

# Detekcija modova prijevoza zasnovana na podacima kretanja korisnika prometnom mrežom

Martina Erdelić

Fakultet prometnih znanosti, Sveučilišta u Zagrebu  
Vukelićeva 4, Zagreb  
merdelic@fpz.hr

---

## Sažetak

Poznavanje mobilnosti korisnika prometne mreže je osnova za određivanje prometne potražnje i planiranje transporta. Porastom broja mobilnih uređaja oni postaju, uz uređaje ugrađene u vozilima, vrijedan izvor prometnih podataka o mobilnosti korisnika u mreži. Takvi podaci mogu se koristiti kao višestruki izvor dodatnih informacija o okruženju u kojem se korisnik nalazi. Dodatne informacije poput modova prijevoza koriste se tijekom raznovrsnih prometnih analiza, stoga se razvijaju metode temeljem kojih se iz prikupljenih podataka dobivaju dodatne informacije koje se ne mogu automatski prikupljati uređajem za prikupljanje podataka.

U radu je definiran problem detekcije modova prijevoza te osnovni dijelovi problema: segmentacija trajektorije, odabir relevantnih značajki te metode klasifikacije. Opisana su dva različita pristupa klasifikaciji modova prijevoza ovisno o tome detektira li se mod prijevoza iz trajektorije korisnika tijekom prikupljanja podataka ili nakon prikupljanja podataka. Također, u radu je dan pregled metoda koje se koriste u svakom dijelu detekcije moda prijevoza pri čemu su odvojeno analizirana istraživanja koja provode klasifikaciju nad povijesnim podacima i istraživanja koja provode klasifikaciju u stvarnom vremenu.

*Ključne riječi:* GNSS trajektorija; multimodalni prijevoz; klasifikacija; mod prijevoza; klasifikacija u stvarnom vremenu; segmentiranje trajektorije; detekcija moda prijevoza; mobilna osjetila; prepoznavanje uzoraka; prepoznavanje aktivnosti

---

## 1. Uvod

Za mnoge aplikacije u upravljanju prometom od velikog su interesa profili kretanja te tragovi kretanja korisnika prometnom mrežom. Tragovi kretanja prikazuju niz zapisa nekog osjetila koji nastaju tijekom kretanja korisnika prometnom mrežom, a pod profilima kretanja podrazumijevaju se skup tragova kretanja grupiranih prema sličnosti. Takvi podaci koriste se kao ulazni parametri kod izrade simulacijskih modela. Učestalo se koriste podaci poput odredišnih izvorišnih matrica, podaci o stanju cestovnih segmenata i sl. Prikupljanje podataka o putovanju korisnika prometne mreže provodi se pomoću dvije osnovne metode. Prva metoda odnosi se na ispitivanje korisnika kroz anketne obrasce o dnevnim migracijama. Ovaj pristup je u praksi dugi niz godina te još uvijek ima široku primjenu. Bez obzira na korištenje ove metode ona je kroz godine pokazala niz nedostataka, poput netočnih informacija zbog lošeg sjećanja korisnika ili nedovoljnog odaziva korisnika koji žele dati informacije o svojim dnevnim migracijama. Osim toga, prikupljanje podataka ovom metodom iziskuje mnogo vremena. Druga metoda prikupljanja podataka odnosi se na automatsko prikupljanje podataka o kretanju korisnika s uređaja podržanih satelitskim sustavom za globalno pozicioniranje (engl. *Global Navigation Satellite System*, GNSS) ili drugih mjernih uređaja. Ovom metodom prikuplja se znatno veći i točniji skup podataka u kraćem vremenu. Iz prikupljenih podataka mogu se izdvojiti fizikalne veličine, poput brzine koje se mogu statistički analizirati kroz deskriptivne mjere kao što su aritmetička sredina, medijan, maksimalna vrijednost i inferencijalne mjere kao što su koeficijenti asocijacije i korelacije. Tako obrađeni podaci nazivaju se značajke te se koriste kao ulazni podaci metodama za predviđanje, npr. prometnog toka.

Učestalost korištenja pojedinog moda prijevoza na nekom području koristi se tijekom planiranja infrastrukture u gradskim sredinama, planiranja prometa ili za razumijevanje ponašanja korisnika prometne mreže. Za detekciju moda prijevoza koji se koristi na nekom dijelu prometne mreže mogu se koristiti podaci o kretanju korisnika. Ovisno o uređaju kojim se podaci prikupljaju mogu se dobivati podaci s raznovrsnih senzora. Najčešće se koriste podaci s mobilnih uređaja koji u tom slučaju mogu uključivati podatke s akcelerometra, GNSS uređaja, žiroskopa, magnetometra,

Bluetooth uređaja i drugih senzora ugrađenih u mobilnom uređaju. Iz prikupljenih podataka se izdvajaju značajke koje opisuju korišteni mod prijevoza u trenutnoj trajektoriji korisnika. Na temelju značajki, trajektorije se klasificiraju prema modu prijevoza. Tijekom putovanja korisnika pojavljuju se raznovrsni motorizirani modovi prijevoza koji se ostvaruju vožnjom automobilom, vlakom, autobusom, električnim biciklom i drugim motornim vozilima te nemotorizirani modovi prijevoza koji se ostvaruju pješaćenjem, vožnjom biciklom i drugim sličnim vozilima bez motornog pogona. U svrhu klasifikacije koriste se razne klasifikacijske metode čiji je broj u posljednjih nekoliko godina sve veći. Klasifikacija modova prijevoza zasniva se na dva osnovna pristupa. Prvi pristup odnosi se na klasifikaciju modova prijevoza na povijesnim podacima. Dakle, podaci se određeno vrijeme prikupljaju sa senzora koje korisnici nose, te istovremeno vode dnevnik putovanja (bilježi se vremenski interval i korišteni mod prijevoza). Korisnički dnevnicu spajaju se s podacima sa senzora te zajedno čine skup podataka koji ima pridruženu klasu, u ovom slučaju mod prijevoza. Tako prikupljeni i obrađeni podaci koriste se za učenje klasifikatora. Drugi pristup odnosi se na klasifikaciju modova prijevoza tijekom prikupljanja podataka. Dakle, podaci se prikupljaju u unaprijed određenom vremenskom intervalu (vremenski prozor) nakon čega se podaci prikupljeni u vremenskom prozoru obrađuju kako bi se izračunale značajke na temelju kojih se određuje mod prijevoza (Wang et al., 2019). Ovisno o primijenjenoj vrsti, vremenski prozori se promatraju zasebno ili su u međuovisnosti sa slijednim vremenskim prozorima. Neovisno o kojem se pristupu klasifikacije radi, prethodno se putovanje koje generira jedan korisnik razdvaja na segmente koji sadrže samo jedan mod prijevoza (segmentiranje trajektorije).

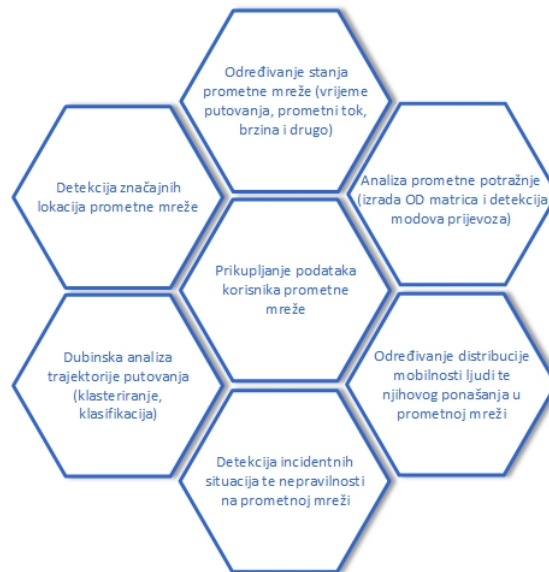
U ovom radu opisan je problem detekcije modova prijevoza iz korisničkih putovanja prikupljenih sa senzora za praćenje korisnika, metode rješavanja problema detekcije te vrste podataka koje se koriste u tu svrhu. U poglavlju 2 dan je pregled senzora čiji se podaci mogu koristiti za predviđanje prometnih parametara. Klasifikacija modova prijevoza detaljno je opisana u poglavlju 3 s posebnim naglaskom na značajke koje se koriste u procesu učenja klasifikatora za detekciju modova prijevoza te pregled metoda klasifikacije modova prijevoza. U poglavlju 4 dan je pregled metoda koje se koriste za segmentiranje trajektorije koja je preduvjet za postizanje veće točnosti klasifikacije. Zaključci rada kao i prijedlozi za buduća istraživanja dani su u poglavlju 5.

## 2. Obrada podataka kretanja korisnika prometnom mrežom i njihovo korištenje u analizi prometa

Prometna mreža predstavlja složeni sustav za čiju su analizu potrebni podaci korisnika prometne mreže. Ti podaci mogu se prikupljati iz više izvora, a dijele se na dvije osnovne vrste: senzori postavljeni na fiksnim lokacijama na prometnoj mreži te pokretni senzori. Podaci se prikupljaju korištenjem satelitskih sustava za određivanje položaja, navođenje i određivanje vremena, npr. GNSS sustava, globalnih sustava za mobilne komunikacije, npr. GSM sustava (engl. *Global System for Mobile Communications*, GSM) te raznovrsnih senzora s mobilnih uređaja. Također, podaci s drugih izvora se mogu koristiti za praćenje kretanja korisnika, primjerice podaci s društvenih mreža.

U današnje vrijeme broj GNSS uređaja je znatno veći u odnosu na prethodno desetljeće. Mnogi korisnici imaju navigacijske uređaje u automobilima te pametne telefone koji imaju ugrađen GNSS modul. Podaci koje generiraju korisnici uređaja s GNSS modulom tijekom korištenja prometnog sustava mogu se koristiti za prometne analize te zamijeniti podatke dobivene iz fiksnih senzora postavljenih na prometnoj mreži. Prometne analize nad takvim podacima su raznovrsne te su neka od područja primjene znanja dobivenih iz podataka praćenja kretanja korisnika prikazana na Slici 1 (Dabiri and Heaslip, 2018b).

