njihova prikupljanja. Posebno se osvrnuti na mogućnost korištenja mobilnih platformi za prikupljanje takvih podataka. U praktičnom dijelu rada programski ostvariti alat za prikupljanje podataka iz mnoštva. (Tema rezervirana za: <u>Luka Jukić</u>)
Prijedlog ocjene završnog rada:	Izvrstan (5)
Kratko obrazloženje ocjene prema Kriterijima za ocjenjivanje završnih i diplomskih radova:	Primjena znanja stečenih na fakultetu: 3 bod/boda Postignuti rezultati u odnosu na složenost zadatka: 3 bod/boda Jasnoća pismenog izražavanja: 3 bod/boda Razina samostalnosti: 3 razina
Datum prijedloga ocjene od strane mentora:	28.08.2023.
Datum potvrde ocjene od strane Odbora:	08.09.2023.
Potvrda mentora o predaji konačne verzije rada:	<i>Mentor elektronički potpisao predaju konačne verzije.</i>
	Datum:

**FERIT**FAKULTET ELEKTROTEHNIKE, RAČUNARSTVA
I INFORMACIJSKIH TEHNOLOGIJA OSIJEK**IZJAVA O ORIGINALNOSTI RADA**

Osijek, 09.09.2023.

Ime i prezime studenta:

Luka Jukić

Studij:

Programsko inženjerstvo

Mat. br. studenta, godina upisa:

R4506, 27.07.2020.

Turnitin podudaranje [%]:

3

Ovom izjavom izjavljujem da je rad pod nazivom: **Prikupljanje podataka iz mnoštva uporabom mobilnih platformi**

izrađen pod vodstvom mentora doc. dr. sc. Bruno Zorić

i sumentora ,

moj vlastiti rad i prema mom najboljem znanju ne sadrži prethodno objavljene ili neobjavljene pisane materijale drugih osoba, osim onih koji su izričito priznati navođenjem literature i drugih izvora informacija. Izjavljujem da je intelektualni sadržaj navedenog rada proizvod mog vlastitog rada, osim u onom dijelu za koji mi je bila potrebna pomoć mentora, sumentora i drugih osoba, a što je izričito navedeno u radu.

Potpis studenta:

SADRŽAJ

1. UVOD	1
1.1. Sadržaj završnog rada	1
2. PODACI IZ MNOŠTVA	2
2.1. Oblici i svrha prikupljanja podataka iz mnoštva.....	2
2.2. Sustavi za prikupljanje resursa iz mnoštva	3
2.2.1. Osnovni pojmovi	3
2.2.2. Načini prikupljanja podataka iz mnoštva.....	5
2.3. Dizajn i modeliranje sustava za prikupljanje podataka iz mnoštva	5
2.3.1. Općeniti koraci pri izradi sustava za prikupljanje podataka iz mnoštva	6
2.3.2. Ključni aspekti pri dizajniranju mobilnih sustava za prikupljanje resursa iz mnoštva	8
2.3.3. Arhitekture sustava za mobilno prikupljanje podataka iz mnoštva	12
2.4. Izazovi vezani uz prikupljanje podataka	15
2.4.1. Kontrola kvalitete	15
2.4.2. Kontrola troška	15
2.4.3. Kontrola latencije	16
2.4.4. Izazovi vezani uz mobilno prikupljanje iz mnoštva	16
2.5. Primjeri primjena prikupljanja podataka.....	18
2.5.1. ImageNet	18
2.5.2. City Soundscape	18
2.5.3. Let Malaysia Airlinesa 370.....	20
2.5.4. HealthMap	20
2.5.5. Waze	21
2.6. Postojeće mobilne platforme i alati za prikupljanje podataka	22
2.6.1. Ushahidi.....	22
2.6.2. Crowdsorce by Google	23
2.6.3. Sleep as Android.....	24
2.6.4. Open Data Kit.....	25
3. PROGRAMSKO RJEŠENJE ZA PRIKUPLJANJE PODATAKA IZ MNOŠTVA.....	26
3.1. Zahtjevi na programsko rješenje.....	26
3.2. Korišteni alati i tehnologije	27
3.2.1. Android Studio	27
3.2.2. Firebase Authentication i Realtime Database	27
3.2.3. Rowy	28

3.2.4. Python.....	28
3.3. Prikaz načina rada aplikacije	28
3.3.1. Autentifikacija i početni zaslon	28
3.3.2. Ručno mjerenje.....	31
3.3.3. Automatsko mjerenje.....	35
3.3.4. Primjena.....	40
4. ZAKLJUČAK.....	41
LITERATURA	43
SAŽETAK.....	47
ABSTRACT	48
ŽIVOTOPIS.....	49
PRILOZI.....	50

1. UVOD

Razvoj Web 2.0 tehnologija u posljednjih 20-ak godina rezultirao je novim načinima prikupljanja podataka [1]. Korisnici svojim aktivnostima pružaju velike količine informacija web stranicama i online bazama podataka. Tako prikupljeni podaci omogućuju brži razvoj proizvoda i njihovu bolju prilagodbu ljudima i njihovim potrebama. Kao nova vrsta velikih podataka (engl. *big data*) pojavili su se podaci iz mnoštva (engl. *crowdsourced data*). Prikupljanje podataka ili resursa iz mnoštva (engl. *crowdsourcing*) podrazumijeva regrutiranje pojedinaca u svrhu obavljanja zadataka neizvedivih provedbom automatiziranih rutina ili previše zahtjevnih za računalne resurse [2]. Primjeri uključuju raspoznavanje entiteta (engl. *entity resolution*), analizu sentimenta (engl. *sentiment analysis*) i prepoznavanje slika (engl. *image recognition*) [3]. Prikupljanje podataka najčešće se izvodi putem online platformi ili aplikacija. Podaci iz mnoštva variraju u kvaliteti i točnosti jer su izvori resursa ljudi, često nedovoljno obrazovani u području koje se proučava. Pojavljuje se i problem privatnosti i sigurnosti podataka koji mogu sadržavati osjetljive informacije. Razvojem mobilnih uređaja koji su danas opremljeni moćnim procesorima, raznim sensorima, velikom memorijom, brzim modulima za bežičnu komunikaciju itd., podaci iz mnoštva dobili su na značaju i točnosti [4]. Primjeri korištenja mobilnih uređaja za prikupljanje podataka uključuju prikupljanje slika ili videozapisa (pomoću kamere), zvučnih zapisa (pomoću mikrofona), podataka o lokaciji (pomoću GPS modula) te raznih drugih kontekstualnih informacija.

U ovome radu proučavaju se podaci iz mnoštva i njihova primjena te sustavi odgovorni za njihovo prikupljanje. U drugome poglavlju istražuju se koraci u dizajniranju sustava te prepreke na koje se nailazi, uz posebnu pozornost posvećenu mobilnim sustavima. Navode se primjeri primjena prikupljanja podataka i mobilnih platformi za nabavu iz mnoštva. U trećem se poglavlju predstavlja praktično rješenje u kojemu se pomoću mobilnih uređaja prikupljaju podaci iz okoline korisnika. Posljednje poglavlje predstavlja zaključak i smjernice za budući rad.

1.1. Sadržaj završnog rada

U teorijskom dijelu rada potrebno je opisati podatke iz mnoštva (engl. *crowdsourced data*), njihovu primjenu te načine njihova prikupljanja. Posebno se osvrnuti na mogućnost korištenja mobilnih platformi za prikupljanje takvih podataka. U praktičnom dijelu rada programski ostvariti alat za prikupljanje podataka iz mnoštva.

2. PODACI IZ MNOŠTVA

U ovome poglavlju opisani su podaci iz mnoštva i sustavi za njihovo prikupljanje. Takvi podaci nastaju djelovanjem pojedinaca koji svojim radom stvaraju resurse pogodne za analizu i razvoj različitih djelatnosti. Raspravlja se o izazovima koji se tiču osiguranja kvalitete prikupljenih podataka, smanjenja troška i latencije. Nadalje, dani su primjeri nabave i mobilnih platformi za prikupljanje resursa iz mnoštva. Poseban je fokus na mobilnom prikupljanju koje je danas najzastupljenije u ovome području.

2.1. Oblici i svrha prikupljanja podataka iz mnoštva

Nekoliko je tipova prikupljanja podataka iz mnoštva koji uključuju [5] :

- Građanska znanost (engl. *citizen science*) uključuje sudjelovanje javnosti u znanstvenim istraživanjima na način da građani pružaju potrebne podatke kroz online platforme, mobilne aplikacije i terenski rad. Poznati primjer uključuje projekt *The Great Backyard Bird Count* [6] u kojemu građani identificiraju ptice.
- Izrada karti pomoću podataka iz mnoštva (engl. *crowdsourced mapping*) proces je u kojem javnost doprinosi kreiranju karti slanjem geografskih podataka o svome okruženju koristeći alate poput *OpenStreetMap* i *Google Map Maker*.
- Grupno financiranje (engl. *crowdfunding*) koristi se kada je potrebno iz javnosti prikupiti novčana sredstva za potporu nekakvog projekta, pomoći nekome itd. Primjeri kada se putem raznih platformi prikuplja novac za financiranje liječenja oboljelih osoba svakodnevno se susreću.
- Izrada sadržaja uz pomoć mnoštva (engl. *crowdsourced content creation*) uključuje doprinos javnosti u stvaranju članaka, videozapisa ili slika u posebne svrhe kao što su marketing ili obrazovanje. Sve društvene mreže i platforme poput *Youtube*-a nastale su kao posljedica navedenog.
- Testiranje od strane mnoštva (engl. *crowdsourced testing*) korisno je u slučaju kada je potrebno od javnosti dobiti povratnu informaciju o proizvodu ili usluzi testirajući ju te slanjem podataka o mogućim greškama.
- Generiranje ideja iz mnoštva (engl. *idea crowdsourcing*) slučaj je kada javnost pruža svoje ideje i rješenja na određene probleme, prepreke i izazove, koje mogu biti razmotrene i sprovedene u djelo. Najbolji primjeri pronalaze se u grafičkom dizajnu gdje mnoštvo pruža razne ideje za natpise na odjeći.

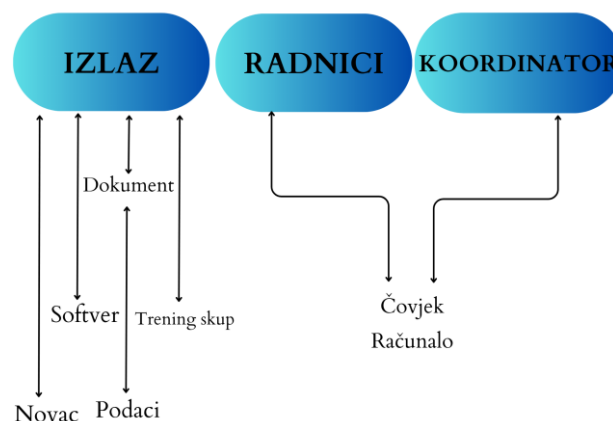
- Grupno istraživanje tržišta (engl. *crowdsourced market research*) postupak je u kojem mnoštvo pruža informacije o vlastitim preferencijama, ponašanjima i mišljenjima kako bi se potpomoglo razvoju određenog tržišta. *Google Play* se uvelike oslanja na korisnike u izradi prijedloga i lista popularnosti te koristi i sustav ocjenjivanja.
- Glasanje mnoštva (engl. *crowdvoting*) predstavlja sudjelovanje pojedinca koji svojim glasom doprinosi rješavanju problema. Primjer je glasanje publike povodom događaja poput *Eurosonga*.

2.2. Sustavi za prikupljanje resursa iz mnoštva

Pojam sustava za prikupljanje resursa iz mnoštva odnosi se na proces kreiranja i pružanja zadataka velikoj skupini ljudi [2]. Taj proces uključuje mehanizme i alate korištene za upravljanje mnoštvom (engl. *crowd*), prikupljanje i rad s podacima te pružanje nagrada i poticaja suradnicima (engl. *contributors*). Sustavom mogu upravljati određene organizacije tj. može biti centraliziran, dok se decentraliziran oslanja na mrežu nezavisnih sudionika.

2.2.1. Osnovni pojmovi

Neophodni za funkciju sustava za prikupljanje iz mnoštva su radnici, koordinatori i izlaz ili rezultat (engl. *output*) [2]. Radnici su pojedinci ili jedinice koje obavljaju ponuđene zadatke, koordinatori su entiteti koji kontroliraju radnike, a krajnji rezultat odnosi se na tip ishoda koji je generiran pomoću prikupljenih resursa iz mnoštva. Prikaz ovih triju dimenzija dan je slikom 2.1. Iako je na



Slika 2.1. Radnici, koordinatori i izlaz u sustavima za prikupljanje podataka iz mnoštva, izrađeno po uzoru na [2]

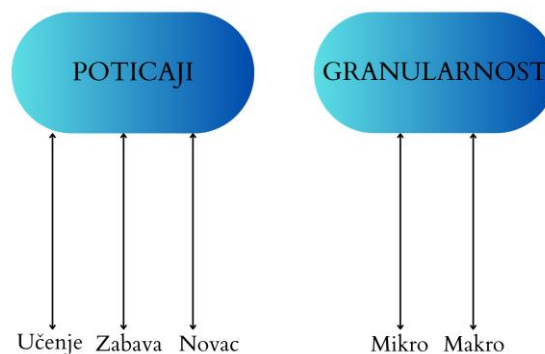
skici prikazan jedan, centralni koordinator, zadacima može upravljati veći broj distribuiranih koordinatora. Dvije su alternative vezane za koordinate i radnike – čovjek (engl. *human*) i računalo (engl. *computer*).

Kada su i koordinator i radnik ljudi, osoba nadzire izvršavanje zadataka i zapošljava druge ljude da rade dijelove ukupnog posla ili zadatke (engl. *tasks*). Ovaj tip prikupljanja podataka iz mnoštva korišten je dugi niz godina, primjerice, kada su u Sjedinjenim Američkim Državama prilikom traganja za osobom šerifi postavljali „*Wanted*“ postere u nadi da će osobe koje vide poster i koje imaju saznanja o osobi za kojom se traga šerifu javiti korisne informacije. U ovome primjeru, šerif je koordinator, a radnici su ljudi kojima će biti plaćena nagrada (engl. *bounty*) ukoliko ispune zadatak.

Kada je koordinator računalo, a radnici ljudi, računalo problem razdvaja u manje zadatke te ih daje ljudima na rješavanje. Primjer je označavanje fotografija gdje računalo raspodijeli slike tako da svaku sliku dobije jedan ili više radnika kojima je zadatak da ju označe. Ono zatim prikuplja odgovore, razrješava potencijalne probleme koji nastaju zbog nekonzistentnosti u označavanju i sl. Preostale dvije kombinacije nisu toliko relevantne u ovome području.

Treća dimenzija ovog skupa odnosi se na izlaz prikupljanja iz mnoštva. Podaci, softver, dokumenti i novac neke su od mogućnosti, a naglasak je na podacima, digitalnim zabilješkama koje predstavljaju ljude, proizvode, mape, slike i sl. Kod primjera označavanja slike, izlaz je skup slika s oznakom. Trening skup podataka specifičan je slučaj izlaza koji se koristi za neke algoritme strojnog učenja, a ukoliko se u taj kontekst stavi problem označavanja slika, podaci bi mogli biti korišteni za treniranje algoritma strojnog učenja kojemu je svrha da automatski označava slike. Čest primjer izlaza je i novac, kada je projekt, film ili tvrtka financirana od strane velike grupe ljudi (engl. *crowdfunding*), a jedan od najpoznatijih svjetskih primjera platformi ovoga tipa je online platforma *GoFundMe*.

Sljedeće stavke koje je potrebno definirati su poticaji (engl. *incentives*) i granularnost (engl. *granularity*) kao što je prikazano na slici 2.2. Pojam granularnosti odnosi se na veličinu zadataka koje radnik obavlja. U primjeru s označavanjem slika, granularnost pojedinog zadatka je mala, pa



Slika 2.2. Granularnost i poticaji, izrađeno po uzoru na [2]

ih se često naziva mikro zadacima (engl. *micro-tasks*), jer svaki radnik obavlja relativno malen dio ukupnog problema te je sukladno tome i plaćen. Nasuprot tomu, kada radnici razvijaju komponente velikog softvera, tada se radi o velikoj (engl. *large*) granularnosti pa se takvi zadaci često nazivaju makro zadaci (engl. *macro-tasks*).

Dimenzija poticaja opisuje ono što motivira radnika da sudjeluje. Primjeri poticaja uključuju novac, razne usluge, zabavu, učenje i slično. Većina platformi koristi novac kao poticaj, tj. radnici dobiju novac kao nagradu za sudjelovanje u obavljanju zadataka.

Važno je spomenuti i neke kompromise ili izazove o kojima će i kasnije biti riječi. Postoje 3 važna faktora koja su u konfliktu [3]:

- Nesigurnost (engl. *uncertainty*) se odnosi na činjenicu da ljudski radnici često rade pogreške, što rezultira mogućnošću da prikupljeni podaci ili konačan odgovor nisu ispravni.
- Trošak (engl. *cost*) se odnosi na monetarne troškove obavljanja svih zadataka, pojavljuje se zbog potrebe da ljudski radnici budu plaćeni za ono što odrade.
- Latencija (engl. *latency*) ili kašnjenje posebno je važno zbog mogućnosti da obavljanje zadataka potraje vrlo dugo, mnogo duže nego što računalu treba za obradu istih.