Trenutno stanje prometne mreže može se odrediti analizom podataka o kretanju više korisnika na nekom području. Trenutno stanje prometne mreže predstavljaju vrijednosti relevantnih varijabli za opis stanja prometnog toka na nekom dijelu prometne mreže. Primjerice ako je na nekom cestovnom segmentu poznato ograničenje brzine, a trenutna brzina koja se prikuplja od praćenih vozila je znatno niža, onda se može zaključiti da je trenutno na promatranom cestovnom segmentu prisutno zagušenje. Mnoga istraživanja koriste prikupljene podatke za izračun budućeg stanja prometne mreže na temelju trenutnog stanja (de Fabritiis et al., 2008; Ma et al., 2017) ili izradu profila brzina (Erdelić et al., 2015). Posljedično, kada je poznato stanje prometne mreže, rute putovanja se određuju ovisno o predviđenom stanju prometne mreže (Rožić et al., 2015) ili se može predviđati odabir rute putovanja (Sun et al., 2014). Stanje prometne mreže na razini lokacije može osim zagušenja ukazivati na značajne lokacije na prometnoj mreži koje su usko povezane sa detekcijom incidentnih situacija te nepravilnosti na prometnoj mreži. U sva tri područja detektiraju se lokacije na mreži za koje je prostorno-vremenskom analizom utvrđeno da je na tim lokacijama veća frekvencija zapisa korisnika. Obično su zagušenije lokacije na kojima se odvija određeni događaj ili poslovne zone grada za vrijeme radnog vremena. Tako Andrienko et al. (2011) razvijaju proceduru za vizualnu analizu podataka o kretanju korisnika kako bi se detektirale značajne lokacije pomoću ponavljajućih događaja u tragovima kretanja korisnika. Kasnije se



Slika 1. Primjena znanja dobivenih iz podataka kretanja korisnika prometnom mrežom u prometnom sustavu (Dabiri and Heaslip, 2018b)

ove informacije mogu koristiti za predviđanje lokacija koje će u nekom periodu biti češće posjećene koristeći Markovljeve lance (Ashbrook and Starner, 2003). Incidentne situacije su jedan tip nepravilnosti na prometnoj mreži i mogu se detektirati pomoću GNSS podataka. Naime, na nekom cestovnom segmentu mogu se uočiti usporenja prometnog toka koja nisu uobičajena (ne očekuju se ponavljajuća zagušenja na tom cestovnom segmentu u tom periodu dana) što može ukazivati na incidentnu situaciju (D'Andrea and Marcelloni, 2017). U detekciji incidentnih situacija primjenjuju se i GSM podaci (Steenbruggen et al., 2011). Vrlo često se promatra vrijeme provedeno na području pokrivanja jedne bazne stanice, prije pojave preuzimanja poziva susjedne bazne postaje (engl. *handover*). Ako je promatrano vrijeme duže od uobičajenog ono može ukazivati na višu razinu zagušenja koja je uzrokovana incidentnom situacijom na području promatrane bazne postaje (Pattara-atikom and Peachavanish, 2007).

Podaci kretanja korisnika se mogu promatrati i na razini trajektorije, tj. putanje kretanja korisnika. U tom slučaju trajektorija opisuje mobilnost korisnika prometnom mrežom. Dubinska analiza trajektorije putovanja, koja može biti podržana metodama klasteriranja ili metodama klasifikacije, uključuje promatranje kretanja korisnika prometne mreže na razini trajektorije. Klasteriranjem se određuje sličnost između odabranih trajektorija, pa se mogu predvidjeti buduća kretanja neke grupe korisnika prometnom mrežom. Klasifikacija se također može koristiti u kontekstu predviđanja budućih mjesta koje će korisnik posjetiti temeljem do sada posjećenih mjesta u trajektoriji (Lee et al., 2011). Kod analiza ovog tipa vrlo često se primjenjuje metoda asocijacije (Cao et al., 2005). Određivanje distribucije mobilnosti ljudi te njihova ponašanja primjena je u kojoj se ponašanje korisnika kroz vrijeme vrlo brzo mijenja. Dakle, trajektorije korisnika je potrebno staviti u kontekstualni okvir. Primjerice kako bi se detektirala potražnja korisnika za autobusnim prijevozom na nekoj relaciji jedna od analiza je analiza izbora moda putovanja. Tako je uostalom odgovoreno na pitanje zašto korisnici na nekoj relaciji biraju određeni mod prijevoza (Thomas et al., 2018). Promatranje mobilnosti vlasnika automobila koristi se za određivanje cijene police osiguranja, tzv. način plaćanja *plati kako voziš* (engl. *Pay How you Drive*, PHD) u kojem se cijena police automobilskog osiguranja određuje na temelju GNSS trajektorije promatranog vozača. Iz GNSS trajektorija se izračunaju značajke koje dobro opisuju faktor rizika vozača, koji onda postaje jedan od parametara za izračun police osiguranja (Bian et al., 2018). Također, za svaki prometni model potrebno je definirati prometnu potražnju koja se izražava kroz više modova prijevoza: automobil, pješak, biciklist, autobus, vlak, tramvaj i slično. Kako bi se iz GNSS trajektorija izračunala prometna potražnja za pojedini mod prijevoza, prethodno je potrebno detektirati kojim modom prijevoza se korisnik kretao tijekom prikupljanja GNSS tragova. Za detekciju modova prijevoza prikupljene GNSS trajektorije se obrađuju i klasificiraju, pri čemu je izlazni parametar klasifikatora korišten mod prijevoza. Detekcija modova prijevoza koja se može uostalom koristiti i za dobivanje prometne potražnje za pojedini mod prijevoza je tema ovog rada. Prometna potražnja se također definira određeno izvorišnim matricama (engl. origin destination matrix, OD). Krajnji cilj OD matrica je izračun prosječnog

broja putovanja između svaka dva para prometnih zona na mreži tijekom nekog perioda. Za izračun OD matrica vrlo često se koriste GSM podaci (Iqbal et al., 2014).

Osim raznovrsnih primjena podataka kretanja korisnika u prometu, raznovrsni su i izvori s kojih se podaci prikupljaju. Ovisno o problemu koji se proučava, zbog specifičnosti informacija koje prikupljeni podaci sadrže, neke tehnologije prikupljanja podataka su više primjenjive od drugih. U sljedećim poglavljima opisani su ukratko sustavi za prikupljanje prometnih podataka korištenjem GNSS, GSM i Bluetooth tehnologija te njihova primjena u prometnoj analizi s posebnim naglaskom na detekciju modova prijevoza.

### 2.1. GNSS podaci kao izvor podataka

GNSS kao osjetilo prikuplja podatke o lokaciji (geografskoj širini, dužini, visini i trenutnom vremenu) iz čega se izračunava brzina korisnika. Na temelju udaljenosti GNSS prijavnika na Zemlji od satelita određuje se pozicija prijavnika, pa posljedično i njegova brzina. GNSS prijavnici, radi mogućnosti višestrukog odbijanja elektromagnetskih valova, rade s manjom točnošću u gradskim sredinama gdje su okruženi visokim zgradama te u zatvorenim prostorima (Lin and Hsu, 2014). GNSS podaci prikupljeni od korisnika prometne mreže mogu se koristiti za različite analize prometnog sustava jer se njima dobiva uvid u kretanje motoriziranog prometa (prijavnici ugrađeni u vozila), ali i nemotoriziranog prometa, kao što su biciklisti i pješaci (telefonski uređaj kao GNSS prijemnik).

Kada se govori o GNSS podacima koji se prikupljaju za detekciju modova prijevoza uglavnom se misli na prikupljanje podataka s mobilnih uređaja. U tom slučaju prikupljati se mogu samo GNSS podaci ili i podaci s drugih senzora poput inercijalno mjernog uređaja sastavljenog od akcelerometara za mjerenje linearnog odnosno žiroskopa za mjerenje rotacionog gibanja. Često se koristi jedan od ta dva skupa podataka ili oba zajedno. Niže su navedeni radovi koji podatke zasnivaju na jednom od navedenih skupova podataka.

#### 2.1.1. Podaci zasnovani samo na GNSS-u

Nekoliko istraživanja koristi samo GNSS podatke za detekciju moda prijevoza (Gonzalez et al., 2010; Xiao et al., 2015; Bantis and Haworth, 2017; Dabiri and Heaslip, 2018a). U tim istraživanjima značajke za klasifikaciju temelje se isključivo na podacima dobivenim iz GNSS zapisa poput brzine i udaljenosti. Dodatno, u nekim je istraživanjima izračunata akceleracija iz podataka prikupljenih GNSS-om. Kako bi se izbjegao nedostatak relevantnih značajki, neki autori uključuju geografski relevantne podatke, poput stajališta autobusa ili vlakova korištenjem geografskog informacijskog sustava (engl. Geographic Information System, GIS) (Stenneth et al., 2011; Gong et al., 2012; Das et al., 2014).

#### 2.1.2. Podaci zasnovani samo na akcelerometru

Kada se podaci prikupljaju s mobilnih uređaja, mogu se bilježiti samo podaci s akcelerometra. Često autori ističu da se izbjegavanjem prikupljanja GNSS podataka postižu značajne uštede baterije na mobilnom uređaju (Hemminki et al., 2013). Podaci s akcelerometra se prikupljaju kontinuirano tijekom putovanja korisnika, a značajke se uglavnom zasnivaju podacima prikupljenim u vremenskom i frekvencijskom području (Figo et al., 2010). Mjerenja dobivena iz akcelerometra odnose se na akceleraciju u smjeru osi  $x$ ,  $y$  i  $z$ , pa se značajke izračunavaju zasebno za svaku os (Wang et al., 2010; Yu et al., 2014; Hemminki et al., 2013; Shafique and Hato, 2015b).

#### 2.1.3. Podaci zasnovani na GNSS-u i akcelerometru

Velik broj istraživanja u području klasifikacije modova prijevoza koristi združivanje podataka s GNSS-a i akcelerometra (Reddy et al., 2010; Zhang et al., 2012; Nitsche et al., 2014; Byon and Liang, 2014; Ansarilari and Golroo, 2015). Podaci s akcelerometra se prikupljaju između slijednih GNSS lokacija, te se izračunavaju značajke temeljene na mjerenjima GNSS-a i akcelerometra. Potreba za oba podatka u klasifikaciji modova prijevoza razvija se u situacijama u kojima su izlazne vrijednosti akcelerometra slične, ali najčešće je tada brzina različita i obrnuto. Kada se koristi samo jedan od ova dva podatka, pad točnosti klasifikacije je između 10-20% (Reddy et al., 2010).