2.2.2. Načini prikupljanja podataka iz mnoštva

Podaci mogu biti prikupljeni iz mnoštva na dva načina [2]:

- Implicitno prikupljanje podataka izvršava se bez da je pojedinac toga svjestan, a odličan primjer je prikupljanje podataka iz mnoštva uporabom mobilnih platformi jer ljudi većinom nisu svjesni da putem brojnih senzora unutar njihovih mobilnih uređaja platforme mogu konstantno prikupljati podatke.
- Eksplicitno prikupljanje podataka je slučaj kada radnik dobije specifičan zahtjev za nekakvom uslugom.

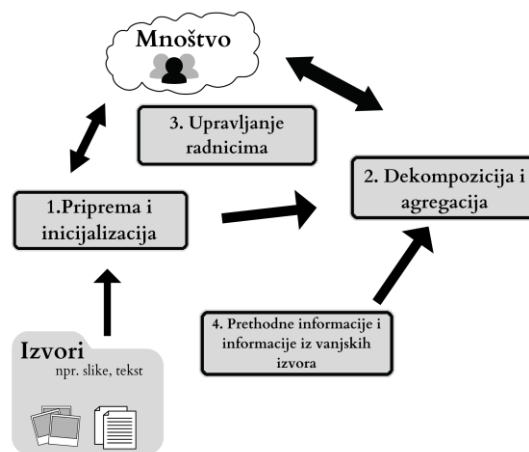
2.3. Dizajn i modeliranje sustava za prikupljanje podataka iz mnoštva

Dizajniranje sustava za prikupljanje podataka iz mnoštva definira se ovisno o potrebi, svrsi, platformi itd. Općeniti postupak se može grupirati u nekoliko koraka [2]:

1. Priprema i inicijalizacija (engl. *preparation and initialization*) – uključuje prepoznavanje svih dijelova potrebnih za realizaciju sustava.

2. Dekompozicija i agregacija (engl. *decomposition and aggregation*) – odabir tehnika za dekompoziciju ukupnog problema u manje zadatke.
3. Upravljanje radnicima (engl. *worker management*) – manipulacija odnosa s radnicima
4. Prethodne informacije i informacije iz vanjskih izvora (engl. *prior and external information*) – pripajanje vanjskih informacija rješenju.

Navedeni postupak slikovito je prikazan na slici 2.3. U nastavku će svaki od koraka biti pobliže opisan [2].



Slika 2.3. Dizajniranje sustava za prikupljanje podataka iz mnoštva, izrađeno po uzoru na [2]

2.3.1. Općeniti koraci pri izradi sustava za prikupljanje podataka iz mnoštva

Priprema i inicijalizacija

Priprema i inicijalizacija zahtijeva mnogo eksperimentiranja s raznim metodama prije nego se dođe do uspjeha. Prvi je korak prepoznati cilj ili problem (npr. sortirati skup slika, procijeniti razinu buke i sl.) i ograničenja kao što su budžet, vrijeme i točnost. Određivanje zadataka sljedeći je korak koji predstavlja specificiranje vrste zadataka zadanih radnicima. Zadaci predstavljaju male dijelove koji će biti korišteni za rješenje ukupnog problema, a najčešći tipovi uključuju zadatke s jednim ili višestrukim izborom, ocjenjivanje, grupiranje i označavanje [3].

Dekompozicija i agregacija

Dekompozicija problema predstavlja korak u kojemu se ukupan problem raspoređuje na skup zadataka. Najčešće se problem podijeli na nekoliko faza tako da svaka faza uključuje skup zadataka. Jedan od uobičajenih obrazaca je slijed eliminacijskih faza. Npr., ukoliko postoji određeni broj podataka među kojima treba pronaći maksimum (po nekom mjerilu koje zahtijeva čovjeka kao ocjenjivača), u prvoj se fazi podaci grupiraju u parove koji se daju radnicima na

usporedbu. Sljedeće faze zasnivaju se na eliminaciji gubitnika i ponavljanju prve faze kako bi se došlo do konačnog maksimuma.

Ocjena trenutnog stanja kombinira trenutno dostupne informacije, tj. odgovore mnoštva, procjene mogućnosti radnika i vanjske dokaze kako bi se moglo u svakom trenutku dati najbolju procjenu konačnog rješenja problema uz postotak sigurnosti.

Odabir sljedećeg zadatka komponenta je koja služi za odlučivanje o tome koje zadatke dati radnicima u sljedećoj fazi, uz poznavanje trenutnog stanja. Zadaci za koje je procjena sigurnosti odgovora mala, daju se radnicima na rješavanje kako bi se sigurnost povećala ili se kreiraju novi zadaci drugačije strukture s istim ciljem.

Optimizacija uključuje prilagodbu parametara sustava da se postigne najbolja točnost, najniža cijena ili najmanja latencija. Najčešće se rješava kreiranjem više planova izvršavanja kojima se procjene parametri te se na temelju toga odabere plan koji najbolje odgovara dostupnim resursima.

Upravljanje radnicima

Radnici često netočno rješavaju zadatke, a pojedinci čak do te mjere da daju više netočnih nego točnih odgovora. Posljedica je to djelovanja radnika kojima je cilj zarada bez velikog truda, a nazivaju se spameri (engl. *spammers*). Kako bi se riješili ovi problemi, sustavi implementiraju mehanizme za procjenu kvalitete radnika.

Izvlačenjem značajki radnika proučava se i ocjenjuje radnika na temelju njegovog obavljanja zadataka kako bi se procijenile značajke koje bi mogle biti korisne za prepoznavanje kvaliteta radnika. U nekim se slučajevima radnicima daju testni zadaci za koje su točni odgovori poznati kako bi se procijenile njihove sposobnosti ili se jednostavno procijeni koliko se često odgovori radnika razlikuju od odgovora ostalih radnika.

Prethodne informacije i informacije iz vanjskih izvora

Kao dodatak podacima iz mnoštva, moguće je da postoje informacije dostupne od prije ili podaci iz automatiziranih algoritama koji mogu biti od pomoći. Važno je iskoristiti ove informacije kao potvrdu rezultata dobivenih iz mnoštva te za lakši i kvalitetniji izbor zadataka koji će biti zadani mnoštvu. Čest je slučaj kada može postojati baza podataka vezana uz ciljani problem koju je, umjesto da ju mnoštvo kreira od početka, potrebno nadopuniti određenim poljima kako bi se riješio problem.

2.3.2. Ključni aspekti pri dizajniranju mobilnih sustava za prikupljanje resursa iz mnoštva

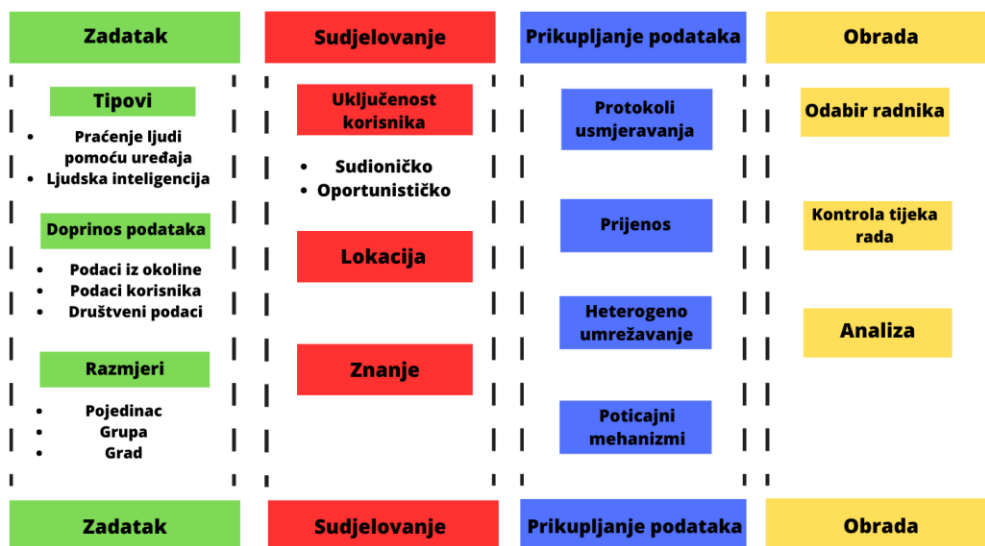
Mobilno prikupljanje podataka iz mnoštva danas je sve popularnije. Ono se odnosi na grupu ljudi koja dobrovoljno prikuplja i dijeli podatke koristeći mobilne uređaje [4]. Prikupljeni se podaci zatim obrađuju i šalju svima koji pokazuju interes. Obično se sustav za mobilno prikupljanje podataka iz mnoštva sastoji od platforme smještene na oblaku i pametnog uređaja. Mobilni uređaji imaju jedinstvene sposobnosti praćenja događaja u okolini jer su opremljeni nizom senzora koji povećavaju potencijal prikupljanja resursa.

Potrebna je infrastruktura koja podržava dizajn, implementaciju i automatsku obradu podataka dobivenih od ljudi. Sustav se uvelike oslanja na ljude, stoga treba imati na umu da će proces prikupljanja možda trajati duže jer je potrebno naći odgovorne radnike i prikupiti obavljene zadatke, što znači da su angažiranje radnika i tehnike motivacije ključni faktori kod razvoja ovih sustava.

Mnogo je različitih podjela sustava za mobilno prikupljanje resursa iz mnoštva. Jedna od podjela je prema svojstvima zadataka i količini ljudskog sudjelovanja [7] gdje se sustavi dijele na 3 vrste:

1. Mobilno računarstvo mnoštva (engl. *mobile crowd computing*) – sustav šalje računске zadatke mobilnim uređajima i prikuplja rezultate putem različitih mreža.
2. Mobilno prikupljanje podataka iz okoline (engl. *mobile crowdsensing*) – upotreba pametnih uređaja kao senzora kako bi se prikupili podaci o okruženju, infrastrukturi i korisnicima.
3. Prikupljanje podataka iz mnoštva uz pomoć ljudi (engl. *human-assisted crowdsourcing*) – cilj je iskoristiti ljudsku inteligenciju za rješavanje određenih zadataka.

Prema ključnim karakteristikama mobilnog prikupljanja podataka iz mnoštva, sustavi se mogu podijeliti na one gdje su u fokusu ljudi (engl. *people-centric*), u kojima se prikupljaju podaci vezani za korisnika kao što su fizička aktivnost i najčešće posjećena mjesta, te one gdje je u fokusu okolina (engl. *environment-centric*), kao što su aplikacije koje prate razinu zagađenja zraka, buke, stanje na cestama i slično [7]. Sustav je moguće analizirati kroz četiri aspekta koji uključuju zadatke (engl. *tasks*), sudjelovanja (engl. *participations*), prikupljanje podataka (engl. *data collection*) i obradu (engl. *processing*) [4]. Aspekti su sažeto prikazani na slici 2.4. i objašnjeni u nastavku.



Slika 2.4. 4 ključna aspekta pri dizajnu sustava za mobilno prikupljanje resursa iz mnoštva, izrađeno po uzoru na [4]

Zadatak (engl. *task*)

Tipovi zadataka koji se najčešće susreću uključuju praćenje ljudi pomoću uređaja (engl. *human-companion device tasks*) i korištenje ljudske inteligencije (engl. *human intelligence tasks*). Zadaci praćenja ljudi uz pomoć uređaja koriste mobilni uređaj kao senzor koji je izvor informacija o čovjeku i njegovoj okolini. Mobilni uređaji uključuju pametne telefone, pametne automobile i slično, a primjene su najčešće u području prikupljanja osobnih podataka kao što su podaci o zdravlju korisnika, sportska iskustva i slično. Zadaci koji zahtijevaju ljudsku inteligenciju koriste se u područjima gdje je problem teško rješiv za računalo, ali trivijalan za čovjeka. Dakako, rješenja ovih zadataka mogu biti subjektivna kao posljedica različitih mišljenja i iskustava.

Cilj sustava je da izdvoji korisne informacije iz velike količine podataka koje su generirali mobilni uređaji. Korisne informacije najčešće podrazumijevaju podatke korisnika (engl. *user data*), podatke iz okoline (engl. *context data*) i društvene podatke (engl. *social data*). Podaci korisnika podrazumijevaju osobnu okolinu (npr. česte lokacije, aktivnosti i sl.), zdravstvene podatke (npr. broj otkucaja srca, količina šećera u krvi i sl.) itd. Podaci iz okoline uključuju stanje stvari u krugu korisnika kao što su količina buke, gustoća prometa i sl. Društveni podaci odnose se na podatke koje korisnik kreira putem društvenih mreža, a omogućuju bolji uvid u korisnikove zanimacije, najdraža mjesta, tipove aktivnosti itd.

Razmjer (engl. *scale*) mnoštva u koje je radnik (korisnik mobilnog uređaja) uključen može biti malen ili velik, od pojedinca gdje svaki korisnik pojedinačno rješava zadatke ili pruža nekakvu uslugu, preko grupe u kojoj više pojedinaca čini čvrsto povezanu grupu čiji članovi dijele

informacije dobivene iz grupnih aktivnosti, do gradskog razmjera gdje su zadaci namijenjeni korisnicima mobilnih uređaja na razini grada.

Sudjelovanje (engl. *participation*)

Uključenost korisnika (engl. *user involvement*) zasniva se na njegovom sudjelovanju u prikupljanju podataka iz okoline, a može se podijeliti na oportunističko (engl. *opportunistic*) i sudioničko (engl. *participatory*). Kod oportunističkog sudjelovanja, korisnik je minimalno uključen, a mobilni uređaj donosi odluke na temelju podataka koje je prikupio (npr. prikupljanje podataka o lokaciji). Sudioničko sudjelovanje uključuje doprinos pojedinca prikupljanju podataka iz okoline (npr. slikanje fotografija, prijava kvarova i sl.) te često zahtijeva ljudsku inteligenciju.

Svijest o lokaciji (engl. *location-awareness*) područje je koje je zaživjelo zbog činjenice da svaki mobilni uređaj pruža lokacijske i mrežne usluge, što omogućuje svakoj osobi da glumi senzor koji prikuplja i dijeli različite tipove prostorno-vremenskih resursa. Osim lokacije, svaki pametni telefon danas je opremljen akcelerometrom, žiroskopom, kamerom i mikrofonom koji omogućuju proširenje tipova podataka koje je moguće prikupiti od korisnika.

Znanje (engl. *knowledge*) pojedinaca u mnoštvu ovisi o kompleksnosti i vještinama koje zadaci zahtijevaju. Neke aplikacije zahtijevaju pojedince koji nisu eksperti u području koje se istražuje (npr. označavanje slike), dok neke traže pojedince s točno određenim vještinama kako bi prikupili podatke (npr. dijagnosticanje neobičnih slučajeva u medicini).

Prikupljanje podataka (engl. *data collection*)

Tradicionalne metode upravljanja podacima u području nabave iz mnoštva uglavnom su u obzir uzimale samo centralnu ili klijent-poslužitelj komunikaciju. Mobilne *ad-hoc* mreže uključuju i *multi-hop* usmjeravanje i decentraliziranu obradu. Općenito, čvor šalje zadatke drugim čvorovima pomoću mobilne *peer-to-peer* komunikacije i prikuplja odgovore od njih u stvarnom vremenu. Prikupljanje podataka nije trivijalan postupak te se zbog toga razdvaja u četiri različita područja navedena u nastavku.

Protokoli usmjeravanja (engl. *routing protocols*) odnose se na strategiju propagacije zadataka među radnike iz mnoštva. Dvije su osnovne strategije, *point-to-point* koja se odnosi na komunikaciju između dva čvora na način da u jednom trenutku jedan čvor može komunicirati ili slati informacije jednom čvoru (npr. *Bluetooth* koji implementira *master-slave* odnos) i *multicast* koji se zasniva na povezanosti oblika jedan naprema više (npr. WiFi, WiFi Direct, LTE itd.).

Prijenos (engl. *transmission*) podrazumijeva transfer resursa koje su prikupili radnici prema poslužitelju u pozadini. Podaci mogu biti preneseni odmah pomoću mobilnih mreža ili s kašnjenjem (čekanje na dostupnu WLAN mrežu).

Heterogeno umrežavanje (engl. *heterogeneous networking*) daje mobilnoj nabavi iz mnoštva veliki potencijal olakšavanjem prijenosa. Danas dostupni mobilni uređaji pružaju mnoga bežična mrežna sučelja kao što su GSM/4G/5G, WiFi i *Bluetooth*. GSM/4G/5G i WiFi tehnologije pružaju stabilnu mrežnu konekciju ukoliko u okruženju postoji odgovarajuća infrastruktura koja ih omogućuje. S druge strane, *Bluetooth* i slične tehnologije pružaju konekciju između mobilnih uređaja čime formiraju samostalnu mrežu za distribuciju i dijeljenje podataka.

Dvije su osnovne metode prikupljanja i obrade podataka, centralizirana metoda kod koje se svi prikupljeni podaci šalju na poslužitelj u oblaku gdje se odvija obrada, te decentralizirana, kod koje se obrada i komunikacija odrađuje lokalno na mobilnim uređajima. U zadnje su vrijeme oblaci računala (engl. *cloud computing*) često korištena tehnika za prikupljanje i analizu podataka iz mnoštva, što nekada može izazvati veliko kašnjenje te time postati usko grlo sustava kojima je presudan rad u stvarnom vremenu. Jedno od mogućih rješenja je računalstvo na rubovima (engl. *edge computing*) kod kojega se najosjetljiviji podaci prikupljaju i analiziraju blizu mjesta gdje su generirani, čime se uvelike smanjuje kašnjenje.