### 2.2. GSM podaci kao izvor podataka

Za uspostavljanje poziva i prijenos podataka u mobilnoj mreži koriste se mobilni uređaji. Prvi opće prihvaćeni i zato najrasprostranjeniji standard za mobilnu mrežu u svijetu je GSM. GSM standard se zasniva na baznim stanicama čiji radiokomunikacijski sustav određuje domet pokrivanja stanice, tj. ćeliju koja je zahvaćena tom baznom stanicom,

a svaka bazna stanica je lokacijski određena (svaka ćelija je zapravo jedan poligon koji ima svoje geografske koordinate). Dakle, svaki korisnik mobilne mreže je u nekom trenutku prijavljen na baznu stanicu u čijem se području pokrivanja nalazi. Prelasci korisnika s jedne bazne stanice u drugu, tj. slijed jedinstvenih identifikatora baznih stanica te jačina signala mogu se pratiti te se tako može locirati mobilni uređaj. Ako se poveže slijed prelazaka između baznih stanica tada se može rekonstruirati trajektorija jednog korisnika prometnom mrežom. Osim toga, podaci se mogu koristiti i za klasifikaciju modova prijevoza. [Sohn et al. \(2006\)](#) primijenili su GSM podatke na klasifikaciju kretanja korisnika kroz 3 stanja: stajanje, hodanje i vožnja. Iste klase koriste [Mun et al. \(2008\)](#), no uz GSM podatke koriste i podatke bežične mreže u slučaju njezine dostupnosti, na primjer dijelovi prometne mreže s javno dostupnom bežičnom mrežom, opskrbna središta, rekreacijska i slična područja. Podatke bežične mreže koriste i [Montoya et al. \(2015\)](#) za klasifikaciju modova prijevoza. Ipak, zastupljenost podataka prikupljenih iz mobilnih mreža u prometnim analizama je veća kod određivanja OD matrica.

### 2.3. Bluetooth podaci kao izvor podataka

Podaci prikupljeni iz bežične Bluetooth mreže nisu zastupljeni u klasifikaciji modova prijevoza. Razlog je mala zastupljenost Bluetooth senzora u vanjskom okruženju, pa se podaci mogu prikupljati eventualno s mobilnih uređaja. Takvi podaci pokazuju nedostatak u raspoznavanju kretanja korisnika ili okoline koja ga okružuje ([Reddy et al., 2010](#)), dok u zatvorenim prostorima Bluetooth senzori pronalaze svoju primjenu jer je u tim slučajevima GNSS često nedostupan i neprecizan ([Madhavapeddy and Tse, 2005](#); [Aparicio et al., 2008](#)). U situacijama kada je prometna mreža pokrivena s fiksnim Bluetooth sensorima, moguće je detektirati mod prijevoza korisnika koji imaju uređaj koji odašilje Bluetooth signal ([Yang and Wu, 2017](#)).

## 3. Klasifikacija modova prijevoza

Podaci kretanja korisnika sadrže informaciju o prometnoj potražnji na nekom dijelu prometne mreže. Prometna potražnja je različita za svaki mod prijevoza, a informacija o korištenom modu prijevoza nije sadržana u podacima koji se automatski prikupljaju. Stoga je informaciju o modu prijevoza potrebno odvojeno prikupljati ili pronaći metodu za detekciju moda prijevoza iz podataka. Dugo je u praksi primjenjivana metoda prikupljanja podataka o putovanju korisnika prometne mreže ispitivanjem korisnika kroz anketne obrasce o dnevnim migracijama. Ovaj pristup prikupljanja podataka uključuje i informaciju o korištenom modu prijevoza, pa daljnje metode utvrđivanja načina prijevoza odnosno prijevoznog sredstva nisu potrebne. Kroz godine je ovaj pristup pokazao niz nedostataka, poput netočnih informacija zbog lošeg sjećanja korisnika ili nedovoljni odaziv korisnika koji žele dati informacije o svojim dnevnim migracijama. Stoga se sve češće detektira mod prijevoza iz podataka kretanja korisnika. Za detekciju moda prijevoza koriste se metode klasifikacije trajektorije, pri čemu se klasificira dio trajektorije u kojem je korišten jedan mod prijevoza. Danas se pri tome koriste sustavi strojnog učenja. U fazi učenja se pomoću značajki gradi klasifikator koji će što bolje razlikovati modove prijevoza. Značajke se izračunavaju iz trajektorija svakog korisnika. Klasifikator se nakon učenja testira, te se zatim može koristiti na podacima kojima nije poznat mod prijevoza za detekciju moda prijevoza. U cijelom procesu je vrlo bitna faza učenja klasifikatora te odabir značajki kojima se opisuje trajektorija. Dakle, značajke su osnovni parametri na temelju kojih klasifikator razlikuje modove prijevoza jedan od drugoga, zato je bitno da odabrani skup značajki što bolje opisuje problem. Značajke se mogu odrediti njihovim odabirom od strane čovjeka korištenjem metoda i tehnika strojnog učenja ili samostalnim odabirom od strane sustava za učenje zasnovanog na umjetnim neuronskim mrežama. Klasifikacija se može provoditi nad već prikupljenim podacima (povijesni podaci) ili tijekom prikupljanja podataka u stvarnom vremenu. U nastavku su opisane relevantne značajke za problem klasifikacije modova prijevoza te je dan pregled klasifikacijskih metoda nad povijesnim podacima i u stvarnom vremenu.

### 3.1. Odabir relevantnih značajki za klasifikaciju modova prijevoza

Osnovni cilj odabira relevantnih značajki je što realističniji opis ponašanja uzorka koji se klasificira kako bi klasifikator strukturom što točnije opisivao klase obuhvaćene u klasifikaciji. U mnogim istraživanjima je prihvaćeno da su značajke zasnovane na brzini vrlo bitne za klasifikaciju modova prijevoza ([Reddy et al., 2010](#); [Shen and Stopher, 2014](#)). Osim brzine, vrlo često se uzima u obzir i akceleracija ([Nick et al., 2010](#)). Naravno, značajke koje se uzimaju u obzir za klasifikaciju ovise i o tehnologiji koja se koristi za prikupljanje podataka. Autori [Reddy et al.](#)

(2010) su usporedili točnost klasifikacije u slučajevima primjene podataka prikupljenih s različitih senzora mobilnog uređaja: Bluetooth, GSM i WiFi. Bluetooth s mobilnog uređaja daje lošije rezultate u odnosu na ostale podatke jer je na temelju Bluetooth signala teško detektirati kreće li se korisnik ili se okolina mijenja (drugi ljudi koji nose uređaje se kreću). Usporedbom GSM značajki s značajkama dobivenim iz sustava za globalno pozicioniranje (engl. *Global Positioning System*, GPS) s uključenim podacima s akcelerometra autori su dobili 22% veću točnost koristeći GPS podatke. Uključivanje dodatnih podataka poput GSM-a i radijusa poslužitelja neće značajno doprinijeti točnosti, a utječe na veću potrošnju baterije mobitela. [Dabiri and Heaslip \(2018a\)](#) dodaju da korištenje višestrukih senzora može onemogućiti primjenu klasifikatora u nekim gradovima zbog smanjene dostupnosti podataka. Dio istraživanja GNSS podatke dopunjuje geoprostornim podacima kao npr. GIS podacima i obradom u programskom okruženju geoinformacijskog sustava kako bi se povećala točnost klasifikacije modova prijevoza koje karakteriziraju fiksne lokacije na mreži ([Nour et al., 2016](#)). Geoprostorni podaci koji se koriste u klasifikaciji modova prijevoza najčešće uključuju značajke poput udaljenosti do stanice autobusa i stanice vlaka ([Stenneth et al., 2011](#)) ili dopunu podataka za podzemnu željeznicu gdje nema GPS signala ([Gong et al., 2012](#)).

U istraživanjima je najčešće korišten vlastiti skup podataka. Ipak, neka istraživanja koriste objavljene skupove podataka od kojih su najpoznatiji GeoLife ([Zheng et al., 2010b](#)) skup podataka od 9043 putovanja pri čemu su prikupljeni GPS podaci; i Susex-Huawei skup podataka kojeg čine 2800 sati prikupljenih podataka kretanja korisnika s 16 senzora s mobilnog uređaja ([Gjoreski et al., 2018](#)).

Neke od najčešćih značajki u radovima su prosječna vrijednost, standardna devijacija, mod, medijan, maksimalna i minimalna vrijednost brzine i akceleracije ([Xiao et al., 2017](#); [Gong et al., 2018](#)). Međutim, zbog promjenjivog ponašanja brzine prometnog toka tijekom dana ([Erdelić et al., 2015](#)) vozilo može imati jednaku maksimalnu brzinu kao i pješak ili biciklist. Jedno rješenje problema je uvođenje novih značajki poput prvih 5 najvećih vrijednosti brzine umjesto maksimalne brzine. Ipak, kod dodavanja novih značajki mora se efikasno upravljati s dimenzionalnosti podataka ([Dabiri and Heaslip, 2018a](#)). [Zheng and Xie \(2008\)](#) prvi uvode osnovne značajke poput prijedene udaljenosti, brzine i akceleracije. Kasnije, [Zheng et al. \(2008\)](#) uvode dodatne značajke: stupanj promjene brzine, stupanj promjene smjera te stupanj stajanja u ruti. [Eftekhari and Ghatee \(2016\)](#) koriste značajke dobivene na temelju  $L^2$ -Norm vektora žiroskopa po  $x$ ,  $y$ ,  $z$  osi za detekciju kretanja u ne motoriziranim modovima prijevoza te kako bi poboljšali rezultate detekcije stajanja. [Gong et al. \(2018\)](#) razmatraju dodatne značajke koje se odnose na meteorološko vrijeme, poput prosječne temperature ili temperature zraka na početku i kraju putovanja. Autori uključuju značajke tog tipa jer razmatraju i svrhu putovanja. Dio istraživanja uključuje demografske podatke (spol, godine, vozačka dozvola i slično) kao i geoprostorne podatke ([Gong et al., 2012](#)). Osim uobičajenih statističkih mjera [Feng and Timmermans \(2013\)](#) koriste i značajke poput vlasnik bicikla, automobila ili motora (binarna varijabla s vrijednostima  $Da$  ili  $Ne$ ), broj dostupnih satelita te pogrešku uzorkovanja po prijedenoj udaljenosti. [Etemad et al. \(2018\)](#) koriste značajke promjene kuta u odnosu na sjever, pri čemu promatraju promjenu kuta u vremenu te učestalost takvih promjena. [Jahangiri and Rakha \(2015\)](#) promatraju distribuciju značajki za detekciju značajnijih promjena u uzorcima (spektralna entropija). Isti autori u radu [Jahangiri and Rakha \(2014\)](#) promatraju četiri skupa atributa s različitih senzora: akcelerometra, žiroskopa, GPS-a te vektor rotacije, s izračunatim osnovnim statističkim mjerama u sve tri osi. Stupanj promjene akceleracije prvi put koriste [Dabiri and Heaslip \(2018a\)](#). Dodatne značajke poput temperature tijela, otkucaja srca te intenziteta svjetlosti koriste [Parkka et al. \(2006\)](#).

Kako bi se izbjegao gubitak podataka izostavljanjem bitnih značajki, dio istraživanja koristi sirove podatke za klasifikaciju pri čemu se metodama dubokog učenja automatski izračunavaju značajke iz sirovih podataka trajektorije ([Endo et al., 2016](#)).

### 3.2. Metode klasifikacije modova prijevoza na povijesnim podacima

Svrha razvoja metoda klasifikacije modova prijevoza je automatizacija prikupljanja dnevnika putovanja. Uz dovoljno točan klasifikator, podaci kretanja korisnika se mogu automatski razvrstati prema modu prijevoza te nije potrebno vođenje dnevnika putovanja od strane korisnika. Prva istraživanja koja proučavaju automatsko prikupljanje dnevnika putovanja su [Wolf \(2000\)](#), [Schöenfelder et al. \(2002\)](#) i [Axhausen et al. \(2003\)](#). Podaci su prikupljeni GPS prijamlnikom ugrađenim u vozilo, što isključuje velik broj modova prijevoza. Stoga se istraživanja usmjeravaju prema prikupljanju podataka uređajima izvan vozila. Jedno od prvih istraživanje tog tipa provode [Draijer et al. \(2000\)](#) pri čemu koriste GPS prijamlnike nošene u ruci. Naglasak istraživanja je, obzirom na težinu dostupnih uređaja, na teretu nošenja uređaja te pogrešci uređaja. Kasnije se razvijaju prilagođene platforme za prikupljanje podataka kretanja korisnika prometnom mrežom, ali se ne razmatraju metode klasifikacije ([Bao and Intille, 2004](#); [Chung and Shalaby,](#)

2005; Consolvo et al., 2008; Brezmes et al., 2009; Bieber et al., 2009; Bolbol et al., 2012). Nakon toga su istraživanja usmjerena ka razvoju klasifikatora i značajki kako bi točnost klasifikacije bila što veća.

Klasifikacija modova prijevoza može se podijeliti na dva osnovna zadatka: i) određivanje postoji li kretanje korisnika te u slučaju da se korisnik kreće ii) detekcija prijevoznog sredstva koje korisnik koristi za kretanje. Broj modova prijevoza uključenih u klasifikaciju, razmatran u dostupnoj literaturi, nije uniformiran i varira između 2 i 6. Klasifikacija 2 moda prijevoza najčešće uključuje klasifikaciju na motorizirane i ne motorizirane modove prijevoza, dok klasifikacija više modova prijevoza uključuje razlikovanje između motoriziranih modova prijevoza. Hemminki et al. (2013) navode da klasifikacija između motoriziranih modova prijevoza omogućava detaljnije informiranje o ponašanju korisnika u prometu. Motorizirani modovi prijevoza uključeni u klasifikaciju su najčešće: automobil, autobus, tramvaj i vlak. Novija istraživanja, Wang et al. (2018) uključuju električni bicikl kao mod prijevoza. Većina istraživanja kod razvoja klasifikatora dijeli prikupljeni skup podataka na dva dijela: skup podataka za učenje te skup podataka za testiranje klasifikatora. Shafique and Hato (2015a) su testirali utjecaj dva načina formiranja skupa podataka za učenje i testiranje na razvoj klasifikatora za detekciju moda prijevoza. Autori pokazuju da podjela podataka svih praćenih korisnika na slučajan način daje znatno veću točnost nego korištenje skupa podataka od jednog dijela korisnika za učenje, a drugog dijela korisnika za testiranje.

Do sada su primijenjene različite metode za klasifikaciju modova prijevoza. Stabla odluke (engl. *Decision Tree*, DT), algoritam  $k$  najbližih susjeda (engl. *k Nearest Neighbor*, kNN), naivan Bayesov klasifikator (engl. *Naive Bayes*, NB) te algoritam slučajnih šuma (engl. *Random Forest*, RF) su najčešće korištene metode strojnog učenja. Metode koje se još spominju u literaturi su neizrazita logika (engl. *Fuzzy Logic*, FL) (Biljecki et al., 2013), metoda potpornih vektora (engl. *Support Vector Machine*, SVM) (Pereira et al., 2013) te metode zasnovane na samostalnom određivanju značajki učenjem, kao što su neuronske mreže (engl. *Neural Network*, NN) (Byon et al., 2007; Gonzalez et al., 2010), duboke neuronske mreže (engl. *Deep Neural Network*, DNN) (Wang et al., 2017), konvolucijske neuronske mreže (engl. *Convolutional Neural Network*, CNN) (Dabiri and Heaslip, 2018a). Također razvijeno je nekoliko pristupa koji koriste združene (engl. *ensemble*) klasifikatore, višeslojne klasifikatore ili obradu rezultata nakon klasifikacije. Kod združenih klasifikatora najčešće se koriste AdaBoost (Yu et al., 2014; Xiao et al., 2017) i RF (Jahangiri and Rakha, 2015; Shafique and Hato, 2015b; Sonderen, 2016; Xiao et al., 2017; Stenneth et al., 2011) te se korištenjem takve kombinacije postižu bolji rezultati klasifikacije. Višeslojna klasifikacija se odnosi na dvije razdvojene faze odlučivanja kako bi se došlo do konačne odluke. Prvo se provodi odlučivanje kreće li se korisnik motoriziranim ili nemotoriziranim modom prijevoza. Ako je u prvoj fazi detektiran motorizirani mod prijevoza, u drugoj fazi se odlučuje o kojem se motoriziranom modu prijevoza radi (Yu et al., 2014; Hemminki et al., 2013; Siirtola and Röning, 2012). Obrada nakon klasifikacije odnosi se na proučavanje korelacije između susjednih vremenskih prozora (Yu et al., 2014) ili primjenu skrivenih Markovljevih lanaca (engl. *Hidden Markov Model*, HMM) za dobivanje vjerojatnosti tranzicije iz jednog moda prijevoza u drugi (Feng and Timmermans, 2013; Reddy et al., 2010; Nitsche et al., 2014). U Tablici 1 navedeno je 8 radova koji provode klasifikaciju modova prijevoza na povijesnim podacima te je u nastavku svaki od radova detaljnije opisan.

Tablica 1. Istraživanja u području klasifikacije modova prijevoza na povijesnim podacima

Autor	Podaci		Metoda klasifikacije	Modovi prijevoza	Točnost
	Senzori	Veličina uzorka			
Gonzalez et al. (2010)	GPS	114 uzoraka	NN	automobil, pješak i autobus	88,6 %
Gong et al. (2012)	GPS i GIS	63 volontera	FL	pješak, vlak, automobil i autobus	82,6%
Zhang et al. (2012)	GPS i akcelerometar	125 trajektorija	FL i SVM	pješak, bicikl, automobil, autobus, tramvaj i vlak	92,59%
Byon and Liang (2014)	GPS, akcelerometar i magnetometar	5 volontera, 100 sati	NN	automobil, autobus, bicikl, hodanje	-
Xiao et al. (2017)	GPS	GeoLife	RF, DT i XGBoost	pješak, bicikl, automobil, vlak i tramvaj	90,77%
Bantis and Haworth (2017)	GPS	5 volontera	Bayesova mreža	stajanje, pješak, autobus, vlak	78%
Gong et al. (2018)	GPS i podaci o vremenu	20 volontera tijekom 10 mjeseci	RF	automobil, autobus, tramvaj i pješak	20% - 96%
Dabiri and Heaslip (2018a)	GPS	GeoLife	CNN	pješak, bicikl, autobus, automobil i vlak	84,8%

**Gonzalez et al. (2010)** koriste NN za automatsku detekciju moda prijevoza. Podaci korišteni u istraživanju prikupljeni su s dva tipa mobilnih uređaja na koje je instalirana aplikacija kojom se prikupljaju podaci o kretanju korisnika s uzorkovanjem svake 4 sekunde. Podatak o stvarno korištenom modu je unesen ručno kako bi se podaci mogli koristiti za učenje i testiranje mreže. Promatrana su tri moda prijevoza: automobil, autobus i pješak. U procesu učenja i testiranja mreže koriste  $k$ -struku unakrsnu provjeru (engl. *k-folded cross validation*), a vrijednost parametra  $k$  je 10. Značajke korištene za učenje su: prosječna i maksimalna brzina, točnost GPS signala, postotak lokacija izračunatih pomoću bazne stanice, standardna devijacija udaljenosti između stajanja te prosječno vrijeme između dva stajanja. Kako bi smanjili količinu podataka koja se prenosi prema poslužitelju, autori koriste algoritam kritičnih točaka kojim detektiraju i odbacuju podatke koji ne donose novu informaciju. Time se smanjuje trošak prijenosa podataka, a i štedi baterija na mobilnom uređaju. Slučajevi u kojima se primjenjuje algoritam kritičnih točaka su primjerice pogreška GPS-a ili podaci koji ne donose novu informaciju. Pogreška GPS-a se događa u zatvorenim prostorima primjerice autobusu (metalna kutija), tijekom koje nastaje niz višestrukih zapisa na istoj lokaciji. Podaci koji ne donose novu informaciju mogu biti generirani tijekom stajanja korisnika na jednoj lokaciji ili kod pravocrtnog kretanja prometnom mrežom. Kod stajanja korisnika bilježi se jedna lokacija više puta, a bitnu informaciju donose samo trenutak početka i kraja stajanja. Isto se može primijeniti na pravocrtno kretanje korisnika, pri čemu prva i posljednja GPS točka, prije promjene smjera kretanja, daju dovoljnu informaciju o kretanju korisnika, a sve ostale točke stvaraju zalihost. Za algoritam kritičnih točaka autori koriste odvojene značajke: prosječna i maksimalna akceleracija, prosječna i maksimalna brzina, broj kritičnih točaka po udaljenosti i vremenu, ukupna udaljenost te prosječna udaljenost između kritičnih točaka. Autori postižu točnost 88,6% kod promatranja svih GPS točaka, a 91,23% kod promatranja samo kritičnih točaka.

**Zhang et al. (2012)** dijele proces klasifikacije u dva dijela. U prvom dijelu autori dodjeljuju svakoj trajektoriji korisnika jednu od tri moguće klase, pravilima kreiranim neizravnom logikom: i) pješak, ii) bicikl ili iii) motorizirani mod prijevoza. Sve trajektorije kojima je nakon prvog dijela klasifikacije detektiran motorizirani mod prijevoza ulaze u drugi dio klasifikacije, gdje se SVM metodom svakoj trajektoriji dodjeljuje jedna od klasa: i) automobil, ii) autobus, iii) tramvaj ili iv) vlak. Za prvi dio klasifikacije, uz standardne statističke mjere, autori uvode specifične značajke kojima će bolje prikazati razliku između modova. Za razlikovanje pješaka i biciklista opisuju trajektoriju polinomom drugog stupnja, a zatim koriste koeficijent polinoma kao značajku, gdje obično za hodanje vrijedi da je koeficijent konstantan. Za razlikovanje bicikla i motoriziranog prijevoza korištena je maksimalna akceleracija. Autori postižu točnost 96,95% u prvom dijelu klasifikacije, a 92,59% u drugom dijelu klasifikacije.