Poticajni mehanizmi (engl. *incentive mechanisms*) su aspekt koji ima ulogu poticanja sudjelovanja radnika u mobilnom prikupljanju resursa iz mnoštva. Poznato je da motivacije za rješavanje zadataka mogu biti razne, poput užitka i poticaja zajednice, pa sve do novca i brzine njegove isplate. Ispostavlja se da je novac, kao i kod tradicionalnih platformi, najveći motivator radnicima za sudjelovanje u mobilnom prikupljanju podataka.

Obrada (engl. *processing*)

Odabir radnika (engl. *worker selection*) odnosi se na obradu i podjelu zadataka na takav način da svaki skup zadataka bude dodijeljen onim pojedincima koji su ga u stanju riješiti. Ukoliko se npr. analizira platforma koja se zasniva na tome da radnici moraju biti u krugu određene lokacije kako bi bili u stanju riješiti zadatak (engl. *spatial crowdsourcing*), važno je da sustav te platforme potencijalni posao na određenom području pokazuje samo radnicima koji se nalaze u blizini (npr. *Glovo*).

Kontrola tijeka rada (engl. *workflow controls*) podrazumijeva postupak planiranja i izvršavanja zadataka koje podgrupa mnoštva može efikasno riješiti. Ovakva okruženja u stanju su rastaviti

veći zadatak na manje zadatke i kasnije objediniti rješenja manjih zadataka u jedinstveno ukupno rješenje, a koriste se kao temelj za algoritme prikupljanja resursa iz mnoštva.

Analiza (engl. *analysis*) igra ključnu ulogu u pružanju povratne informacije podnositelju zahtjeva. Ona omogućuje kvalitetnije rezultate i lakši odabir najboljih rezultata iz velikog skupa rješenja koje je mnoštvo pružilo. Najčešći scenarij je taj da se zadaci pošalju radnicima, koji potom rješenja šalju na poslužitelj u oblaku. Zatim se odvija analiza podataka koja može uključivati neke jednostavne metode poput odabira najčešće danog odgovora ili računanja prosjeka ukoliko je u pitanju problem ocjenjivanja. Nasuprot jednostavnima, postoje i naprednije metode obrade podataka poput rudarenja podataka, strojnog učenja i ostalih prilagođenih algoritama poput maksimizacije očekivanja (engl. *expectation maximization*).

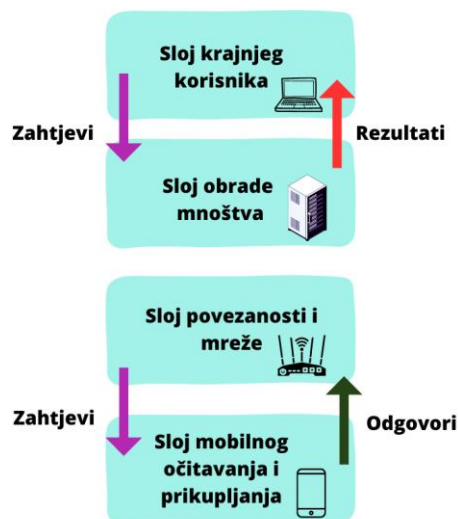
2.3.3. Arhitekture sustava za mobilno prikupljanje podataka iz mnoštva

Nužno je pojasniti dvije generalizirane arhitekture mobilnih sustava za prikupljanje resursa iz mnoštva, zajedno sa svojim funkcionalnostima i aspektima dizajna [4]. Centralizirana arhitektura odnosi se na postupak u kojemu se svi prikupljeni podaci obrađuju na poslužitelju u oblaku, dok decentralizirana arhitektura inzistira na obavljanju izračuna i komunikacije lokalno na mobilnim uređajima.

Centralizirana arhitektura mobilnog prikupljanja podataka iz mnoštva

Sustavi za mobilno prikupljanje podataka iz mnoštva obično su implementirani uporabom komunikacije putem Interneta i klijent-poslužitelj komunikacije koje omogućuju pristup putem pametnih telefona ili radnih stanica. Ovaj tip arhitekture zasniva se na klijent-poslužitelj modelu, gdje je poslužitelj komponenta koja pruža servise, funkcionalnosti i resurse za jednog ili više klijenata koji zatraže zahtjev za nekom uslugom. Svi izračuni obavljaju se na centralnom poslužitelju, a svi se resursi drže u jedinstvenom podatkovnom centru. Krajnji korisnici šalju zahtjeve centru, a rezultat dobiju od svoje usluge u oblaku. Na slici 2.5. ilustrirana je centralizirana arhitektura za mobilno prikupljanje resursa iz mnoštva koja se sastoji od 4 osnovna sloja: sloj mobilnog očitavanja i prikupljanja (engl. *mobile sensing and gathering layer*), sloj povezanosti i mreže (engl. *connectivity and network layer*), sloj obrade mnoštva (engl. *crowd processing layer*) i sloj krajnjeg korisnika (engl. *end-user layer*) [4].

Sloj mobilnog očitavanja i prikupljanja (engl. *mobile sensing and gathering layer*) ima ulogu generiranja i slanja prikupljenih podataka iz mnoštva ili odgovora centralnom poslužitelju putem bežičnih mreža. Zbog velikog broja različitih uređaja koji podatke šalju u različitim oblicima i enormnim količinama, javlja se potreba za standardom koji će riješiti problem heterogenih izvora



Slika 2.5. Centralizirana arhitektura nabave iz mnoštva, izrađeno po uzoru na [4]

podataka. Jedno od rješenja je uvođenje dodatne platforme koja olakšava suradnju i razmjenu informacija (npr. apstrakcija hardvera i kreiranje API-ja (engl. *Application Programming Interface*) koji će pružiti potrebne funkcionalnosti).

Sloj povezanosti i mreže (engl. *connectivity and network layer*) osigurava mrežnu povezanost sustavu za mobilno prikupljanje resursa iz mnoštva. Mehanizam koji ovaj sloj pruža je otkrivanje uređaja koji je potreban za pristup dodatnim informacijama o okolini. Izazovi vezani uz ovaj sloj uključuju očuvanje privatnosti lokacije radnika i informacija koje šalje mrežom, konzumiranje hardverskih resursa (memorija, baterija) koje nadmašuju učinak motivacijskih mehanizama koje sustav pruža i sl.

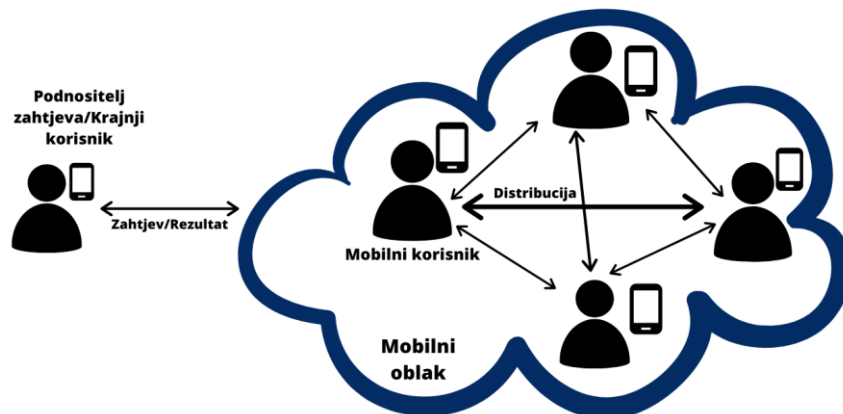
Sloj obrade mnoštva (engl. *crowd processing layer*) zadužen je za analizu, obradu i pohranu zadataka i senzorskih podataka. Oblak računala (engl. *cloud computing*) trenutno je najkorištenija infrastruktura za realizaciju ovoga sloja u centraliziranim arhitekturama. Njome se postiže dulje trajanje baterije, povećanje kapaciteta pohrane, poboljšana snaga obrade i pouzdanost.

Sloj krajnjeg korisnika (engl. *end-user layer*) odnosi se na podnositelje zahtjeva koji kupuju ili iznajmljuju usluge sustava za mobilno prikupljanje resursa iz mnoštva. Krajnji korisnik može biti pojedinac, organizacija ili pametni uređaj (npr. pametni automobil) koji šalje zahtjev za uslugom centralnom poslužitelju koji mu vraća rezultat. Krajnji korisnici imaju mogućnost kreiranja zadataka namijenjenih mnoštvu i pristupa rezultatima kroz korisničko sučelje (engl. *user interface*, UI).

Decentralizirana arhitektura mobilnog prikupljanja podataka iz mnoštva

Nasuprot centraliziranoj, u decentraliziranim arhitekturama svi izračuni i komunikacija odrađuju se lokalno na mobilnim uređajima. Svaki čvor u takvom sustavu (mobilni uređaj) jednako je uključen u doprinos ukupnom rezultatu i može biti smješten na bilo kojoj lokaciji s time da geografska lokacija uređaja postaje relevantna za proces izračuna. To znači da svaki čvor ima mogućnost obrade i distribucije svojih informacija bez oslanjanja na bilo kakav centralni autoritet. Uloga sustava je da propagira zahtjev ili zadatak velikom broju radnika s mobilnim uređajima na kojima se obavljaju svi izračuni te koji imaju ovlasti da u potpunosti manipuliraju i distribuiraju zahtjev ostalim čvorovima putem raznih komunikacijskih kanala. Zbog snage današnjih procesora i velike memorije, mobilni uređaji pogodni su i za kompleksnije izračune te međusobnu razmjenu s niskim kašnjenjem i visokom brzinom. Zbog toga je decentralizirana arhitektura pogodna za korištenje kada je potrebno realizirati mobilni sustav za prikupljanje podataka iz mnoštva koji će biti korišten u scenarijima katastrofe i hitnim slučajevima (npr. potresi, traženje nestale osobe itd.) [8].

Na slici 2.6. prikazana je decentralizirana arhitektura u kojoj su mobilni uređaji radnici, vrlo slično *peer-to-peer* mreži. Svaki čvor ima mogućnost komunikacije s ostalima koji se nalaze u njegovoj blizini korištenjem bežičnih sučelja, što omogućuje zajedničko rješavanje problema koji je



Slika 2.6. Decentralizirana arhitektura nabave iz mnoštva, izrađeno po uzoru na [4]

proizišao iz zahtjeva nekog drugog mobilnog uređaja. Prikupljeni podaci i rješenja zadataka pohranjuju se lokalno na uređaju u bazu podataka. Decentralizirana arhitektura zapravo je distribuirana verzija centralizirane, no za razliku od nje decentralizirana uključuje kompleksnije elemente za upravljanje zadacima i njihovu propagaciju, odabir radnika te obradu zadataka u *peer-to-peer* okruženju. Tako se izbjegava spajanje na centralni poslužitelj koji je usko grlo

centralizirane arhitekture, no uz puno kompliciraniju strukturu suradnje među čvorovima. U decentraliziranoj arhitekturi nema strane koja kontrolira cijeli proces prikupljanja i izračuna, a kako bi sustav uopće funkcionirao potrebno je da čvorovi međusobno surađuju u distribuciji i rješavanju zadataka.

2.4. Izazovi vezani uz prikupljanje podataka

U nastavku su opisani neki od najvažnijih izazova poput kontrole kvalitete, troška, latencije i privatnosti [3], [9], s posebnim naglaskom na izazove specifične za mobilne sustave koji donose novu dimenziju prepreka vezanih uz njihovu ideju i strukturu [4], [7], [10].

2.4.1. Kontrola kvalitete

Kontrola kvalitete (engl. *quality control*) naglasak stavlja na nadzor rezultata dobivenih iz mnoštva [3]. Kvaliteta rezultata od presudne je važnosti za podnositelja zahtjeva jer doprinosi relevantnosti i iskoristivosti prikupljenih rezultata. Ona omogućuje detekciju zlonamjernih sudionika koji ne pružaju korisne odgovore i njihovu eliminaciju iz kruga radnika te zanemarivanje odgovora koje pružaju. Mnoge su tehnike koji se koriste u ovu svrhu što rezultira implementacijom složenijih okvira za fleksibilno upravljanje kvalitetom podataka iz mnoštva [11], a primjeri uključuju vjerojatnost radnika (engl. *worker probability*) i matricu konfuzije (engl. *confusion matrix*).

2.4.2. Kontrola troška

Iako su platforme za prikupljanje podataka iz mnoštva danas vrlo dostupne te iako pružaju puno jeftiniji način na koji će se od čovjeka tražiti da obavi neki posao, i dalje je trošak vrlo velik kada je količina zadataka i posla velika. Nekoliko je klasa korisnih tehnika koje se koriste za rješavanje ovoga problema, a najbitnije su navedene u nastavku [3].

Obrezivanje (engl. *pruning*) koristi računalne algoritme za predobradu svih zadataka, odbacujući zadatke za čije rješavanje nisu potrebni ljudi.

Odbitak odgovora (engl. *answer deduction*) zasniva se na činjenici da su zadaci međusobno povezani. Prikupljanjem odgovora za dio zadataka, moguće je iz prikupljenih informacija dati odgovore na neke druge zadatke. Ove tehnike sprječavaju redundanciju, no u situacijama kada ljudi rade greške može se dogoditi njena propagacija na veliki broj zadataka.

Tehnike uzorkovanja (engl. *sampling*) koriste mnoštvo za rješavanje uzorka zadataka, a zatim podatke prikupljene na uzorku primjene na cijeli skup podataka. Nisu primjenjive u svim

situacijama, tj. kada je problem takav da rješenja uzorka ne smiju biti primijenjena na cijeli skup jer bi došlo do velike pogreške (npr. traženje maksimalne vrijednosti).

2.4.3. Kontrola latencije

Kontrola latencije (engl. *latency control*) izazov je vezan uz vrijeme potrebno da se zadaci obave. Jednostavan način rješenja ovog problema je povećati trošak. Stoga se u praksi često događa dinamično mijenjanje cijena zadataka, ovisno o broju radnika koji prihvaćaju riješiti zadatke po danoj cijeni i vremenskim rokovima. Latencija je u mobilnoj nabavi iz mnoštva manje izražena, poglavito zbog velikog broja mobilnih uređaja, prirode zadataka i činjenice da ih ljudi nose svuda sa sobom [4].

2.4.4. Izazovi vezani uz mobilno prikupljanje iz mnoštva

Mobilni kontekst i prostorna nabava iz mnoštva

Mobilni uređaji nose važne informacije o korisniku, što omogućuje aplikacijama prilagodbu na njegove zahtjeve. Mobilni kontekst (engl. *mobile context*) uključuje svijest o uređaju i njegovoj okolini, a primjeri su lokacija, vrijeme, tip kretanja korisnika (npr. šetnja, trčanje...), korisnički profil (npr. godine, spol...), socijalni status (npr. korisnik priča na telefonu, na sastanku je...) ili pojava nekih događaja [4]. Raspoznavanje konteksta postaje izazov jer je potrebna visoka razina točnosti kako bi se olakšala dodjela zadataka (koji imaju nekakav zahtjev vezan uz kontekst mobilnog uređaja) adekvatnim radnicima. Primjer su aplikacije koje kao preduvjet za rješavanje zadataka zahtijevaju specifično stanje mobilnog konteksta, tj. svakom se radniku nude samo oni zadaci koje je u mogućnosti riješiti s obzirom na kontekst njegovog mobilnog uređaja. Mobilni kontekst ključan je za funkcionalnost skoro svake mobilne aplikacije za nabavu iz mnoštva, no s njegovim poboljšanjem, pogoršava se problem privatnosti, posebno kod centraliziranih pristupa prikupljanju podataka iz mnoštva gdje središnji ustupitelj mora poznavati kontekst uređaja radnika. Decentraliziran pristup ovaj problem izbjegava, no po cijenu lošije efikasnosti.

Potrošnja energije

Potrošnja energije mobilnih uređaja kao što su pametni telefoni postala je zabrinjavajuća stavka [4]. Senzori ugrađeni u mobilne uređaje veliki su potrošači energije. Iako se u zadnjih nekoliko godina prosječni kapacitet baterije povećao, to i dalje nije dovoljno dobro da pokrije širok spektar različitih aplikacija. Mobilne mreže također su veliki potrošači baterije, među kojima prednjače 5G i 4G koji su vrlo energetski zahtjevni u usporedbi s Wi-Fi-em, posebno kada se odvija prijenos

velike količine podataka. Rješenja ovoga problema zasnivaju se na prikupljanju i slanju manjih količina podataka te izvršavanju potrebnih operacija u pozadini.

Dodjela zadataka

Dodjela zadataka kod mobilne nabave slično kao i u tradicionalnim sustavima podrazumijeva proces dodjele zadataka podskupu mobilnih korisnika [4]. Problemi koji se javljaju vezani su uz dinamiku pridruživanja i odlaska radnika, što dovodi do stalnog mijenjanja mrežne topologije. Osim toga, može doći do nedostupnosti podataka ili nestanka signala. Stoga se često implementiraju mreže koje su vođene pohrani-nosi-proslijedi paradigmom (engl. *store-carry-forward*), tj. dopuštaju komunikaciju i kada je mrežni put nedostupan ili nepouzdan, a nazivaju se mreže tolerantne na kašnjenje (engl. *Delay Tolerant Networks*).