**Byon and Liang (2014)** koriste NN za klasifikaciju između različitih modova prijevoza. Cilj rada je usporediti točnost klasifikacije NN na podacima prikupljenim sa standardnih GPS uređaja i podacima s akcelerometra i magnetometra ugrađenih u mobilni uređaj. Rezultati klasifikacije pokazuju da dostupnost izmjerenih vrijednosti ubrzanja i elektromagnetskih očitavanja s mobilnih uređaja poboljšava točnost. Osim toga autori prikazuju utjecaj različitih parametara na točnost klasifikacije: kraće vrijeme uzorkovanja te detekcija moda prijevoza tijekom zagušenja povećavaju točnost klasifikacije.

Prema iskustvenom radu **Hansen and Salamon (1990)** i teorijskom radu **Schapire (1990)**, pokazano je da kombinacija skupa klasifikatora često daje bolje rezultate nego jedan klasifikator samostalno. Stoga, **Xiao et al. (2017)** koriste kombinaciju više klasifikatora kako bi poboljšali rezultate klasifikacije. Korišteni su klasifikatori RF, Gradient Boosting DT i XGBoost. Autori koriste GeoLife skup podataka za učenje i testiranje klasifikatora. Na temelju podataka izračunate su 72 globalne i 39 lokalnih značajki. Globalne značajke se odnose na cijelu trajektoriju, dok se lokalne značajke odnose na dio trajektorije kojim je korišten jedan mod prijevoza. Odabrane klasifikatore autori uspoređuju s kNN metodom, SVM i DT klasifikatorom, pri čemu je najveća točnost od 90,77% postignuta XGBoost klasifikatorom.

Klasifikacija modova prijevoza uglavnom uključuje prostorno-vremenske podatke o kretanju korisnika, no takvoj dimenziji mogu se dodati i društveno-demografski podaci. Tada se problem klasifikacije modova prijevoza može promatrati individualno ovisno o ponašanju pojedinog korisnika. **Bantis and Haworth (2017)** uzimaju u obzir zdravstveno stanje korisnika prometne mreže, pa tako razlikuju prikupljene podatke od korisnika u invalidskim kolicima i korisnika koji koriste pomoćno medicinsko pomagalo. Autori koriste Bayesovu mrežu za klasifikaciju te promatraju rezultate ovisno o vrsti korisnika. Točnost klasifikacije za korisnike s medicinskim pomagalom je 78%, a korisnika u invalidskim kolicima 71%.

Osim GPS i GIS podataka, **Gong et al. (2018)** uzimaju u obzir demografske podatke i podatke o meteorološkom vremenu. Autori koriste skup podataka koji je prikupljen od 20 volontera tijekom 5 zimskih i 5 ljetnih mjeseci, pa tako i dijele skup podataka za testiranje. Testiraju korišteni RF klasifikator pomoću tri skupa podataka: ljetni

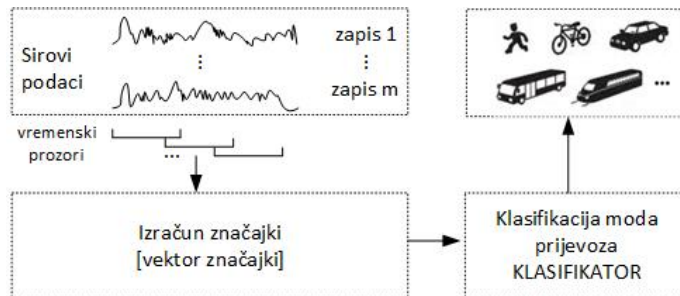


skup podataka, zimski skup podataka te skup podataka koji sadrži oba razdoblja. Osim toga, autori promatraju svrhu putovanja svakog korisnika u svakom od testnih skupova podataka. Analizirana matrica konfuzije pokazuje točnost predikcije između 20% i 96%, pri čemu se najveća točnost odnosi na pješake, a najmanja točnost je zabilježena kod klasifikacije autobusa. Autori zaključuju da GIS podaci imaju pozitivan utjecaj na klasifikaciju, dok podaci o vremenu imaju negativan utjecaj na klasifikaciju.

Konvolucijsku neuronsku mrežu koriste [Dabiri and Heaslip \(2018a\)](#) pri čemu koriste za učenje i testiranje mreže podatke prikupljene u sklopu GeoLife projekta. Za klasifikaciju modova prijevoza autori koriste sirove GPS podatke te postižu točnost od 84,8%.

### 3.3. Klasifikacija modova prijevoza u stvarnom vremenu

Krajnji cilj klasifikacije modova prijevoza je razvrstavanje prikupljenih trajektorija putovanja prema korištenim modovima prijevoza. Podaci koji se koriste za učenje klasifikatora moraju uz prikupljene podatke sa senzora sadržavati podatak o korištenom modu prijevoza. Vrlo često tijekom prikupljanja podataka sa senzora praćeni korisnik paralelno vodi dnevnik u kojem bilježi korištene modove prijevoza prema vremenskim intervalima. Kako bi se taj proces automatizirao, u nekim istraživanjima izrađene su aplikacije na mobilnim uređajima koje u pozadini prikupljaju podatke s GPS-a ili akcelerometra ili oboje, a na grafičkom sučelju korisnik može odabrati trenutni mod prijevoza. Tako su razvijene aplikacije Future Mobility Survey ([Cottrill et al., 2013](#)), MEILI ([Prelipcean et al., 2014](#)), Move Smarter ([Geurs et al., 2015](#)) te Peacox ([Montini et al., 2015](#)). Tijekom prikupljanja podaci se klasificiraju prema samostalno određenom modu prijevoza bez obzira na korisničku potvrdu, a u slučaju korisničke potvrde rezultat klasifikacije se uspoređuje sa stvarnim modom prijevoza. Proces klasifikacije modova prijevoza u stvarnom vremenu prikazan je na Slici 2. Sirovi podaci prikupljaju se s različitih senzora tijekom trajanja jednog vremenskog prozora. Vektor značajki izračunava se za svaki vremenski prozor te predstavlja ulazne podatke klasifikatora. Izlaz klasifikatora je jedan od razmatranih modova prijevoza. Osnovna razlika ovog pristupa klasifikaciji u odnosu na klasifikaciju povijesnog skupa podataka je primjena vremenskih prozora. Značajke koje opisuju trajektoriju izračunavaju se nakon isteka definiranoga vremenskog intervala (vremenski prozor), a ne za cijelu trajektoriju jednog korisnika. Vremenski prozori mogu se promatrati svaki samostalno ili se dodatno obrađuje tranzicija između vremenskih prozora ([Reddy et al., 2008](#)).



Slika 2. Proces klasifikacije modova prijevoza u stvarnom vremenu

Prikupljanje podataka i klasifikacija modova prijevoza koristeći mobilne uređaje otvara novi problem potrošnje baterije. Tako, pojedini autori razmatraju utjecaj senzora na mobilnom uređaju na potrošnju baterije ([Reddy et al., 2010](#); [Gonzalez et al., 2010](#); [Hemminki et al., 2013](#); [Bloch et al., 2015](#); [Su et al., 2016](#); [Sonderer, 2016](#); [Prelipcean et al., 2018](#)). Podaci prikupljeni s mobilnih uređaja u svrhu detekcije modova prijevoza najčešće uključuju GPS podatke i podatke s akcelerometra. Mnoga istraživanja potvrđuju da GPS troši značajno više energije od akcelerometra ([Prelipcean et al., 2014](#)). Međutim, akcelerometar može dati podatke na temelju kojih se mod prijevoza može detektirati, ali tada korišteni mod prijevoza nije povezan s lokacijama na kojima je korišten detektirani mod prijevoza. Nedostatak prostorne komponente podatka može predstavljati problem za prometne analize koje su lokacijski zavisne. Stoga, se razvijaju optimizacijske metode kojima se regulira kada će koji senzor biti korišten, kako bi podaci što bolje opisivali stanje prometne mreže, a da je istovremeno što veća ušteda potrošnje energije ([Cottrill et al., 2013](#); [Prelipcean et al., 2014](#)). U Tablici 2 naveden je dio istraživanja koji proučava klasifikaciju modova prijevoza u stvarnom vremenu. Radovi navedeni u tablici opisani su u nastavku.

Tablica 2. Istraživanja u području klasifikacije modova prijevoza u stvarnom vremenu

Autor	Podaci		Metoda klasifikacije	Modovi prijevoza	Točnost
	Senzori	Veličina uzorka			
Reddy et al. (2010)	GPS i akcelerometar	120 sati	DT, DHMC	stajanje, pješak, trčanje, bicikl, motor	93, 7%
Stenneth et al. (2011)	GPS i GIS	6 volontera, 3 tjedna	NB, BN, DT, RF, ML	vlak, autobus, automobil, pješak, bicikl, stajanje	93, 70%
Hemminki et al. (2013)	akcelerometar	16 volontera, 150 sati	Hijerarhijska metoda	stajanje, pješak, autobus, vlak, tramvaj, metro	80, 1%
Nitsche et al. (2014)	GPS i akcelerometar	15 volontera 2 mjeseca	Združena metoda, DHMM	pješak, bicikl, motor, automobil, autobus, tramvaj, metro i vlak	65% - 95%
Van Dijk (2018)	umjetno generirane GPS trajektorije	200 trajektorija	NB, boosted C5.0, SVM i RF	stajanje i kretanje	>99%
Prelipcean et al. (2018)	GPS i akcelerometar	51 korisnik 9 dana	kNN	pješak, bicikl, motor, automobil (vozač, putnik), autobus, tramvaj, vlak, katamaran, zrakoplov, taksi	75%

Reddy et al. (2010) uspoređuju točnost više klasifikatora: DT, algoritam  $k$  sredine (engl. *k Means Clustering*, kMC), NB, kNN, SVM i kontinuirani skriveni Markovljev model (engl. *Continuous Hidden Markov Model*, CHMM). Za klasifikaciju odabiru dvorazinski klasifikator koji je kombinacija diskretnog skrivenog Markovljevog modela (engl. *Discrete Hidden Markov Model*, DHMM) i jednog od uspoređenih klasifikatora koji daje najbolje rezultate. Klasifikator je razvijen za mobilni uređaj na kojem tijekom prikupljanja podataka DT na temelju izračunatih značajki predviđa mod prijevoza, a DHMM ispituje vjerojatnost tranzicije s obzirom na prethodni mod prijevoza. Prednost ovog pristupa je ispravak klasifikacije kada se neki od modova prijevoza ponaša izvan svojih prosječnih vrijednosti značajki. Primjerice kod trčanja i vožnje bicikla vrlo često se pojavljuje zamijenjena predviđanja (brzo trčanje, sporija vožnja na biciklu), no vjerojatnost tranzicije između ta dva moda prijevoza je mala, pa DHMM zaključuje da je mod prijevoza i dalje jednak.