Očuvanje privatnosti

Kako bi sustavi za prikupljanje podataka iz mnoštva bili globalno prihvaćeni, potrebno je da zadovoljavaju uvjete privatnosti koji uključuju i pružatelja usluge i radnike. Zadaci koje mnoštvo rješava mogu sadržavati osjetljive informacije, osobne interese, često posjećena mjesta, prostorno-vremenske informacije o radniku i slično [7]. Stoga je zaštita takvih podataka uvjet bez kojega većina radnika ne bi pristala na sudjelovanje u nabavi. Informacije koje se prenose mrežom moguće je zaštititi uporabom kriptografije, čime se sprječava mogućnost da ih zlonamjerni sudionici iskoriste. Privatnost podnositelja zahtjeva i zadataka koje objavljuje također je bitna stavka, jer zadaci koje on daje na rješavanje mnoštvu mogu odavati privatne informacije o njemu. Kako bi se osjetljivi podaci zaštitili, provode se tehnike anonimizacije podataka. Anonimizacija podataka zapravo je pretvorba osobnih podataka u anonimizirani oblik [12]. U anonimiziranom obliku podataka uklonjeni su svi osobni identifikatori koji su na neki način mogli dovesti do osobe, a povratak u originalni oblik podataka nije moguć. Navedeno se ostvaruje uklanjanjem dijelova podataka i primjenom kriptografskih metoda, s ciljem da podaci ostanu iskoristivi. Kako bi se provjerio stupanj anonimizacije, koriste se mjere kao što su K-anonimnost (engl. *k-anonymity*), L-raznolikost (engl. *l-diversity*) i T-bliskost (engl. *t-closeness*).

Heterogeni formati podataka

Porastom broja i vrsta mobilnih uređaja, raste i broj različitih tipova i formata podataka koji mogu biti prikupljeni pomoću takvih uređaja [4]. Osim toga, količine podataka su enormne, što dovodi do teže manipulacije i obrade. Raznolikost formata može predstavljati problem pri obradi zbog nemogućnosti da se sustav prilagodi na sve moguće formate različitih tipova podataka. Stoga se obrada nekada provodi i lokalno u svrhu pretvorbe podataka u oblik potreban sustavu, što dovodi

do povećane potrošnje energije. Prikupljanje podataka može biti izvršeno pomoću različitih mreža (4G, 5G, WiFi, *Bluetooth* i sl.), a izazov je osigurati da podaci prikupljeni na razne načine budu adekvatno obrađeni.

2.5. Primjeri primjena prikupljanja podataka

Najčešća područja primjene uključuju praćenje događaja u okolini uređaja, pomoć ljudima koje je zadesila katastrofa, praćenje zdravlja korisnika, rješenja raznih problema na području grada, socijalnu problematiku i mnoge druge. Neki primjeri primjene bave se i problematikom vezanom za kvalitetu prikupljenih podataka, uspoređujući kvalitetu podataka koje su prikupili eksperti u području prikupljanja i ostalih [13] te dolaze do zaključka da bitnija razlika ne postoji.

2.5.1. ImageNet

ImageNet danas je poznat kao jedna od najvećih baza podataka koja sadrži označene slike na svijetu [14]. Za stvaranje *ImageNet-a* zaslužna je podatkovna znanstvenica Fei-Fei Li koja je tvrdila da bi velike količine podataka iz stvarnoga svijeta mogle donijeti algoritmima veću efikasnost:

“Promjena paradigme razmišljanja ImageNet-a je ta da iako mnogo ljudi većinu pažnje pridaje modelu, obratimo pozornost na podatke. Podaci će potpuno promijeniti naše razmišljanje o modelima.”

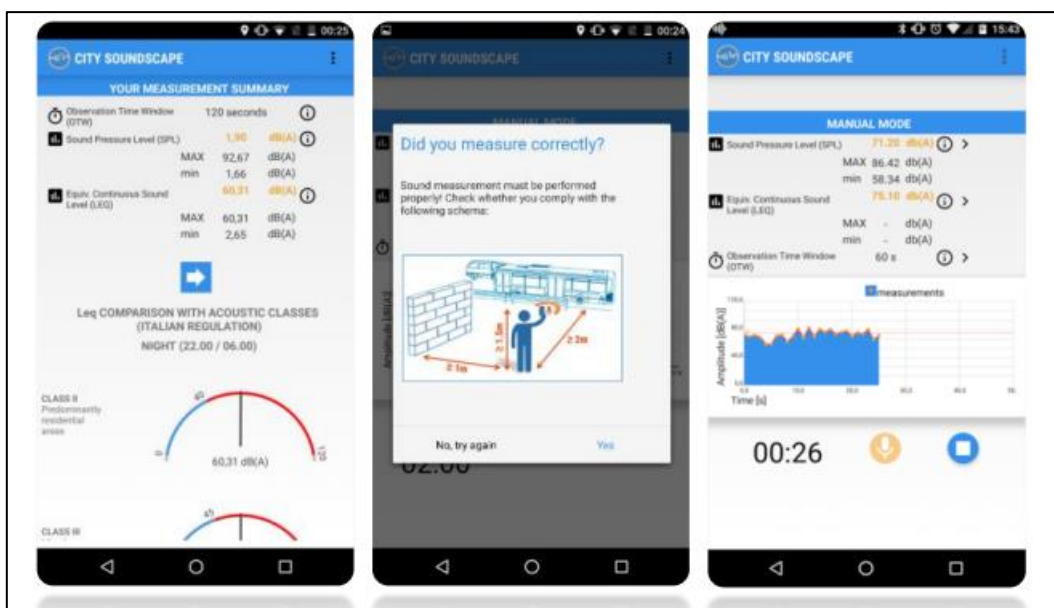
Fei-Fei Li [15]

U srpnju 2008., *ImageNet* je imao 0 slika. Već u prosincu iste godine, imao je više od 3 milijuna označenih slika u više od 6000 kategorija. Slike su prikupljene pomoću različitih tražilica za slike koristeći upite formirane pomoću sinonima iz *WordNet-a*. Danas ova baza slika broji preko 14 milijuna označenih slika i 20 000 različitih kategorija, a najzaslužnije za njeno postojanje je upravo mnoštvo. Označavanje slika odvijalo se pomoću jedne od najvećih platformi za nabavu iz mnoštva, *Amazon Mechanical Turk-a*. Prikupljene slike od velike su koristi u istraživanjima vezanima uz algoritme strojnog učenja i duboke neuronske mreže gdje ih se koristi kao podatke za trening i testiranje, a također su besplatno dostupne istraživačima u nekomercijalne svrhe.

2.5.2. City Soundscape

U gradu Brindisi 2016. godine održan je pilot projekt koristeći aplikaciju *City Soundscape* objavljenu na *Google Play Store-u* u kojemu su građani slali informacije o razini buke na određenim lokacijama unutar grada [16]. Projekt je trajao 4 mjeseca, od veljače do svibnja, a

prikupljeno je ukupno 16 000 mjerenja, od kojih je iskoristivih bilo 13 500. U gradu su dostupni gotovo svi oblici prijevoza, što je bilo povoljno za projekt kojemu je glavni cilj bio istražiti buku koju izaziva promet. Razina buke može bitno narušiti ne samo kvalitetu života stanovnicima gradova u kojima je ona na visokoj razini, nego i turistima koji su vrlo česta pojava diljem Italije. Kao glavni cilj projekta ističe se osvještavanje gradskih i državnih vlasti o problemima koje buka izaziva, te njihova reakcija kako bi se taj efekt smanjio. U projektu su mogli sudjelovati svi stanovnici Brindisia, s posebnom pozornošću posvećenom gradskim službenicima, studentima, učenicima i profesorima, za koje je naglašena potreba za edukacijom u području korištenja mobilnih uređaja i aplikacija za praćenje zvuka u okolini. Aplikacija je imala sučelje za objavu informacija o buci, toplinsku kartu (engl. *heat map*) razina buke te prikaz dodatnih informacija vezanih uz potencijalna rješenja za smanjenje razine buke. Također, dodana je i opcija koja pruža uvid u utjecaj primjene različitih rješenja za smanjenje buke u obliku nove toplinske karte. Na slici 2.7. prikazane su snimke zaslona aplikacije *City Soundscape* s prethodno opisanim funkcionalnostima. Prvo bi se mikrofonom pametnog uređaja snimio audio zapis, zatim bi se korisnika pitalo o pojedinostima vezanima za izvor zvuka, lokaciju (unutar ili izvan građevine) i udaljenost od izvora te bi se nakon toga podaci zajedno sa lokacijom korisnika prikupljenom pomoću GPS-a poslali na poslužitelj uporabom mobilnih mreža ili WiFi-a. Alternativno se umjesto mikrofona mogla subjektivno specificirati razina buke vrijednošću između 1 i 10. Obradom prikupljenih podataka popunjavala se toplinska mapa, a detaljniji uvid i analiza razina buke prezentirani su gradskim vlastima. Cijeli projekt dao je obećavajuće rezultate i pobudio svijest o

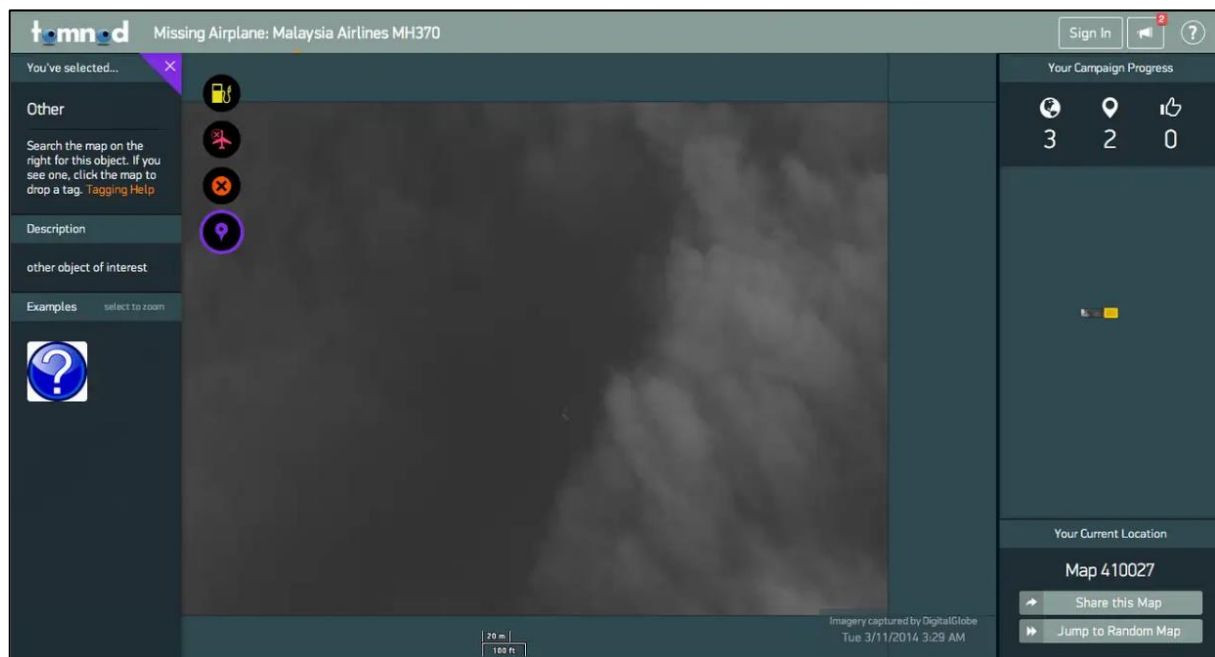


Slika 2.7. Korisničko sučelje aplikacije *City Soundscape*

ovome problemu, što je rezultiralo ulaganjima u smanjenje buke i dodatnim promicanjem ove teme u različitim područjima.

2.5.3. Let Malaysia Airlinesa 370

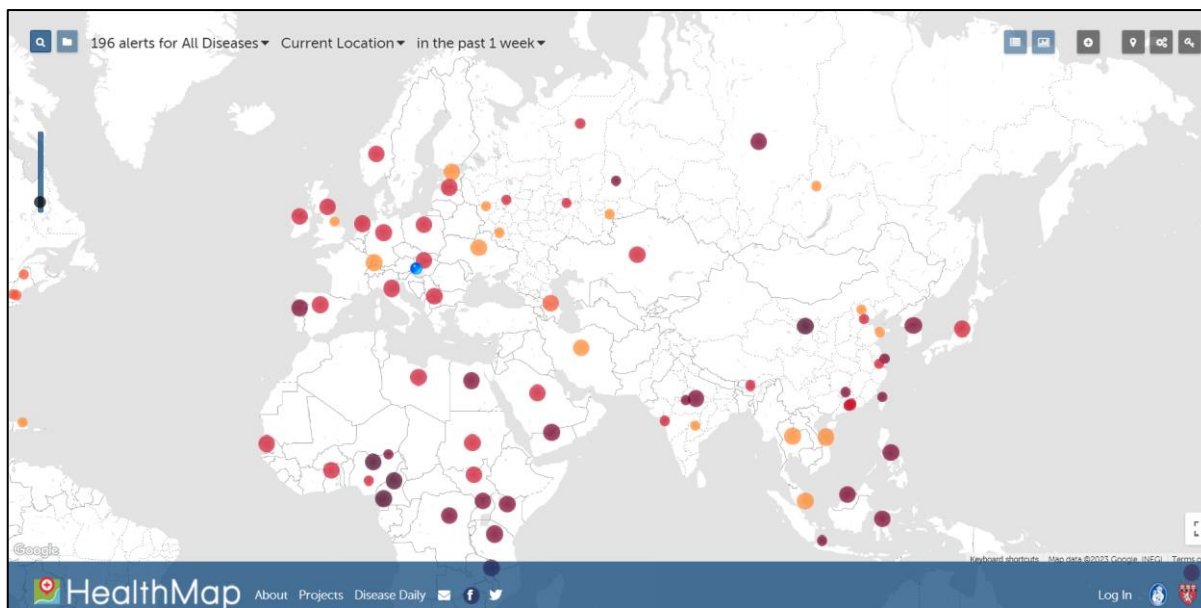
Let Malaysia Airlinesa 370 (MH370) internacionalni je putnički let koji je nestao 8. ožujka 2014. godine na putu iz Kuala Lumpura u Beijing [17]. Kompanija *DigitalGlobe* pomoću svojih 5 satelita zabilježila je slike visoke razlučivosti područja u vrijeme nesreće na kojemu se pretpostavljalo da je avion pao te pomoću njih pokrenula potragu uz pomoć mnoštva. Slike su bile dostupne javnosti na web stranici *Tomnod* (tomnod.com), a svatko zainteresiran mogao je početi pretraživati slike na kojima je svaki piksel predstavljao pola metra površine oceana, označavajući područja koja su im se činila sumnjivima. *DigitalGlobe* koristio je računalni algoritam koji je bilježio područja koja su češće označena nego druga te tako usmjeravao stvarnu potragu za avionom. Potrazi se pridružilo više od 8 milijuna volontera, a od nje se putem slika odustalo krajem 2014. godine. Iako su neki dijelovi aviona isplivali iz Indijskog oceana 2015. i 2016. godine, potraga za avionom nikada nije uspjela. Slika 2.8. prikazuje sučelje iz kojega su korisnici označavali sumnjive lokacije na ponuđenoj slici.



Slika 2.8. Izgled sučelja za označavanje područja mogućeg pada MH370 s web stranice *Tomnod* (tomnod.com)

2.5.4. HealthMap

HealthMap koristi podatke iz različitih izvora (nastalih prikupljanjem iz mnoštva) kako bi se stvorila jedinstvena zdravstvena karta [18] prikazana na slici 2.9. Inicijativu su pokrenuli John Brownstein i Clark Freifeld 2006. godine. *HealthMap* prikuplja podatke iz različitih besplatno

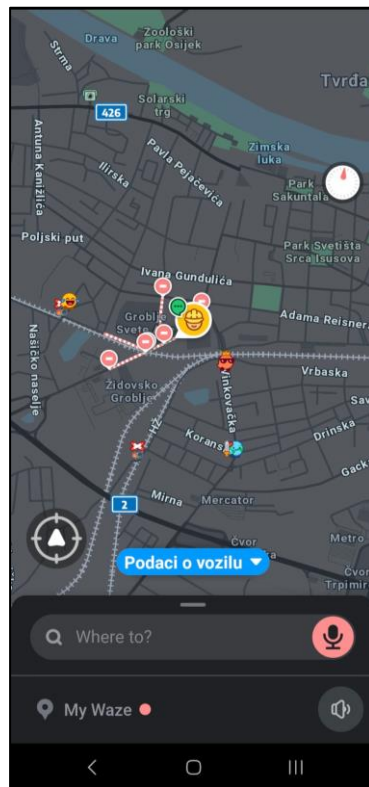


Slika 2.9. HealthMap (<https://www.healthmap.org/en/>)

dostupnih izvora kao što su *ProMED-mail*, *Eurosurveillance* i *Wildlife Disease Information Node* kako bi se izradila karta (u suradnji s *Google-om*) s prikazom globalne prisutnosti zaraznih bolesti. Kartu koriste velike zdravstvene organizacije poput Svjetske zdravstvene organizacije (WHO) i Europskog centra za prevenciju i kontrolu bolesti, primarno u svrhu rane detekcije potencijalne pandemije i zaraznih bolesti prisutnih na određenom području. Njena efikasnost posebno se istaknula kada je vrlo rano detektirala mjesta pojave COVID-a i Ebole te tako potaknula preventivne mjere diljem svijeta.