Stenneth et al. (2011) koriste GPS i geoprostorne podatke. Za klasifikaciju koriste geoprostorne podatke u obliku zemljopisnih koordinata stajališta autobusa i vlaka te autobusnih i željezničkih linija, iz kojih izračunavaju značajke poput udaljenosti GPS lokacije trenutnog korisnika od fiksnih geoprostornih lokacija. Uz navedene značajke, koriste GPS značajke zasnovane na brzini i akceleraciji. Autori uspoređuju više metoda navedenih u Tablici 2 te najveću točnost klasifikacije postižu metodom RF.

Hemminki et al. (2013) koriste u procesu klasifikacije samo podatke s akcelerometra. Značajke grupiraju u tri odvojene kategorije ovisno o dužini uzorkovanja: kratko uzorkovanje (detektiraju se vršne vrijednosti akceleracije), uzorkovanje jednog okvira te uzorkovanje cijele trajektorije jednog moda prijevoza. Različite duljine uzorkovanja doprinose razlikovanju između motoriziranih modova prijevoza. Za klasifikaciju autori koriste hijerarhijsku metodu klasifikacije pri čemu prvo klasificiraju kretanje i stajanje u trajektoriji, a potom ako nije detektirano niti kretanje niti stajanje, prelazi se u treći dio klasifikacije u kojemu se klasificira između motoriziranih modova prijevoza.

Nitsche et al. (2014) koriste za klasifikaciju modova prijevoza metodu koja je zasnovana na združenoj metodi u kombinaciji s DHMM. Više klasifikatora predviđa mod prijevoza, a konačni mod prijevoza je mod prijevoza koji je imao najveći broj glasova svih klasifikatora. Zatim se DHMM metodom provjerava vjerojatnost tranzicije s prethodnog moda prijevoza. Klasifikacija je zasnovana na podacima prikupljenim s mobilnih uređaja što uključuje GPS podatke te podatke s akcelerometra. Autori rezultate prikazuju matricom konfuzije pri čemu je vidljiva samo točnost klasifikacije po pojedinom modu prijevoza. Najveća točnost klasifikacije je 95% za bicikl, dok je najlošiji rezultat 65% za podzemnu i nadzemnu željeznicu.

Van Dijk (2018) koristi umjetno generirane GPS trajektorije za prepoznavanje aktivnosti u ruti. Za prepoznavanje aktivnosti koristi metode NB, boosted C5.0, SVM i RF pri čemu razlikuje dvije aktivnosti (klase): stajanje i kretanje. Za usporedbu je korištena metoda zasnovana na pravilima koju su razvili autori u prethodnim istraživanjima. Primjerice Miluzzo et al. (2008) koriste algoritam zasnovan na pravilima kako bi detektirali sjedenje, stajanje, hodanje ili trčanje korisnika mobitela. Van Dijk (2018) za prvo pravilo primjenjuje pravilo koje su razvili Bohte and Maat (2009) u kojem koriste prag udaljenosti i vremena kako bi odvojili stajanje od kretanja. Drugo pravilo, određivanje moda stajanja, je preuzeto od Schuessler and Axhausen (2009a) u kojem se dio trajektorije klasificira kao stajanje ako je brzina manja od 3,6 km/h dulje od 120 sekundi, a zatim se promatra gustoća točaka na nekom području. Za treće pravilo prilagođeno je drugo pravilo tako da se detektira stajanje ako je udaljenost od najbliže prometnice veća od 10 metara i

trajanje duže od 120 sekundi. Klasifikatori su testirani na podacima koji su zabilježeni u vremenskim prozorima fiksne duljine: 30 i 60 sekundi. Rezultati pokazuju značajno bolje rezultate gore navedenim metodama klasifikacije (točnost veća od 99%) nego primjenom pravila (točnost između 54% i 77%). Također, vremenski prozor od 30 sekundi daje nešto bolje rezultate u odnosu na vremenski prozor duljine 60 sekundi.

Prelipcean et al. (2014) razvijaju aplikaciju za mobilne uređaje koja bilježi podatke o GPS trajektoriji korisnika i akceleraciji. Tijekom prikupljanja podataka u stvarnom vremenu klasificiraju se segmenti trajektorije prema modovima prijevoza. Implementirana je metoda kNN u kontekstu aktivnog učenja što znači da se parametri klasifikatora tijekom prikupljanja podataka ispravljaju. Početna točnost klasifikatora od 53,5% je nakon procesa aktivnog učenja ispravljena na 75%. U klasifikaciju je uključeno 14 modova prijevoza navedenih u Tablici 2.

#### 4. Detekcija točke promjene moda prijevoza unutar GPS trajektorije

Trajektorija jednog korisnika sastoji se od niza uzastopnih GPS zapisa koji opisuju kretanje tog korisnika. Jedna trajektorija korisnika može uključivati više povezanih modova prijevoza koje korisnik koristi kako bi završio putovanje od polazišta prema odredištu. Takve trajektorije nije moguće klasificirati jer je izlaz klasifikatora uvijek jedan mod prijevoza. Stoga je trajektorije, koje sadrže više modova prijevoza, potrebno prethodno segmentirati kako bi se razdvojili segmenti koji će sadržavati samo jedan mod prijevoza. Većina istraživanja u području segmentacije modova prijevoza zasnovana je upravo na subjektivnim metoda u kojima autori postavljaju iskustvene zaključke na temelju kojih se kreiraju pravila detekcije točke promjene moda prijevoza (engl. *Mode Changing Point*, MCP). Metode su uglavnom zasnovane na neizrastoj logici (Rasmussen et al., 2015; Schuessler and Axhausen, 2009b; Schussler et al., 2011; Tsui and Shalaby, 2006; Biljecki et al., 2013) ili na nizu heurističkih pravila (Bohte and Maat, 2009; Chung and Shalaby, 2005).

Najzastupljenija metoda segmentiranja trajektorije je neizrastita logika. Neizrastita logika zasniva se na iskazima o stanju i nizu pravila zaključivanja zasnovanim na viševrijednosnoj mjeri istinitosti, čiji je cilj, u određivanju moda prijevoza, detektirati neuobičajen zapis tijekom prikupljanja podataka. Ako se jedno od pravila aktivira, točka koja je uzrokovala aktivaciju smatra se mjestom tranzicije s jednog moda prijevoza u drugi. Vrlo često su pravila zasnovana na gubitku signala tijekom kojeg GPS daje neuobičajene zapise, tj. pokazuje značajna odstupanja u bilježenju. Drugo često pravilo detektira višestruke GPS zapise na istoj lokaciji što vrlo često signalizira stajanje. Gubitak signala se događa kod ulaska u autobus, tramvaj ili vlak radi metalne konstrukcije prijevoznog sredstva koja ometa GPS signal (Geurs et al., 2015).

Postupak segmentiranja trajektorije gotovo uvijek prethodi klasifikaciji moda prijevoza. Među prvim istraživanjima koja su promatrala segmentiranje trajektorije su Chung and Shalaby (2005). Autori detektiraju MCP koristeći dva pravila, pri čemu se prvim pravilom detektiraju brze tranzicije iz jednog moda prijevoza u drugi (pješač-automobil), a drugim pravilom se detektiraju sporije tranzicije (pješač-autobus). U prvom pravilu, ako je povećanje brzine iznad 10 km/h i razlika u vremenu između dva zapisa veća od 5 s točka se smatra MCP. Brzina od 10 km/h se smatra maksimalnom brzinom pješačaka. Drugim pravilom se prate mjesta na kojima se pojavljuje gubitak signala, primjerice ulazak u autobus. Ako takvo stanje traje duže od 150 metara detektirana je MCP. Pravila ovog tipa su učestala kod segmentiranja trajektorije, te se pravila uglavnom generiraju iskustveno (Xiao et al., 2015).

Tsui and Shalaby (2006) definiraju točku promjene moda prijevoza (engl. *Mode Transfer Point*, MTP) kao geografsku lokaciju na kojoj korisnici prometne mreže prelaze iz jednog moda prijevoza u drugi. Autori definiraju tri tipa MTP-a: i) početak pješačenja (engl. *Start of Walk*, SoW), točka u kojoj korisnik prelazi iz bilo kojeg moda prijevoza na pješačenje; ii) kraj pješačenja (engl. *End of Walk*, EoW), točka u kojoj korisnik prelazi s pješačenja na bilo koji drugi mod prijevoza te iii) kraj perioda bez zapisa o kretanju korisnika (engl. *End of Gap*, EoG), točka u kojoj prestaje duži period u kojem se nije pojavio niti jedan zapis o kretanju korisnika. Sličnu metodologiju koriste Zheng et al. (2010a) na podacima prikupljenim u sklopu GeoLife projekta u kojima segmentiraju trajektorije koristeći pravilo pješačkog segmenta koji signalizira tranziciju s jednog moda prijevoza na drugi. Pješačke segmente detektiraju pomoću brzine i akceleracije, tj. ako je brzina manja od 2,5 m/s, a akceleracija manja od 1,5 m/s<sup>2</sup>.

Nadalje, Xiao et al. (2015) razvijaju dva pravila pomoću kojih detektiraju krajeve trajektorije. Prvo pravilo se odnosi na kraj putovanja zbog stajanja na nekoj lokaciji duže vrijeme. Lokacije se dijele na uobičajene lokacije (dom, posao) i neuobičajene lokacije (kupovina, ambulanta i slično). Zaustavljanje na lokaciji se najčešće detektira pomoću gubitka signala radi ulaska u zatvoreni prostor. U trenutku u kojem je zaustavljanje detektirano, započinje se mjeriti vrijeme provedeno na detektiranoj lokaciji. Ako je vrijeme na uobičajenoj lokaciji veće od  $t_1$  ili na neuobičajenoj veće od  $t_2$ , točka zaustavljanja se smatra krajem trajektorije. Drugo pravilo se odnosi na kraj trajektorije uzrokovan

ostavljanjem ili skupljanjem nekoga na nekoj lokaciji. Autori objašnjavaju da su ove pojave, u cestovnom prometu, obično praćene U skretanjima nakon kojih se korisnik vraća na početnu lokaciju. Stoga, se prvo detektiraju U skretanja te se nakon detekcije izračunava prijeđena udaljenost prije i poslije skretanja. Ako je prijeđena udaljenost na cestovnim segmentima veća od kritične udaljenosti  $\mu$ , lokacija se smatra krajem trajektorije. U drugom slučaju, kada se za putovanje nakon U skretanja koriste isti cestovni segmenti, izračunava se udaljenost između zabilježenih točaka. Ako je udaljenost manja od zadanog praga, lokacija se označava kao kraj trajektorije. U konačnici autori korigiraju dobivene krajeve trajektorija koje su detektirane pomoću oba pravila, te ih spajaju u jedan kraj.

Yang et al. (2015) za segmentiranje trajektorije koriste metodu valne transformacije maksimuma modula. Ova metoda se često koristi kod procesiranja signala jer može detektirati singularitet signala. Ako se u problemu segmentiranja trajektorije brzina proglasi sekvencom signala, može se primijeniti isti način detektiranja nepravilnosti signala. Te nepravilnosti zapravo predstavljaju prelazak s jednog moda prijevoza na drugi. Prije segmentiranja trajektorije, autori uklanjaju šumove iz trajektorije. U suprotnom bi ovom metodom stajanja u gužvi ili na semaforu također bila detektirana kao tranzicija s jednog moda prijevoza na drugi.

Dabiri et al. (2019) provode segmentaciju trajektorije u dva koraka. U prvom koraku dijele sve trajektorije na segmente istih duljina. Autori navode da je razlog ove podjele klasifikator koji kasnije koriste za detekciju moda prijevoza, a koji kao ulaz mora primiti ulazne podatke iste dužine. Zatim formuliraju problem kao problem diskretne optimizacije. Za problem detekcije diskretne točke promjene predložene su mnoge ciljne funkcije i algoritmi pretraživanja (Oudre and Vayatis, 2018). Autori odabiru model pomaka srednje vrijednosti (engl. *mean shift*) za ciljnu funkciju i algoritam pretraživanja (engl. *Pruned Exact Linear Time*, PELT) koji su predložili Killick et al. (2012).

## 5. Zaključak

U ovom radu prikazan je pregled istraživanja u području klasifikacije modova prijevoza s naglaskom na tri glavna zadatka tijekom detekcije moda prijevoza: i) odabir relevantnih značajki, ii) segmentiranje trajektorije te iii) klasifikacija modova prijevoza. Osnovne statističke značajke nisu dovoljna mjera razlikovanja između modova prijevoza, pa mnoga istraživanja razvijaju nove značajke kojima će bolje opisati ponašanje pojedinih modova prijevoza. Za detekciju modova prijevoza, kao posljedica sve veće upotrebe mobilnih uređaja, najčešće se koriste GNSS podaci u kombinaciji s podacima s akcelerometra. Trajektorije najčešće uključuju korištenje više modova prijevoza, pa je trajektorije prije klasifikacije potrebno segmentirati na segmente koji sadrže samo jedan mod prijevoza. Prvi radovi u području segmentacije trajektorije koriste metode zasnovane na pravilima, koje su i dalje najzastupljenije metode u istraživanjima. Pravila su uglavnom kreirana prema iskustvenim spoznajama. Suprotno tome, za klasifikaciju modova prijevoza primijenjen je velik broj metoda, no najveću točnost postižu metode temeljene na stablu odluke, osobito RF. U području klasifikacije modova prijevoza razlikuju se dva pristupa klasifikacije: klasifikacija nad povijesnim podacima i klasifikacija u stvarnom vremenu. Osnovna metodološka razlika je korištenje vremenskih prozora u klasifikaciji u stvarnom vremenu pri čemu je zaključivanje otežano jer se temelji na puno manjem dijelu trajektorije, pa je veća vjerojatnost zamjene sličnih modova prijevoza. Stoga se često nakon klasifikacije promatraju tranzicije modova prijevoza između vremenskih prozora.

U području segmentacije trajektorije kao i u području klasifikacije modova prijevoza postoji potreba za unaprjeđivanjem metoda kako bi se postupak prikupljanja dnevnika putovanja u potpunosti automatizirao. Do sada primijenjene metode za segmentaciju trajektorije uglavnom su zasnovane na iskustvenim pravilima pri čemu je osnovni nedostatak izostavljanje dijela pravila koja mogu povećati točnosti postupka segmentiranja trajektorije. Tek jedan rad istražuje primjenu optimizacijskih metoda, koje su jedan od načina za automatizaciju postupka segmentiranja trajektorije. Također, u području klasifikacije modova prijevoza javlja se potreba za stvarno vremenskom klasifikacijom. Primjenom aktivnog učenja klasifikatora tijekom prikupljanja podataka postižu se bolji rezultati klasifikacije, a takve metode su vrlo malo zastupljene u literaturi. S obzirom na to da veća točnost klasifikacije znači kvalitetnije podatke, implementacija aktivnog učenja klasifikatora jedno je od budućih smjerova istraživanja u području klasifikacije modova prijevoza.

## Literatura

- Andrienko, G., Andrienko, N., Hurter, C., Rinzivillo, S., Wrobel, S., 2011. From movement tracks through events to places: Extracting and characterizing significant places from mobility data, in: 2011 IEEE Conference on Visual Analytics Science and Technology (VAST), pp. 161–170. doi:10.1109/VAST.2011.6102454.

- Ansarilari, Z., Golroo, A., 2015. Automated Transportation Mode Detection Using Smart Phone Application via Machine Learning: Case Study Mega City of Tehran.
- Aparicio, S., Perez, J., Bernardos, A.M., Casar, J.R., 2008. A fusion method based on bluetooth and wlan technologies for indoor location, in: 2008 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems, pp. 487–491. doi:10.1109/MFI.2008.4648042.
- Ashbrook, D., Starner, T., 2003. Using GPS to learn significant locations and predict movement across multiple users. *Personal and Ubiquitous Computing* 7, 275–286. URL: <https://doi.org/10.1007/s00779-003-0240-0>, doi:10.1007/s00779-003-0240-0.
- Axhausen, K.W., Schönfelder, S., Wolf, J., Oliveira, M., Samaga, U., 2003. 80 weeks of GPS-traces: approaches to enriching the trip information. submitted to the 83rd transportation research board meeting. doi:10.3929/ethz-a-004570614. weitere Autoren: M. Oliveira, U. Samaga.
- Bantis, T., Haworth, J., 2017. Who you are is how you travel: A framework for transportation mode detection using individual and environmental characteristics. *Transportation Research Part C: Emerging Technologies* 80, 286–309. doi:10.1016/j.trc.2017.05.003.
- Bao, L., Intille, S.S., 2004. Activity recognition from user-annotated acceleration data, in: Ferscha, A., Mattern, F. (Eds.), *Pervasive Computing*, Springer Berlin Heidelberg, Berlin, Heidelberg. pp. 1–17.
- Bian, Y., Yang, C., Zhao, J.L., Liang, L., 2018. Good drivers pay less: A study of usage-based vehicle insurance models. *Transportation Research Part A: Policy and Practice* 107, 20–34. URL: <http://www.sciencedirect.com/science/article/pii/S096585641730561X>, doi:<https://doi.org/10.1016/j.tra.2017.10.018>.
- Bieber, G., Voskamp, J., Urban, B., 2009. Activity recognition for everyday life on mobile phones, pp. 289–296. doi:10.1007/978-3-642-02710-9\_32.
- Biljecki, F., Ledoux, H., van Oosterom, P., 2013. Transportation mode-based segmentation and classification of movement trajectories. *International Journal of Geographical Information Science* 27, 385–407. URL: <https://doi.org/10.1080/13658816.2012.692791>, doi:10.1080/13658816.2012.692791.
- Bloch, A., Erdin, R., Meyer, S., Keller, T., Spindler, A., 2015. Battery-efficient transportation mode detection on mobile devices, in: 2015 16th IEEE International Conference on Mobile Data Management, pp. 185–190. doi:10.1109/MDM.2015.16.
- Bohte, W., Maat, K., 2009. Deriving and validating trip purposes and travel modes for multi-day GPS-based travel surveys: A large-scale application in the Netherlands. *Transportation Research Part C-emerging Technologies - TRANSPORT RES C-EMERG TECHNOL* 17, 285–297. doi:10.1016/j.trc.2008.11.004.
- Bolbol, A., Cheng, T., Tsapakis, I., Haworth, J., 2012. Inferring hybrid transportation modes from sparse GPS data using a moving window SVM classification. *Computers, Environment and Urban Systems* 36, 526–537. URL: <http://www.sciencedirect.com/science/article/pii/S0198971512000543>, doi:<https://doi.org/10.1016/j.compenvurbsys.2012.06.001>. special Issue: Advances in Geocomputation.
- Brezmes, T., Gorricho, J., Cotrina, J., 2009. Activity recognition from accelerometer data on a mobile phone, pp. 796–799. doi:10.1007/978-3-642-02481-8\_120.
- Byon, Y.J., Abdulhai, B., Shalaby, A.S., 2007. Impact of sampling rate of GPS-enabled cell phones on mode detection and GIS map matching performance. Technical Report.
- Byon, Y.J., Liang, S., 2014. Real-time transportation mode detection using smartphones and artificial neural networks: Performance comparisons between smartphones and conventional global positioning system sensors. *Journal of Intelligent Transportation Systems* 18, 264–272. URL: <https://doi.org/10.1080/15472450.2013.824762>, doi:10.1080/15472450.2013.824762, arXiv:<https://doi.org/10.1080/15472450.2013.824762>.
- Cao, H., Mamoulis, N., Cheung, D.W., 2005. Mining frequent spatio-temporal sequential patterns, in: Fifth IEEE International Conference on Data Mining (ICDM'05), pp. 8 pp.–. doi:10.1109/ICDM.2005.95.
- Chung, E.H., Shalaby, A., 2005. A trip reconstruction tool for GPS-based personal travel surveys. *Transportation Planning and Technology* 28, 381–401. URL: <https://doi.org/10.1080/03081060500322599>, doi:10.1080/03081060500322599, arXiv:<https://doi.org/10.1080/03081060500322599>.
- Consolvo, S., W. McDonald, D., Toscos, T., Y. Chen, M., Froehlich, J., L. Harrison, B., V. Klasnja, P., LaMarca, A., LeGrand, L., Libby, R., Smith, I., Landay, J., 2008. Activity sensing in the wild: A field trial of ubifit garden, pp. 1797–1806. doi:10.1145/1357054.1357335.
- Cottrill, C.D., Pereira, F.C., Zhao, F., Dias, I.F., Lim, H.B., Ben-Akiva, M.E., Zegras, P.C., 2013. Future mobility survey: Experience in developing a smartphone-based travel survey in Singapore. *Transportation Research Record* 2354, 59–67. URL: <https://doi.org/10.3141/2354-07>, doi:10.3141/2354-07, arXiv:<https://doi.org/10.3141/2354-07>.
- Dabiri, S., Heaslip, K., 2018a. Inferring transportation modes from GPS trajectories using a convolutional neural network. *Transportation Research Part C: Emerging Technologies* 86, 360–371. URL: <http://www.sciencedirect.com/science/article/pii/S0968090X17303509>, doi:<https://doi.org/10.1016/j.trc.2017.11.021>.
- Dabiri, S., Heaslip, K., 2018b. Transport-domain applications of widely used data sources in the smart transportation: A survey.
- Dabiri, S., Lu, C., Heaslip, K., Reddy, C.K., 2019. Semi-supervised deep learning approach for transportation mode identification using GPS trajectory data. *IEEE Transactions on Knowledge and Data Engineering*, 1–1doi:10.1109/TKDE.2019.2896985.
- D'Andrea, E., Marcelloni, F., 2017. Detection of traffic congestion and incidents from GPS trace analysis. *Expert Systems with Applications* 73, 43–56. URL: <http://www.sciencedirect.com/science/article/pii/S0957417416306935>, doi:<https://doi.org/10.1016/j.eswa.2016.12.018>.
- Das, R.D., Ronald, N., Winter, S., 2014. Clustering based transfer detection with fuzzy activity recognition from smart-phone GPS trajectories, in: 17th International IEEE Conference on Intelligent Transportation Systems (ITSC), pp. 3138–3143. doi:10.1109/ITSC.2014.6958195.
- de Fabritiis, C., Ragona, R., Valenti, G., 2008. Traffic estimation and prediction based on real time floating car data, in: 2008 11th International IEEE Conference on Intelligent Transportation Systems, pp. 197–203. doi:10.1109/ITSC.2008.4732534.
- van Dijk, J., 2018. Identifying activity-travel points from GPS-data with multiple moving windows. *Computers Environment and Urban*