2.5.5. Waze

Waze je besplatna aplikacija dostupna na pametnim uređajima koji podržavaju GPS [19]. Razvoj softvera započeo je još 2006. godine pod nazivom *FreeMap Israel*, a u lipnju 2013. *Google* je kupio kompaniju. Aplikacija prikuplja kartografske podatke, duljine putovanja i informacije o transportu od svojih korisnika te ih prenosi na *Waze* poslužitelj besplatno. Korisnici mogu prijaviti nesreće, gužve, policiju na cesti i slično, a u mogućnosti su i pomoću alata za online uređivanje karte ažurirati ceste, brojeve kuća itd. Zaslona aplikacije s primjerom izgleda *Waze* karte prikazan je na slici 2.10. Poslužitelju se šalju anonimizirani podaci kao što su brzina i lokacija korisnika kako bi se cijeli sustav dodatno poboljšao. Na osnovu ovih informacija, *Waze* je u mogućnosti pružiti ažurno stanje na cestama u stvarnome vremenu, kao i informacije poput najjeftinije benzinske pumpe u blizini korisnika koje su također ažurne s podacima dobivenima iz mnoštva, uz prezentaciju traženih informacija pomoću *Waze-ove* karte, dizajniranu specifično za *Waze-ov*



Slika 2.10. Primjer Waze karte u gradu Osijeku

algoritam usmjeravanja. Broj korisnika aplikacije neprestano raste, a posljednje brojke kreću se oko 150 milijuna korisnika.

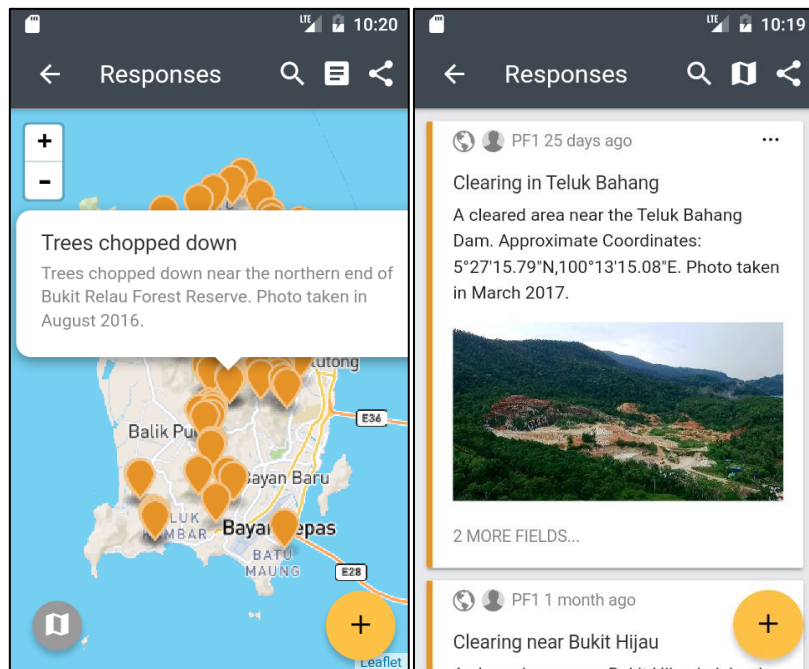
2.6. Postojeće mobilne platforme i alati za prikupljanje podataka

Postoje mnoge mobilne platforme i alati koji korisnicima omogućuju prikupljanje podataka iz mnoštva. Često pružaju mogućnost kreiranja proizvoljnih zadataka ili njihovo rješavanje uzimajući u obzir ranije opisane detalje.

2.6.1. Ushahidi

Platforma *Ushahidi* u početku je bila web stranica nastala kao posljedica nereda u Keniji potaknutih izborima 2007. godine, na kojoj je javnost imala priliku u obliku *Google-ove* karte vidjeti mjesta nasilja koja su prijavili svjedoci putem elektroničke pošte i SMS poruka [20]. Spomenuta web stranica vremenom je prerasla u ekosustav softvera i alata izrađenih kako bi pomogli u raznim krizama i katastrofalnim scenarijima. Alati su realizirani uz pomoć *Python-a*, *PHP-a*, *Javascript-a*, *Ionic-a*, *Ansible-a* i mnogih drugih tehnologija, uz naglasak na prednosti korištenja otvorenog koda (engl. *Open Source*). Platforma prikuplja podatke iz raznih izvora s ciljem poticanja komunikacije i interakcije korisnika, omogućuje organizaciju podataka koristeći

kategorije, filtriranje i pretragu, kontrolu pristupa podacima uz pomoć uloga i dozvola, vizualizaciju podataka i njihovu analizu pomoću karti i dijagrama te njihov izvoz u raznim formatima. Primjer prikupljenog odgovora i vizualizacije na karti uz pomoć mobilne aplikacije *Ushahidi* dan je na slici 2.11. *Ushahidi* nudi podršku u dizajnu upita, trening tima, tehničku



Slika 2.11. Primjer prikupljenog odgovora i vizualizacije u mobilnoj aplikaciji *Ushahidi*

podršku, asistenciju s metrikama i analizom, pomoć u strategiji i dizajn dodatnih opcija ukoliko su potrebne. Pružena je opcija najma dijela platforme na mjesečnoj bazi i besplatno korištenje manjim organizacijama. Naglašava se važnost razumijevanja problema od interesa i podataka koji se žele prikupiti uz pomoć platforme. Potrebno je dizajnirati svoju metodu prikupljanja podataka i ciljeve te kreirati projekt na platformi koja će služiti kao alat mnoštvu za rješavanje zadataka vezanih uz objavljeni projekt. Svaki korisnik platforme ima svoj račun, a platforma je ponuđena na oblaku (bez instalacije), u obliku istoimene mobilne aplikacije, API-ja i ručne instalacije na vlastitom poslužitelju (uz veću fleksibilnost), što čini *Ushahidi* ne samo mobilnom, već platformom dostupnom na skoro svim uređajima.

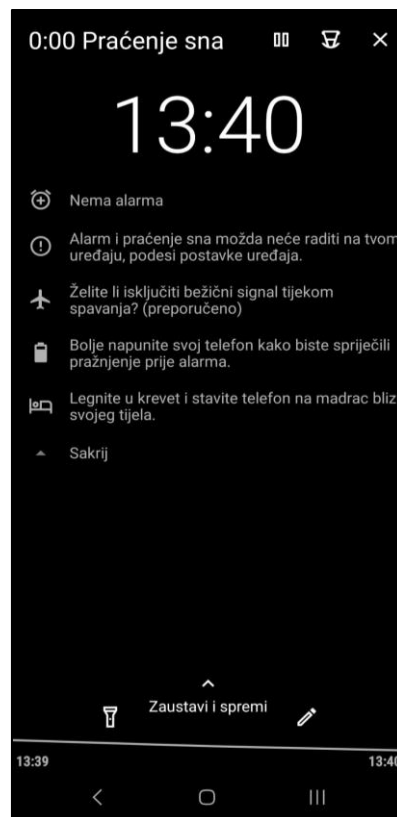
2.6.2. Crowdsorce by Google

Web i mobilna aplikacija *Crowdsorce by Google* izrađena 2016. godine prikuplja podatke iz mnoštva rješavanjem zadataka svrstanih u kategorije kako bi se potpomogao razvoj područja strojnog učenja i umjetne inteligencije [21]. Zadaci dani radnicima vrlo su kratki, a uključuju provjeru prijevoda, identifikaciju objekta na slici, prepoznavanje rukom pisanih znakova itd. Rješavanje je dobrovoljno, tj. isplata novčanih sredstava ne postoji, no postoji poseban oblik

nagrađivanja u obliku razina i znački te rangiranja korisnika prema doprinosu. Prikupljene podatke *Google* također koristi u svrhu razvoja svojih proizvoda poput *Google Maps-a* i *Google Translate-a*.

2.6.3. Sleep as Android

Sleep as Android mobilna je aplikacija koja pomoću senzora u pametnim uređajima prati spavanje korisnika [22]. Aplikacija je u mogućnosti pratiti cijeli tok sna uz podržanu opciju pametnog buđenja (engl. *smart wake up*) koja na temelju očitavanja senzora računa optimalno vrijeme buđenja te tada pali alarm. Pomoću senzora prikupljaju se podaci o snu, a njihovom obradom aplikacija pruža razne povratne informacije o kvaliteti sna, fazama sna i raznim poremećajima. Podržana je integracija s većim zdravstvenim servisima kao što su *Google Fit* i *Zenobase*, a prije korištenja aplikacije korisnik treba dati informacije o svojoj rutini spavanja kako bi se olakšala prilagodba. Kada korisnik u svoju blizinu stavi uređaj za praćenje sna, unutar aplikacije potrebno je jednim klikom uključiti opciju praćenja čime počinje praćenje sna korisnika i izrada bilješki o pojedinostima sna u aplikaciji. Zaslona navedene opcije prikazan je na slici 2.12.



Slika 2.12. Zaslona praćenja sna u aplikaciji *Sleep as Android*

Prikupljeni podaci šalju se *Google-u* koji ih koristi kako bi potpomogao istraživanjima u području ljudskog sna čime se ističe uloga ove aplikacije kao alata za prikupljanje podataka iz mnoštva pomoću senzora. Aplikacija pruža i mehanizme modifikacije privatnosti pomoću kojih korisnik može odabrati želi li da njegovi anonimizirani podaci budu korišteni u statistikama.

2.6.4. Open Data Kit

Open Data Kit (ODK) pruža web i mobilnu aplikaciju za izradu formi koje mogu sadržavati fotografije, GPS lokacije, izračune, vanjske izvore podataka i slično [23]. Koristi se u raznim područjima koja uključuju zdravstvo, krizne odgovore, praćenje klime itd. ODK je *open-source* softver koji korisnik može urediti na koji god način želi. Podaci su prikupljeni i kada je podnositelj zahtjeva izvan mreže, a jednom kada se spoji na Internet odvija se automatska sinkronizacija prikupljenih resursa. Prikupljene je podatke moguće preuzeti ili povezati s aplikacijama kao što su *Excel*, *Power BI* ili programskim jezicima kao što je *Python*. *ODK Collect* naziv je Android aplikacije koja omogućuje popunjavanje formi uz pomoć radnika na nekom projektu. Ispunjene forme šalju se na centralni poslužitelj (najčešće organizacije odabiru *ODK-ov Central* poslužitelj na *ODK Cloudu*) na analizu i dodatnu obradu. *ODK Collect* nudi opciju probavanja aplikacije bez potrebe unošenja projekta pomoću demo projekta koji koristi demo poslužitelj. Prazna forma predstavljena je XML datotekom koju je moguće kreirati koristeći *ODK Build* ili *XLSForm*. Podnositelj zahtjeva kreira projekt i pripadajuće forme te ih šalje na poslužitelj. Zatim svoj projekt podijeli sa željenim radnicima koji mu pristupaju pomoću *QR koda* ili nekim drugim putem unutar *ODK Collect* aplikacije. Mnoštvo putem aplikacije preuzima forme sa poslužitelja, ispunjava ih te ih ispunjene šalje nazad na poslužitelj odakle podnositelj zahtjeva može preuzeti prikupljene i obrađene podatke.

3. PROGRAMSKO RJEŠENJE ZA PRIKUPLJANJE PODATAKA IZ MNOŠTVA

Praktični dio ovog rada obuhvaća *Android* aplikaciju *Life Quality App* koja prikuplja podatke iz okoline korisnika te njihovu analizu prikazom na karti. Svrha aplikacije je prikaz mogućnosti mobilnih uređaja u prikupljanju podataka iz mnoštva.

3.1. Zahtjevi na programsko rješenje

Aplikacija prilikom svakog mjerenja prikuplja korisnikovu lokaciju, razinu buke u okolini te očitavanja senzora za mjerenje temperature, tlaka zraka i relativne vlažnosti (praćenje ljudi pomoću uređaja). Sustav nagrađivanja je poticajni mehanizam koji korisniku dodjeljuje određen broj bodova za svaki podatak koji prikupi, uz dodatak da je informacija o lokaciji obvezna kako bi se prikupljeni podaci kasnije uspješno analizirali. U tablici 3.1. navedene su funkcionalnosti aplikacije *Life Quality App*.

Tablica 3.1. Funkcionalnosti aplikacije *Life Quality App*

	FUNKCIONALNOST	NAZIV	OPIS
F1	autentifikacija korisnika	autentifikacija	korisnik se prijavljuje koristeći Google račun
F2	kontrola nad procesom mjerenja	upravljanje mjerenjem	korisnik u bilo kada može pokrenuti i zaustaviti prikupljanje podataka
F3	prikupljanje podataka o buci u okolini uređaja (ručno)	ručni unos razine buke	korisnik odabire jednu od ponuđenih razina buke u svom okruženju te podatak šalje na poslužitelj
F4	prikupljanje podataka o buci u okolini uređaja (automatski)	mjerenje razine buke	podaci o buci u okolini mobilnog uređaja prikupljaju se pomoću mikrofona i šalju na poslužitelj bez intervencije korisnika
F5	prikupljanje podataka o temperaturi zraka u okolini uređaja	mjerenje temperature	podaci o temperaturi zraka u okolini (°C) mobilnog uređaja šalju se na poslužitelj bez intervencije korisnika
F6	prikupljanje podataka o lokaciji uređaja	lociranje	podaci o lokaciji mobilnog uređaja šalju se na poslužitelj bez intervencije korisnika
F7	prikupljanje podataka o tlaku zraka u okolini uređaja	mjerenje tlaka zraka	podaci o tlaku zraka u okolini (hPa) mobilnog uređaja šalju se na poslužitelj bez intervencije korisnika
F8	prikupljanje podataka o relativnoj vlažnosti u okolini uređaja	mjerenje relativne vlažnosti okoline	podaci o relativnoj vlažnosti u okolini (%) mobilnog uređaja šalju se na poslužitelj bez intervencije korisnika

F9	održavanje stanja računa korisnika	stanje računa	korisnik ima uvid u stanje svojeg računa koje je iskazano u internoj valuti („bodovi“)
F10	nagrađivanje korisnika	nagrađivanje	sustav korisnika nagrađuje dodavanjem određenog broja bodova na račun za svako provedeno mjerenje

3.2. Korišteni alati i tehnologije

3.2.1. Android Studio

Android Studio je integrirano razvojno okruženje za izradu *Android* aplikacija [24]. Aplikacija *Life Quality App* napisana je u *Kotlin*-u, objektno-orijentiranom jeziku koji može komunicirati s *Java* virtualnim strojem i bibliotekama te *Androidom*. Nastao je sažimanjem i uklanjanjem redundancija prisutnih u *Java* programskom jeziku. Za razvoj korisničkog sučelja (engl. *user interface*) korišten je *Jetpack Compose*, skup alata koji preporučuje *Android* u svrhu kreiranja nativnog (engl. *native*) korisničkog sučelja. *Jetpack Compose* ubrzava razvoj korisničkog sučelja uz manje koda i intuitivne *Kotlin API*-je [25]. U svrhu testiranja aplikacije korišten je *Android Emulator* koji je integriran unutar *Android Studija*. *Android Emulator* simulira *Android* uređaj na računalu i nudi razne opcije za modifikaciju, poput mijenjanja razine *Android API*-ja, promjene lokacije uređaja, promjene vrijednosti koju očitava pojedini virtualni senzor itd.

3.2.2. Firebase Authentication i Realtime Database

Firebase je *Google* platforma za razvoj web i mobilnih aplikacija s uslugama poput autentifikacije, stvaranja baza podataka, analitike i pohrane dokumenta [26]. Aplikacija *Life Quality App* koristi uslugu autentifikacije uz pomoć *Google* računa koja je jedna od ponuđenih opcija u okviru autentifikacije putem *Firebase*-a. Glavni razlog korištenja autentifikacije putem *Google* računa je velika raširenost istih te jednostavnost koju ovaj način autentifikacije pruža korisniku (prijava jednim klikom (engl. *One Tap sign-in*)).

Realtime Database primjer je baze podataka u oblaku koja podatke sprema u JSON formatu i sinkronizira u stvarnom vremenu sa svim povezanim klijentima. Iz tog je razloga projekt *Life Quality App* unutar *Android Studija* povezan sa istoimenim projektom u okviru *Firebase*-a. Unutar baze podataka spremaju se podaci o mjerenjima i stanja računa svih korisnika. U svrhu bolje organizacije, zapisi unutar baze podataka razvrstani su u kolekcije (engl. *collections*) ovisno o tipu.

3.2.3. Rowy

Rowy je platforma koja pruža jednostavniji način za upravljanje infrastrukturom u pozadini (engl. *backend infrastructure*) aplikacije [27]. Omogućuje prikaz podataka u obliku tablice i praktično pisanje pozadinske logike. Povezivanjem projekta unutar *Rowy*-a s *Firebase* projektom, moguće je prikazati zapise iz baze podataka u tabličnom obliku te ih preuzeti u nekom od ponuđenih formata (npr. CSV).

3.2.4. Python

Python je objektno-orijentirani programski jezik visoke razine često korišten za analizu podataka, strojno učenje itd. Za analizu prikupljenih mjerenja korištene su biblioteke *Pandas* i *Plotly*. *Pandas* je biblioteka otvorenog koda u okviru *Python*-a za analizu i efikasan rad s tabličnim podacima [28]. Mjerenja zapisana unutar CSV datoteke pomoću odgovarajuće funkcije iz *Pandasa* jednostavno se učitaju u tablični oblik pogodan za analizu.

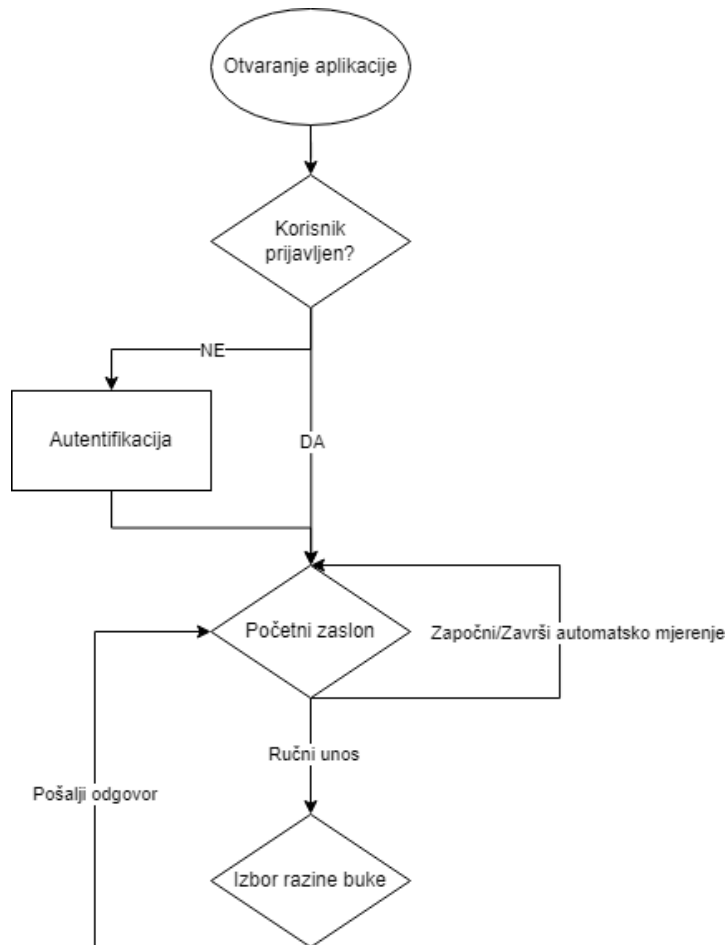
Plotly je također biblioteka otvorenog koda u okviru *Python*-a koja se koristi za grafički prikaz podataka [29]. Pomoću funkcija iz navedene biblioteke moguće je podatke zapisane u tabličnom obliku prikazati na karti. Parametri koje takve funkcije zahtijevaju su zemljopisna širina i dužina, vrsta karte (npr. *OpenStreetMap*), podaci za prikaz prilikom prelaska kursora preko zapisa na karti itd.

3.3. Prikaz načina rada aplikacije

Aplikacija *Life Quality App* omogućuje dvije vrste mjerenja, automatsko i ručno. Automatsko mjerenje podrazumijeva slučaj kada aplikacija samostalno obavlja prikupljanje (implicitno), dok je za ručno mjerenje potrebna interakcija korisnika prilikom svakog unosa (eksplicitno). Na slici 3.2. prikazan je dijagram tijeka aplikacije. Korisnici su međusobno neovisni, a sva se mjerenja šalju u jedinstvenu bazu, tj. sustav prati centraliziranu arhitekturu.

3.3.1. Autentifikacija i početni zaslon

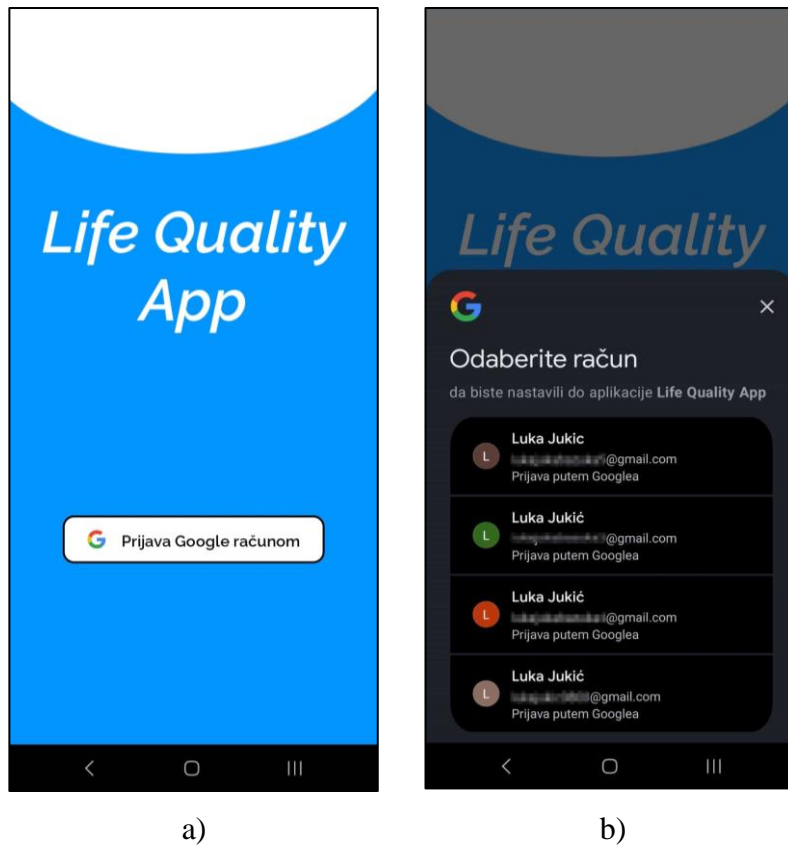
Otvaranjem aplikacije, provjerava se je li korisnik već prijavljen. Ukoliko je prijavljen, odmah se otvara početni zaslon aplikacije, a ukoliko nije, otvara se zaslon za prijavu prikazan na slici 3.3. pod a. Na zaslonu za prijavu nalazi se gumb „Prijava *Google* računom“, a klikom na gumb otvara se dijalog koji nudi sve *Google* račune trenutno dodane na uređaj prikazan na slici 3.3. pod b. Klikom na jedan od njih, korisnik se prijavljuje te se otvara početni zaslon aplikacije.



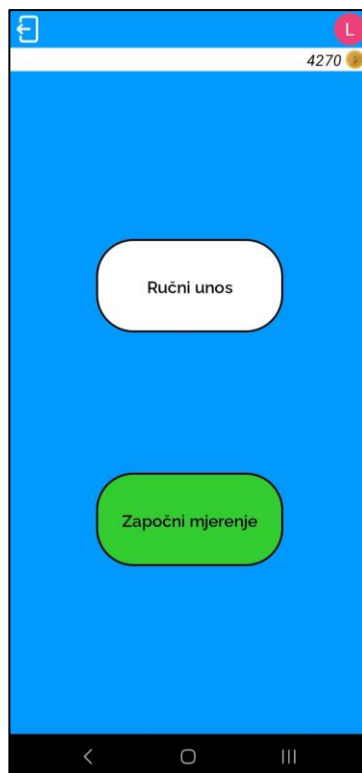
Slika 3.2. Dijagram tijeka aplikacije *Life Quality App*

Svrha postojanja autentifikacije u aplikaciji je omogućiti postojanje valute („bodovi“). Svakom računu je prilikom kreiranja jednoznačno pridružen zapis u bazi podataka koji odgovara stanju računa, tj. broju bodova koje korisnik tog računa posjeduje. Taj se zapis ažurira prilikom promjene stanja računa (npr. korisnik zaradi bodove na mjerenju).

Početni zaslon sastoji se od statusne trake i 2 gumba za mjerenje. Na statusnoj se traci u lijevom kutu nalazi gumb za odjavu koji vodi na zaslon za prijavu, dok se s desne strane nalaze profilna slika trenutno prijavljenog korisnika i broj bodova koji je jednak 0 kod novostvorenih računa. Ukoliko se stanje računa promijeni, broj bodova na statusnoj traci se ažurira. Gumb „Ručni unos“ vodi na zaslon za ručni unos razine buke, dok se drugim gumbom kontrolira automatsko mjerenje. Ovisno o tome je li automatsko mjerenje već započeto ili ne, sadržaj i izgled drugog gumba se mijenjaju. Opisani zaslon prikazan je na slici 3.4.



Slika 3.3. Zaslone za prijavu i dijalog s ponuđenim *Google* računima



Slika 3.4. Početni zaslon aplikacije

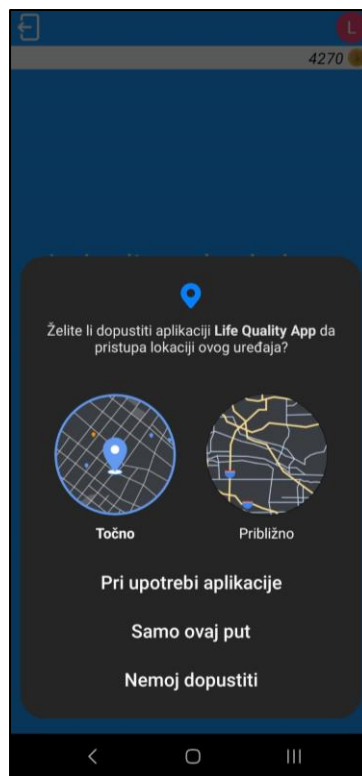
3.3.2. Ručno mjerenje

Dopuštenja

Kako bi ručno mjerenje uspjelo, obvezno je odobriti pristup lokaciji uređaja unutar aplikacije. Lokacija je parametar neophodan za prikaz mjerenja na karti. Lokacijsko dopuštenje jedno je od dopuštenja koje je potrebno zatražiti tijekom izvođenja (engl. *runtime permission*) jer je lokacija uređaja osjetljiv podatak.

Radi zaštite privatnosti, *Android* 12 i novije verzije dopuštaju korisniku da odabere želi li da aplikacija koristi točnu ili približnu lokaciju [30], a uspjeh mjerenja ne ovisi o vrsti lokacije koju korisnik odabere. Dijalog za lokacijsko dopuštenje (engl. *location permission*) pokreće se prilikom pokušaja da se obavi ručno mjerenje ako dopuštenje već ranije nije odobreno.

Počevši od *Androida* 11, kod zahtjeva za korištenjem lokacije, mikrofona i kamere, korisnik može odabrati i opciju „Samo ovaj put“ [31]. U tom će se slučaju dijalog za dopuštenje pojaviti svaki puta kada se iznova pokrene aplikacija. Ukoliko korisnik 2 puta odbije dopuštenje koje je potrebno zatražiti tijekom izvođenja, može ga odobriti samo u postavkama aplikacije u sklopu sustava. Primjer zahtjeva za pristupom lokaciji prikazan je na slici 3.5.



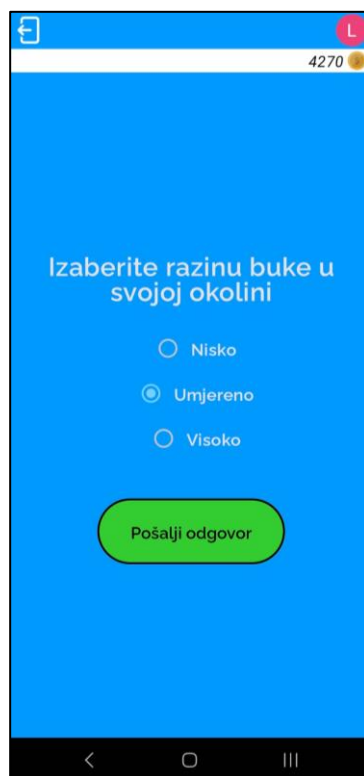
Slika 3.5. Dijalog zahtjeva za pristupom lokaciji

Prilikom razvoja aplikacije, sva je dopuštenja potrebno navesti u *Android Manifest* datoteci [32]. *Android Manifest* je XML datoteka koja sadrži neophodne informacije o aplikaciji poput aktivnosti i servisa, a lokacijsko dopuštenje definira se pomoću ključnih riječi *ACCESS_FINE_LOCATION* (točna lokacija) i *ACCESS_COARSE_LOCATION* (približna lokacija).

Osim lokacijskog dopuštenja, za ručno mjerenje potrebna su i dopuštenja za pristup Internetu i provjeru stanja mreže (*INTERNET* i *ACCESS_NETWORK_STATE*). Za navedena dopuštenja nije potrebno eksplicitno tražiti odobrenje, dovoljno ih je navesti u *Android Manifest* datoteci. Koriste se kako bi se moglo provjeriti je li korisnik spojen na Internet te za slanje odgovora u bazu podataka.

Postupak unosa

Nakon što se na početnom zaslonu klikne gumb „Ručni unos“, otvara se zaslon za unos razine buke prikazan na slici 3.6. Sastoji se od već objašnjene statusne trake, opcija za jačinu buke i gumba „Pošalji odgovor“ za unos mjerenja. Opcije za jačinu buke su „Nisko“, „Srednje“ i „Visoko“ koje se pohranjuju kao brojevi 0, 1 i 2. Takav unos spada u kategoriju zadataka s



Slika 3.6. Zaslon za unos razine buke

višestrukim izborom koji zahtijevaju ljudsku inteligenciju, u ovom slučaju aproksimaciju buke u okolini u najprikladniju od 3 razine.

Klikom na gumb „Pošalji odgovor“ provjerava se je li odobreno lokacijsko dopuštenje te se ukoliko nije otvara dijalog sa zahtjevom za odobrenje. Ako je dopuštenje odobreno, provjerava se je li korisnik spojen na Internet koji je nužan za slanje mjerenja u bazu podataka. Ukoliko nije, mjerenje neće biti započeto i nakratko će se ispisati uputa korisniku da se spoji na Internet. Ukoliko je korisnik spojen na Internet, provjerava se jesu li upaljene lokacijske usluge (GPS). Nužno ih je omogućiti kako bi se uređaj locirao, a navedeno se postiže u postavkama sustava. Ako nisu omogućene, nakratko se ispisuje uputa korisniku da ih uključi.

Kada su svi prethodno navedeni uvjeti ispunjeni, pritiskom na gumb „Pošalji odgovor“ mjerenje se sprema u bazu podataka. Korisnika se vraća na početni zaslone, uz kratku poruku o tome da je mjerenje uspješno provedeno te o broju zarađenih bodova. Za uspješno slanje zemljopisne dužine i širine (lokacije) te intenziteta buke zarađuje se 20 bodova (bez navedenih podataka ručno mjerenje ne može biti uspješno), a za slanje očitavanja svakog senzora zarađuje se po 5 bodova. *Android* ne zahtijeva odobrenje korisnika kako bi se koristili senzori uređaja. Ovakav sustav nagrađivanja potiče korisnika na mjerenje i mehanizam je koji privlači korisnike aplikaciji. Temperatura se sprema u Celzijevim stupnjevima (°C), tlak zraka u hektopaskalima (hPa), a relativna vlažnost okoline u postocima (%). Ukoliko uređaj nema neki od spomenuta 3 senzora, u bazi se očitavanje tog senzora spremi kao vrijednost NULL, a korisnik ne zaradi dodatnih 5 bodova. Primjer ručnog unosa u bazu podataka prikazan je na slici 3.7. Parametri *latitude* i *longitude* odgovaraju zemljopisnoj širini i dužini, *noiseLevel* odabranoj razini buke, *pressure* očitavanju barometra, *relativeHumidity* očitavanju senzora za relativnu vlažnost te *temperature* očitavanju termometra.

```
latitude: 45.28767
longitude: 18.8061867
noiseLevel: 1
pressure: 985
relativeHumidity: 66
temperature: 22
```

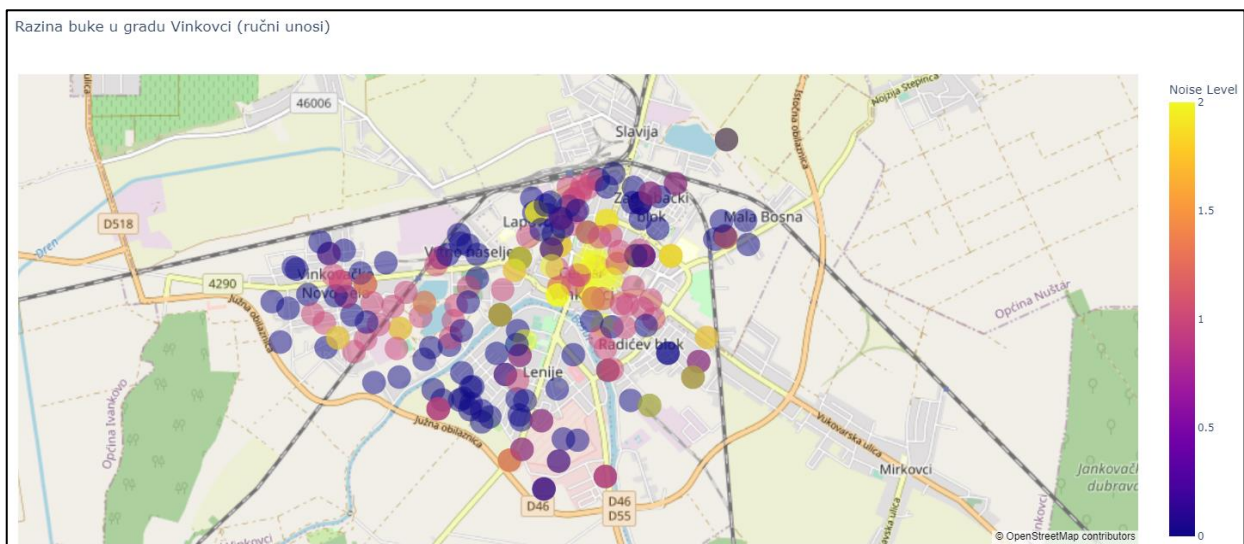
Slika 3.7. Primjer ručnog unosa u *Realtime Database* bazu podataka

Sva mjerenja u bazi su anonimna, tj. iz podataka o mjerenju nikako nije moguće saznati koji ga je korisnik izvršio. Pohrana identifikatora korisnika u sklopu svakog mjerenja omogućila bi kreiranje profila korisnika s procjenom kvalitete njegovih mjerenja, što bi dovelo do lakše detekcije

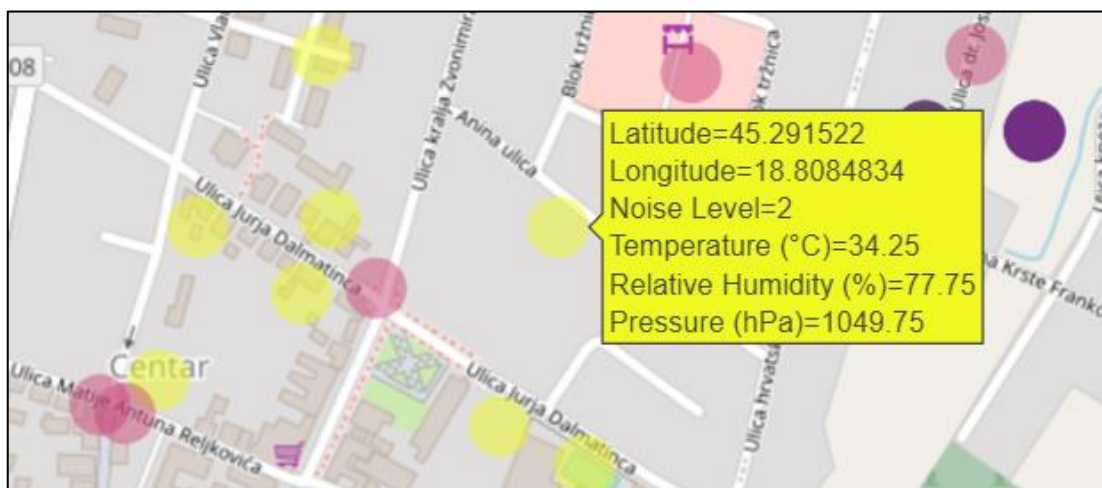
spamera. Dakako, time se narušava korisnikova privatnost jer se lako može doći do osjetljivih informacija o korisniku poput adrese stanovanja i često posjećenih mjesta. U tom bi se slučaju za širu uporabu podataka mogla provesti anonimizacija uklanjanjem identifikatora korisnika kako bi se očuvala privatnost.

Prikaz na karti

Nakon prikupljanja ručnih mjerenja u gradu Vinkovcima i izvoza pomoću *Rowy*-a ista su učitana i prikazana pomoću spomenutih biblioteka u *Pythonu*. Karta koja se pritom koristi je *OpenStreetMap*, a rezultati su prikazani na slici 3.8. Svaka točka na karti predstavlja jedno mjerenje, a prelaskom kursora preko točke ispisuju se svi parametri unosa analogni onima u bazi podataka. Primjer ispisa koji se pojavi kada se kursomom pređe preko točke prikazan je na slici 3.9. Svjetlije točke označavaju područja s višom razinom buke, a tamnije s nižom. Uočljivo je da je intenzitet buke najveći u centru grada, što se može pripisati prometu i količini ljudi.



Slika 3.8. Prikaz ručnih mjerenja na karti



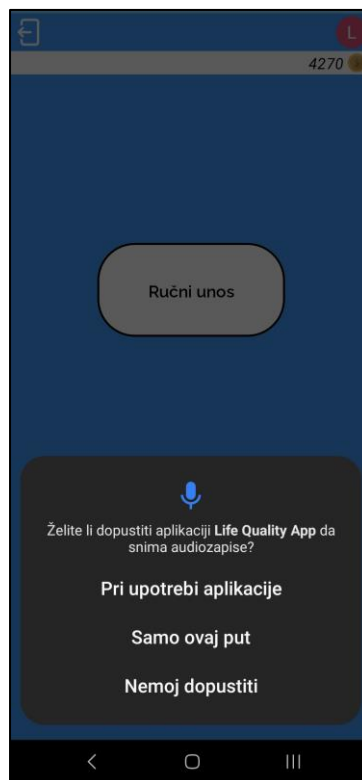
Slika 3.9. Ispis parametara unosa prelaskom kursora preko točke

3.3.3. Automatsko mjerenje

Dopuštenja

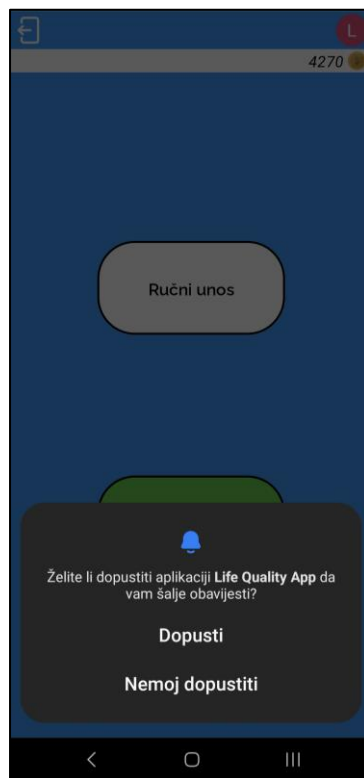
Kako bi automatsko mjerenje bilo uspješno, dodatno je uz sva dopuštenja koja su potrebna za ručno mjerenje nužno odobriti i snimanje zvuka, tj. pristup mikrofonu. Time se omogućuje mjerenje razine buke bez interakcije korisnika. Navedeno dopuštenje spada u istu kategoriju dopuštenja kao i lokacijsko, tj. korisnika je potrebno pitati za odobrenje tijekom izvođenja. Dijalog upita prikazan je na slici 3.10., a dopuštenje je unutar *Android Manifest* datoteke potrebno definirati pomoću ključne riječi *RECORD_AUDIO*.

Dodatno se od korisnika traži i dopuštenje za slanje obavijesti. Počevši od *Androida* 13 slanje obavijesti nije dopušteno nijednoj novoinstaliranoj aplikaciji te je korisnika potrebno pitati da



Slika 3.10. Dijalog zahtjeva za pristupom mikrofonu

navedeno odobri prilikom izvođenja aplikacije [33]. Ovo dopuštenje nije nužno odobriti kako bi automatsko mjerenje uspjelo, no u tom slučaju korisnik neće primati obavijesti vezane uz status mjerenja. Dijalog dopuštenja prikazan je na slici 3.11., a unutar *Android Manifest* datoteke definira se pomoću ključne riječi *POST_NOTIFICATIONS*. Svi se dijalozi, kao i kod ručnog mjerenja, pokreću prilikom pokušaja započinjanja automatskog mjerenja ukoliko dopuštenja već ranije nisu odobrena.



Slika 3.11. Dijalog zahtjeva za slanjem obavijesti

Posljednje obvezno dopuštenje uključuje korištenje servisa u prednjem planu (engl. *foreground service*) koje nije potrebno odobriti tijekom izvođenja. Servis u prednjem planu ključan je za programsko ostvarenje automatskog mjerenja. Ovaj tip servisa se nakon pokretanja izvršava neovisno o životnom ciklusu aplikacije sve dok se ne zadovolje uvjeti prekida ili dok ga ne zaustave korisnik ili sustav [34]. Tijekom izvršavanja korisniku je prikazana statusna obavijest ukoliko je slanje obavijesti odobreno, a u suprotnom pokrenute servise ovoga tipa na novijim inačicama *Androida* može vidjeti izlistane pomoću opcije u sklopu zaslona s nedavnim aplikacijama (engl. *recents screen*). Dopuštenje se navodi unutar *Android Manifest*-a pomoću ključne riječi *FOREGROUND_SERVICE*.

Ograničenja

Kako bi se očuvala privatnost korisnika, od *Androida* 11 uvedena su ograničenja vezana uz pristup mikrofону, lokaciji i kameri unutar servisa u prednjem planu [34]. Servis u prednjem planu više nije moguće pokrenuti dok je aplikacija u pozadini ukoliko nije ispunjen neki od iznimnih uvjeta, a tako pokrenut servis može pristupiti lokaciji ako korisnik odobri pristup lokaciji u pozadini. Samo u iznimnim situacijama, servis u prednjem planu pokrenut iz pozadine može pristupiti lokaciji (bez spomenutog odobrenja), mikrofону i kameri. Navedena ograničenja onemogućuju korištenje *Android* komponenti prikladnijih za periodično obavljanje posla poput *WorkManager*-

a, koji se zasniva na radu u pozadini i manjoj uporabi resursa [35]. Korištenjem takvih komponenti u kombinaciji sa servisom u prednjem planu otežano je prikupljanje lokacije i razine buke pomoću mikrofona u okolini korisnika, stoga se za prikupljanje koristi samo servis u prednjem planu pokrenut prilikom započinjanja mjerenja (dok je aplikacija vidljiva korisniku) što može dovesti do veće potrošnje baterije.

Tijek mjerenja

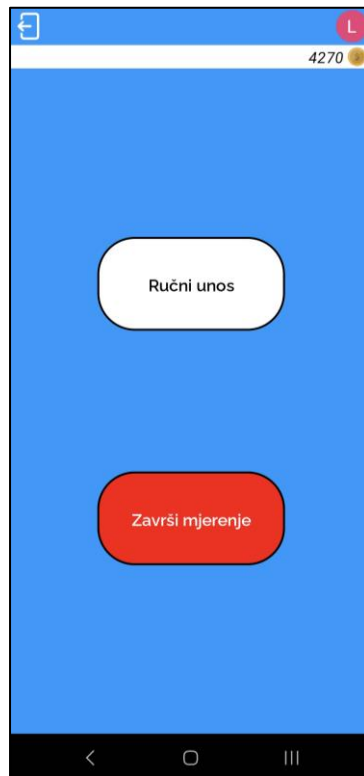
Automatsko mjerenje započinje se klikom na zeleni gumb „Započni mjerenje“ na početnom zaslonu. Pritiskom na gumb, najprije se provjerava jesu li odobrena sva nužna dopuštenja te se pojavljuju dijalozi neodobrenih. Nakon toga, kao i kod ručnog mjerenja, primjenjuje se logika vezana uz uključivanje lokacijskih usluga i Interneta.

Ukoliko su svi uvjeti zadovoljeni, mjerenje se započinje. Gumb „Započni mjerenje“ odmah se nakon klika promijeni u crveni gumb s natpisom „Završi mjerenje“, a funkcionalnost mu postaje zaustavljanje automatskog mjerenja. Izgled početnog zaslona kada je mjerenje započeto prikazan je na slici 3.12. Gumb je takav sve dok je mjerenje u tijeku, bez obzira na to što se radi s aplikacijom u međuvremenu. Nakon klika na takav gumb, mjerenje se zaustavlja, a dizajn i funkcionalnost gumba vraćaju se u prvobitno stanje.

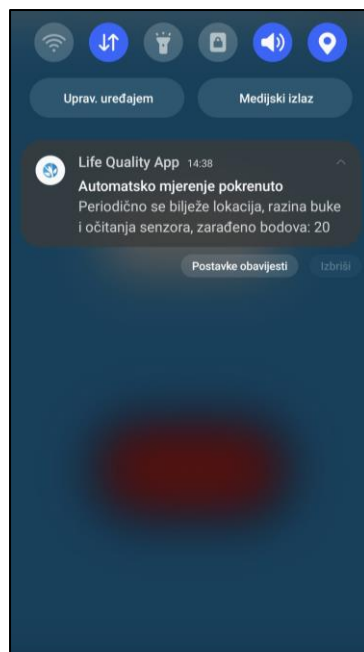
Nakon što protekne 10 sekundi od početka mjerenja, započinje implicitno prikupljanje podataka i pojavljuje se obavijest (ukoliko je korisnik dao dopuštenje) kojom se informira korisnika o postupku mjerenja te o ukupno zarađenom broju bodova. Obavijest je prikazana na slici 3.13., a klikom na nju otvara se početni zaslon aplikacije. Moguće ju je i ukloniti, a razlog odstupanja od 10 sekundi vezan je uz logiku upravljanja obavijestima sustava [34]. Sukladno tome, kako bi se osiguralo da se ništa ne događa bez korisnikova znanja, prikupljanje započinje tek kada se pojavi obavijest.

Mjerenje se odvija tako da se podaci prikupljaju periodično, svakih 15 minuta. Ostvareno je pomoću servisa u prednjem planu koji se izvršava sve dok ono traje. Postupak prikupljanja traje nekoliko sekundi, a jedina razlika u odnosu na parametre ručnog unosa je u razini buke koja se mjeri pomoću mikrofona.

Vrijednost koja predstavlja razinu buke kod automatskog mjerenja je amplituda zvuka. Parametar *noiseAmplitude* automatskog mjerenja odgovara prosječnoj amplitudi zvuka u okolini. Do njegove se vrijednosti dolazi pomoću funkcije *getMaxAmplitude* u sklopu *MediaRecorder* klase koja se koristi za snimanje zvuka i videa [36]. Navedena funkcija vraća vrijednost između 0 i 32767 koja



Slika 3.12. Početni zaslon tijekom trajanja automatskog mjerenja



Slika 3.13. Informativna obavijest o automatskom mjerenju

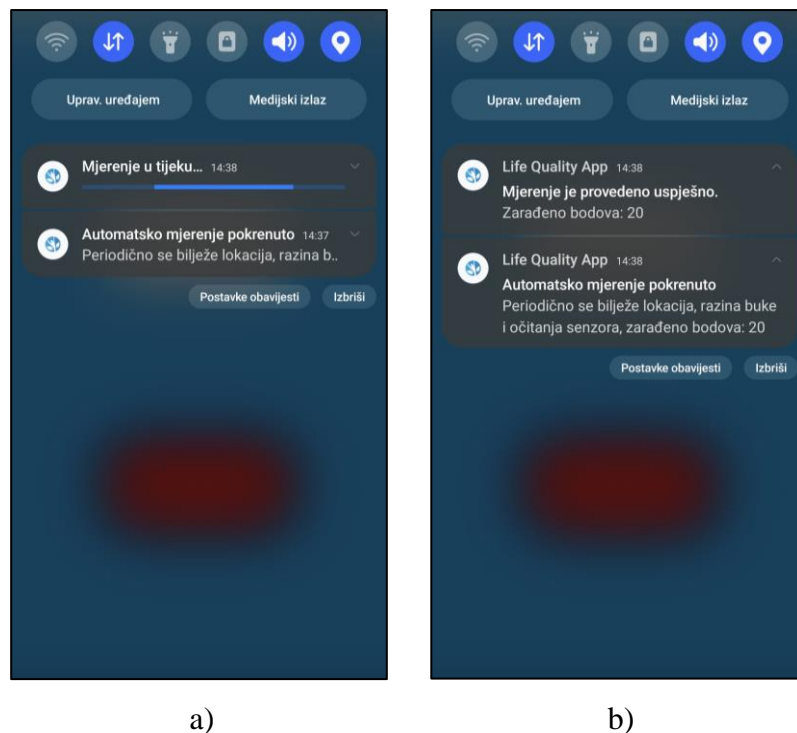
predstavlja maksimalnu zabilježenu amplitudu od zadnjeg poziva. Prosječna amplituda određuje se tako da se zvuk snima 5 sekundi pomoću mikrofona s pozivima *getMaxAmplitude* svakih 100 milisekundi. Nakon što se snimanje završi, prosjek svih zabilježenih amplituda tijekom snimanja

predstavlja prosječnu amplitudu. Primjer automatskog mjerenja u bazi podataka prikazan je na slici 3.14.

```
latitude: 45.2928117
longitude: 18.8038568
noiseAmplitude: 5882
pressure: 1035.25
relativeHumidity: 6.5
temperature: 36.625
```

Slika 3.14. Primjer automatskog mjerenja u *Realtime Database* bazi podataka

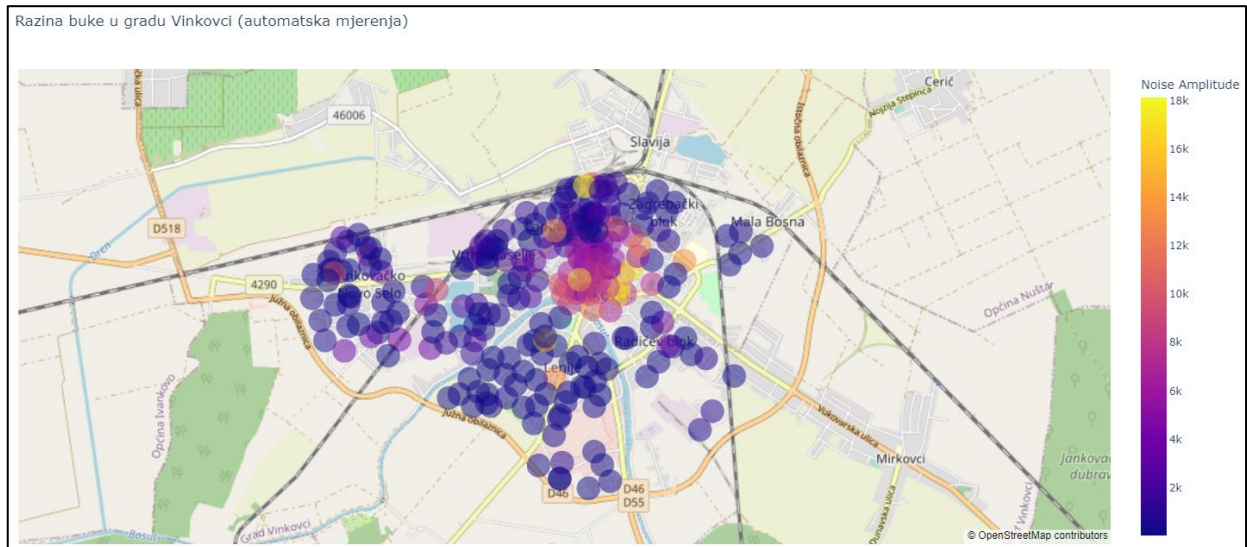
Prilikom svakog prikupljanja, pojavljuje se obavijest s animacijom učitavanja prikazana na slici 3.15. pod a) kojom se korisnika informira da je ono u tijeku, a nakon završetka prikazuje se obavijest na slici 3.15. pod b) s povratnim informacijama o provedenom prikupljanju. Logika nagrađivanja jednaka je kao i kod ručnog mjerenja. Ukoliko prikupljanje iz nekog razloga nije uspjelo (npr. korisnik je u međuvremenu ugasio lokacijske usluge ili Internet), pojavljuje se obavijest s razlogom i rješenjem. Sve navedene obavijesti moguće je ukloniti.



Slika 3.15. Obavijesti prilikom prikupljanja podataka tijekom automatskog mjerenja

Prikaz na karti

Nakon prikupljanja automatskih mjerenja u gradu Vinkovcima, ostvaren je prikaz na karti na isti način kao i kod ručnih mjerenja. Karta s automatskim mjerenjima prikazana je na slici 3.16., a prelaskom kursora preko točke ispisuju se svi parametri mjerenja kao i u slučaju prikaza ručnih unosa. Iz prikaza se ponovno može zaključiti da je intenzitet buke najveći u centru grada.



Slika 3.16. Prikaz automatskih mjerenja na karti

3.3.4. Primjena

Primjena podataka prikupljenih aplikacijom može biti usmjerena u poboljšanje kvalitete života. Kombinacijom lokacije i korištenih senzora pruža se uvid u vremenske prilike na određenim mjestima, a pravilnom vizualizacijom i analizom moguće je primijetiti anomalije te poduzeti mjere na područjima s lošijim životnim uvjetima. Glavna prednost korištenja aplikacije na malom prostoru poput grada upravo je lakša i kvalitetnija analiza jer su neobične pojave lakše uočljive. Vizualizacija korištenja lokacije zajedno s podacima o razini zvuka u okolini ostvarena je na karti, a slična primjena viđena je u projektu *City Soundscape*. Analizom takvih podataka moguće je zaključiti koja su područja više izložena buci te na tim područjima poduzeti mjere koje će pomoći u suzbijanju.

Aplikaciju može koristiti bilo tko s *Android* mobilnim uređajem, a krajnji korisnici podataka mogu biti uprave gradova ili kompanije koje će ih analizirati i odgovorno postupiti. Osim toga, moguć je i doprinos istraživanjima koja se bave intenzitetom zvuka i vremenskim uvjetima.

4. ZAKLJUČAK

Podaci iz mnoštva pridonose razvoju cjelokupnog IT sektora pružajući velike količine informacija nastalih djelovanjem ljudi. Kompanije čije poslovanje direktno ovisi o publici, velike napore ulažu u prikupljanje i obradu informacija dobivenih od ljudi kako bi se prilagodile trendovima i željama gomile. Prikupljanje podataka kompleksan je zadatak koji zahtijeva dizajn specijaliziranog sustava i pažljivo balansiranje parametara o kojima ovisi kvaliteta rezultata. Rad naglašava važnost mobilne nabave iz mnoštva, područja koje je zaživjelo naglim razvojem mobilnih uređaja i njihovih karakteristika. Većina ljudi mobilne uređaje uvijek ima uz sebe, čime se mogućnosti preciznog praćenja i informiranja o stanju korisnika i njegove okoline dodatno povećavaju. Pri tome se mobilni sustavi korišteni za prikupljanje modeliraju po uzoru na dvije osnovne arhitekture opisane u radu, centraliziranoj i decentraliziranoj. U mobilnoj se nabavi javljaju izazovi analizirani u radu od kojih se najviše ističu potrošnja energije i očuvanje privatnosti korisnika. Važno je pružiti visoku razinu privatnosti korisniku zbog postojanja osjetljivih informacija koje uređaji mogu zabilježiti, a u tu se svrhu primjenjuju tehnike anonimizacije podataka. Navedeni su i opisani primjeri prikupljanja podataka i mobilnih platformi za nabavu iz mnoštva.

Kako bi se prikazale mogućnosti mobilnih uređaja u nabavi iz mnoštva, implementirano je i opisano programsko rješenje koje predstavlja *Android* aplikacija *Life Quality App*. Aplikacija prikuplja lokaciju korisnika, razinu buke u okolini i očitavanja senzora uređaja, uz izbor između ručnog i automatskog mjerenja. Za ručno mjerenje potrebno je odabrati jednu od 3 ponuđene razine buke u okolini, dok automatsko mjerenje buku očitava pomoću mikrofona. Kako bi se ručno mjerenje moglo izvesti, potrebno je odobriti lokacijsko dopuštenje, dok je za automatsko potrebno odobriti i pristup mikrofona. Za provedbu automatskog mjerenja nije potrebna interakcija korisnika s aplikacijom osim u pogledu pokretanja i zaustavljanja, no glavni nedostatak u odnosu na ručni unos predstavlja povećana potrošnja energije uzrokovana servisom u prednjem planu koji se izvršava tijekom trajanja mjerenja. Aplikacija je testirana u gradu Vinkovcima te su rezultati za obje vrste mjerenja prikazani na karti što predstavlja moguć slučaj korištenja ovakvog sustava.

Unaprjeđenje aplikacije moguće je u pogledu manje potrošnje energije automatskog mjerenja ugrađivanjem energetski učinkovitijih komponenti i njihovom kombinacijom sa servisom u prednjem planu. Osim toga, dodavanjem opcija za modifikaciju učestalosti prikupljanja podataka i uzorkovanja prilikom snimanja zvuka tijekom automatskog mjerenja moguće je ostvariti veću fleksibilnost i bolju prilagodbu korisničkim potrebama. Dodatno je ostvariv mehanizam koji korisniku omogućuje izbor želi li da se njegovi podaci spremaju anonimno kao do sada, čime je u

potpunosti očuvana njegova privatnost, ili da se spremaju s identifikatorom korisnika što može rezultirati dodatnim nagradama jer je tako omogućena analiza pojedinog korisnika i kreiranje profila, no uz narušenu privatnost. Kako bi valuta unutar aplikacije dobila smisao, potrebno je razviti prilagođenu online trgovinu koja bi posljedično korisnike dodatno motivirala u prikupljanju.

LITERATURA

- [1] H. Niu i E. A. Silva, Crowdsourced Data Mining for Urban Activity: Review of Data Sources, Applications, and Methods, *Journal of Urban Planning and Development*, br. 2, sv. 146, str. 04020007, lipanj 2020.
- [2] H. Garcia-Molina, M. Joglekar, A. Marcus, A. Parameswaran i V. Verroios, Challenges in Data Crowdsourcing, *IEEE Transactions on Knowledge and Data Engineering*, br. 4, sv. 28, str. 901–911, travanj 2016.
- [3] G. Li, J. Wang, Y. Zheng i M. J. Franklin, Crowdsourced Data Management: A Survey, *IEEE Transactions on Knowledge and Data Engineering*, br. 9, sv. 28, str. 2296–2319, rujan 2016.
- [4] J. Phuttharak i S. W. Loke, A Review of Mobile Crowdsourcing Architectures and Challenges: Toward Crowd-Empowered Internet-of-Things, *IEEE Access*, br. 1, sv. 7, str. 304–324, siječanj 2019.
- [5] SendPulse, What is Crowdsourcing: Types and Examples [online], SendPulse, 2023., dostupno na: <https://sendpulse.com/support/glossary/crowdsourcing> [pristupljeno: 31. srpanj 2023.]
- [6] Great Backyard Bird Count, Join us each February when the world comes together to watch, learn about, count, and celebrate birds [online], Great Backyard Bird Count, 2023., dostupno na: <https://www.birdcount.org> [pristupljeno: 31. srpanj 2023.]
- [7] W. Feng, Z. Yan, H. Zhang, K. Zeng, Y. Xiao i Y. T. Hou, A Survey on Security, Privacy, and Trust in Mobile Crowdsourcing, *IEEE Internet Things Journal*, br. 4, sv. 5, str. 2971–2992, kolovoz 2018.
- [8] G. Barbier, R. Zafarani, H. Gao, G. Fung i H. Liu, Maximizing benefits from crowdsourced data, *Computational and Mathematical Organization Theory*, br. 3, sv. 18, str. 257–279, rujan 2012.
- [9] G. Li, Y. Zheng, J. Fan, J. Wang i R. Cheng, Crowdsourced Data Management: Overview and Challenges, u *Proceedings of the 2017 ACM International Conference on Management of Data*, str. 1711–1716, Chicago, Illinois, SAD, 14. - 19. svibnja 2017.
- [10] K. Yang, K. Zhang, J. Ren i X. Shen, Security and privacy in mobile crowdsourcing networks: challenges and opportunities, *IEEE Communications Magazine*, br. 8, sv. 53, str. 75–81, kolovoz 2015.

- [11] S. Meek, M. Jackson i D. Leibovici, A flexible framework for assessing the quality of crowdsourced data, u AGILE 2014 International Conference on Geographic Information Science, str. 1-7, Castellon, Španjolska, 3. - 6. lipnja 2014.
- [12] J. Brickell i V. Shmatikov, Efficient anonymity-preserving data collection, u Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining, str. 76–85, Philadelphia, Pennsylvania, SAD, 20. - 23. kolovoza 2006.
- [13] L. See, A. Comber, C. Salk, S. Fritz, M. v. d. Velde, C. Perger, C. Schill, I. McCallum, F. Kraxner i M. Obersteiner, Comparing the Quality of Crowdsourced Data Contributed by Expert and Non-Experts, PLoS ONE, br. 7, sv. 8, str. e69958, srpanj 2013.
- [14] J. Deng, W. Dong, R. Socher, L.-J. Li, Kai Li i Li Fei-Fei, ImageNet: A large-scale hierarchical image database, u 2009 IEEE Conference on Computer Vision and Pattern Recognition, str. 248–255, Miami, Florida, SAD, 20. - 25. lipnja 2009.
- [15] BrainyQuote, Fei-Fei Li Quotes [online], BrainyQuote, 2023., dostupno na: https://www.brainyquote.com/quotes/feifei_li_914167 [pristupljeno: 31. srpanj 2023.]
- [16] A. Longo, M. Zappatore, M. Bochicchio i S. B. Navathe, Crowd-Sourced Data Collection for Urban Monitoring via Mobile Sensors, ACM Transactions on Internet Technology, br. 1, sv. 18, str. 1–21, veljača 2018.
- [17] C. Ashton, A. Shuster Bruce, G. Colledge i M. Dickinson, The Search for MH370, The Journal of Navigation, br. 1, sv. 68, str. 1–22, siječanj 2015.
- [18] C. C. Freifeld, K. D. Mandl, B. Y. Reis i J. S. Brownstein, HealthMap: Global Infectious Disease Monitoring through Automated Classification and Visualization of Internet Media Reports, Journal of the American Medical Informatics Association, br. 2, sv. 15, str. 150–157, ožujak 2008.
- [19] N. Kela-Madar, Navigating Through Creativity and Innovation Paths: Waze Navigation Application Case Study, Journal of Entrepreneurship Education, br. 6, sv. 24, str. 1-12, lipanj 2021.
- [20] F. Mora, Innovating in the midst of crisis: A case study of Ushahidi, str. 1-37, svibanj 2011. [neobjavljen rad]
- [21] S. Sarin, K. Pipatsrisawat, K. Pham, A. Batra i L. Valente, Crowdsourced by Google: A Platform for Collecting Inclusive and Representative Machine Learning Data, u AAAI Conference on Human Computation and Crowdsourcing 2019, str. 1-3, Washington, SAD, 28. - 30. listopada 2019.
- [22] B. M. Chaudhry, Sleeping with an Android, mHealth, br. 2, sv. 3, str. 3–7, veljača 2017.

- [23] ODK, ODK - Collect data anywhere [online], ODK, 2023., dostupno na: <https://getodk.org> [pristupljeno: 31. srpanj 2023.]
- [24] Android Developers, Meet Android Studio [online], Android Developers, 2023., dostupno na: <https://developer.android.com/studio/intro> [pristupljeno: 1. kolovoz 2023.]
- [25] Android Developers, Jetpack Compose UI App Development Toolkit [online], Android Developers, 2023., dostupno na: <https://developer.android.com/jetpack/compose> [pristupljeno: 1. kolovoz 2023.]
- [26] TechTarget, What is Google Firebase? [online], TechTarget, 2023., dostupno na: <https://www.techtarget.com/searchmobilecomputing/definition/Google-Firebase> [pristupljeno: 1. kolovoz 2023.]
- [27] Rowy, Rowy - Low-code backend on Google Cloud and Firebase [online], Rowy, 2023., dostupno na: <https://www.rowy.io/> [pristupljeno: 1. kolovoz 2023.]
- [28] PyData, pandas - Python Data Analysis Library [online], PyData, 2023., dostupno na: <https://pandas.pydata.org/> [pristupljeno: 1. kolovoz 2023.]
- [29] Plotly, Plotly Open Source Graphing Library for Python [online], Plotly, 2023., dostupno na: <https://plotly.com/python/> [pristupljeno: 1. kolovoz 2023.]
- [30] Android Developers, Request location permissions [online], Android Developers, 2023., dostupno na: <https://developer.android.com/training/location/permissions> [pristupljeno: 4. kolovoz 2023.]
- [31] Android Developers, Request runtime permissions [online], Android Developers, 2023., dostupno na: <https://developer.android.com/training/permissions/requesting> [pristupljeno: 4. kolovoz 2023.]
- [32] Android Developers, App manifest overview [online], Android Developers, 2023., dostupno na: <https://developer.android.com/guide/topics/manifest/manifest-intro> [pristupljeno: 4. kolovoz 2023.]
- [33] Android Developers, Notification runtime permission [online], Android Developers, 2023., dostupno na: <https://developer.android.com/develop/ui/views/notifications/notification-permission> [pristupljeno: 9. kolovoz 2023.]
- [34] Android Developers, Foreground services [online], Android Developers, 2023., dostupno na: <https://developer.android.com/guide/components/foreground-services> [pristupljeno: 9. kolovoz 2023.]
- [35] Android Developers, WorkManager [online], Android Developers, 2023., dostupno na: <https://developer.android.com/reference/androidx/work/WorkManager> [pristupljeno: 9. kolovoz 2023.]

[36] Android Developers, MediaRecorder [online], Android Developers, 2023., dostupno na: <https://developer.android.com/reference/android/media/MediaRecorder> [pristupljeno: 11. kolovoz 2023.]

SAŽETAK

Prikupljanje podataka iz mnoštva područje je koje se ubrzano razvija. Razlog tomu je činjenica da mnoge poslove ne mogu obaviti računala jer zahtijevaju ljudsku inteligenciju. Razvojem mobilnih uređaja koji su opremljeni velikim brojem senzora i moćnim komponentama, nabava pomoću mobilnih platformi postaje sve raširenija. Prilikom dizajniranja sustava za prikupljanje resursa iz mnoštva, pojavljuju se izazovi vezani uz kvalitetu podataka, trošak i latenciju. Veliki izazov predstavlja očuvanje privatnosti korisnika jer prikupljeni podaci mogu biti osjetljivi, a u mobilnom prikupljanju pojavljuje se i problem potrošnje baterije. U praktičnom dijelu rada izrađena je i opisana *Android* aplikacija *Life Quality App* koja koristi mogućnosti mobilnih uređaja kako bi se prikupili podaci iz okoline korisnika. Aplikacija može prikupljati mjerenja bez intervencije korisnika, no postoji i mogućnost ručnog unosa mjerenja. Aplikacija je testirana u gradskim uvjetima s grafičkim prikazom prikupljenih podataka na karti kao slučajem korištenja.

Ključne riječi: mobilne platforme, mobilno prikupljanje iz mnoštva, podaci iz mnoštva, sustavi za prikupljanje podataka iz mnoštva

Crowdsourced data acquisition using mobile platforms

ABSTRACT

Data crowdsourcing is a rapidly evolving area. The reason for this is the fact that many jobs cannot be done by computers because they require human intelligence. With the development of mobile devices which are equipped with many sensors and powerful components, procurement using mobile platforms is becoming more prevalent. When designing a crowdsourcing system, there are challenges related to data quality, cost, and latency. A big challenge is the preservation of user privacy because the collected data can be sensitive, while the problem of battery consumption also appears in mobile collection. As a practical part of this thesis, the Android application Life Quality App which uses the capabilities of mobile devices to collect data from user's environment was created and described. The application collects measurements without user intervention, but there is also the option which allows users to manually input measurement. The application was tested in urban conditions with a graphical representation of the collected data on a map as a use case.

Keywords: mobile platforms, mobile crowdsourcing, crowdsourced data, crowdsourcing systems

ŽIVOTOPIS

Luka Jukić rođen je u Vinkovcima, 8. kolovoza 2001. godine. Osnovnu školu pohađao je u Vinkovcima. Srednjoškolsko obrazovanje stekao je u Gimnaziji Matije Antuna Reljkovića u Vinkovcima 2020. godine. Iste godine upisao je Preddiplomski sveučilišni studij Računarstvo na Fakultetu elektrotehnike, računarstva i informacijskih tehnologija Osijek, smjer Programsko inženjerstvo. Dobitnik je državne STEM stipendije 2020., 2021. i 2022. godine te nagrade za izniman uspjeh u studiranju koju fakultet dodjeljuje svake godine najboljem studentu na pojedinom smjeru.

PRILOZI

1. Završni rad u formatu .docx
2. Završni rad u formatu .pdf
3. Izvorni kod programskog rješenja