- Systems doi:[10.1016/j.compenvurbsys.2018.02.004](https://doi.org/10.1016/j.compenvurbsys.2018.02.004).
- Draijer, G., Kalfs, N., Perdok, J., 2000. Global positioning system as data collection method for travel research. *Transportation Research Record* 1719, 147–153. doi:[10.3141/1719-19](https://doi.org/10.3141/1719-19).
- Eftekhari, H.R., Ghatte, M., 2016. An inference engine for smartphones to preprocess data and detect stationary and transportation modes. *Transportation Research Part C: Emerging Technologies* 69, 313 – 327. URL: <http://www.sciencedirect.com/science/article/pii/S0968090X16300730>, doi:<https://doi.org/10.1016/j.trc.2016.06.005>.
- Endo, Y., Toda, H., Nishida, K., Kawanobe, A., 2016. Deep feature extraction from trajectories for transportation mode estimation, in: Bailey, J., Khan, L., Washio, T., Dobbie, G., Huang, J.Z., Wang, R. (Eds.), *Advances in Knowledge Discovery and Data Mining*, Springer International Publishing, Cham, pp. 54–66.
- Erdelić, T., Vrbanić, S., Rožić, L., 2015. A model of speed profiles for urban road networks using g-means clustering, in: 2015 38th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), pp. 1081–1086. doi:[10.1109/MIPRO.2015.7160436](https://doi.org/10.1109/MIPRO.2015.7160436).
- Etemad, M., Soares Júnior, A., Matwin, S., 2018. Predicting transportation modes of GPS trajectories using feature engineering and noise removal, in: Bagheri, E., Cheung, J.C. (Eds.), *Advances in Artificial Intelligence*, Springer International Publishing, Cham, pp. 259–264.
- Feng, T., Timmermans, H.J., 2013. Transportation mode recognition using GPS and accelerometer data. *Transportation Research Part C: Emerging Technologies* 37, 118 – 130. URL: <http://www.sciencedirect.com/science/article/pii/S0968090X13002039>, doi:<https://doi.org/10.1016/j.trc.2013.09.014>.
- Figo, D., Diniz, P.C., Ferreira, D.R., Cardoso, J.M.P., 2010. Preprocessing techniques for context recognition from accelerometer data. *Personal and Ubiquitous Computing* 14, 645–662. URL: <https://doi.org/10.1007/s00779-010-0293-9>, doi:[10.1007/s00779-010-0293-9](https://doi.org/10.1007/s00779-010-0293-9).
- Geurs, K.T., Thomas, T., Bijlsma, M., Douhou, S., 2015. Automatic Trip and Mode Detection with Move Smarter: First Results from the Dutch Mobile Mobility Panel. *Transportation Research Procedia* 11, 247 – 262. URL: <http://www.sciencedirect.com/science/article/pii/S2352146515003130>, doi:<https://doi.org/10.1016/j.trpro.2015.12.022>.
- transport Survey Methods: Embracing Behavioural and Technological Changes Selected contributions from the 10th International Conference on Transport Survey Methods 16-21 November 2014, Leura, Australia.
- Gjoreski, H., Ciliberto, M., Wang, L., Ordonez Morales, F.J., Mekki, S., Valentin, S., Roggen, D., 2018. The University of Sussex-Huawei Locomotion and Transportation Dataset for Multimodal Analytics With Mobile Devices. *IEEE Access* 6, 42592–42604. doi:[10.1109/ACCESS.2018.2858933](https://doi.org/10.1109/ACCESS.2018.2858933).
- Gong, H., Chen, C., Bialostozky, E., Lawson, C.T., 2012. A GPS/GIS method for travel mode detection in New York City. *Computers, Environment and Urban Systems* 36, 131 – 139. URL: <http://www.sciencedirect.com/science/article/pii/S0198971511000536>, doi:<https://doi.org/10.1016/j.compenvurbsys.2011.05.003>. special Issue: Geoinformatics 2010.
- Gong, L., Kanamori, R., Yamamoto, T., 2018. Data selection in machine learning for identifying trip purposes and travel modes from longitudinal GPS data collection lasting for seasons. *Travel Behaviour and Society* 11, 131 – 140. URL: <http://www.sciencedirect.com/science/article/pii/S2214367X1730039X>, doi:<https://doi.org/10.1016/j.tbs.2017.03.004>.
- Gonzalez, P.A., Weinstein, J.S., Barbeau, S.J., Labrador, M.A., Winters, P.L., Georggi, N.L., Perez, R., 2010. Automating mode detection for travel behaviour analysis by using global positioning systems enabled mobile phones and neural networks. *IET Intelligent Transport Systems* 4, 37–49. doi:[10.1049/iet-its.2009.0029](https://doi.org/10.1049/iet-its.2009.0029).
- Hansen, L., Salamon, P., 1990. Neural network ensembles. *Pattern Analysis and Machine Intelligence*, *IEEE Transactions on* 12, 993 – 1001. doi:[10.1109/34.58871](https://doi.org/10.1109/34.58871).
- Hemminki, S., Nurmi, P., Tarkoma, S., 2013. Accelerometer-based transportation mode detection on smartphones, in: *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems*, ACM, New York, NY, USA, pp. 13:1–13:14. URL: <http://doi.acm.org/10.1145/2517351.2517367>, doi:[10.1145/2517351.2517367](https://doi.org/10.1145/2517351.2517367).
- Iqbal, M.S., Choudhury, C., Wang, P., Gonzalez, M.C., 2014. Development of origin–destination matrices using mobile phone call data. *Transportation Research Part C: Emerging Technologies* 40, 63–74. doi:[10.1016/j.trc.2014.01.002](https://doi.org/10.1016/j.trc.2014.01.002).
- Jahangiri, A., Rakha, H., 2014. Developing a support vector machine (SVM) classifier for transportation mode identification by using mobile phone sensor data.
- Jahangiri, A., Rakha, H., 2015. Applying machine learning techniques to transportation mode recognition using mobile phone sensor data. *IEEE Transactions on Intelligent Transportation Systems* , 1–12doi:[10.1109/TITS.2015.2405759](https://doi.org/10.1109/TITS.2015.2405759).
- Killick, R., Fearnhead, P., Eckley, I., 2012. Optimal detection of changepoints with a linear computational cost. *Journal of the American Statistical Association* 107, 1590–1598. doi:[10.1080/01621459.2012.737745](https://doi.org/10.1080/01621459.2012.737745).
- Lee, J., Han, J., Li, X., Cheng, H., 2011. Mining discriminative patterns for classifying trajectories on road networks. *IEEE Transactions on Knowledge and Data Engineering* 23, 713–726. doi:[10.1109/TKDE.2010.153](https://doi.org/10.1109/TKDE.2010.153).
- Lin, M., Hsu, W.J., 2014. Mining GPS data for mobility patterns: A survey. *Pervasive and Mobile Computing* 12, 1 – 16. URL: <http://www.sciencedirect.com/science/article/pii/S1574119213000825>, doi:<https://doi.org/10.1016/j.pmcj.2013.06.005>.
- Ma, X., Dai, Z., He, Z., Ma, J., Wang, Y., Wang, Y., 2017. Learning traffic as images: A deep convolutional neural network for large-scale transportation network speed prediction. *Sensors* 17, 818. doi:[10.3390/s17040818](https://doi.org/10.3390/s17040818).
- Madhavapeddy, A., Tse, A., 2005. A study of bluetooth propagation using accurate indoor location mapping, in: *Proceedings of the 7th International Conference on Ubiquitous Computing*, Springer-Verlag, Berlin, Heidelberg, pp. 105–122. URL: [http://dx.doi.org/10.1007/11551201\\_7](http://dx.doi.org/10.1007/11551201_7), doi:[10.1007/11551201\\_7](https://doi.org/10.1007/11551201_7).
- Miluzzo, E., Lane, N.D., Fodor, K., Peterson, R., Lu, H., Musolesi, M., Eisenman, S.B., Zheng, X., Campbell, A.T., 2008. Sensing meets mobile social networks: The design, implementation and evaluation of the ceneme application, in: *Proceedings of the 6th ACM Conference on Embedded Network Sensor Systems*, ACM, New York, NY, USA, pp. 337–350. URL: <http://doi.acm.org/10.1145/1460412.1460445>, doi:[10.1145/1460412.1460445](https://doi.org/10.1145/1460412.1460445).
- Montini, L., Prost, S., Schrammel, J., Rieser-Schussler, N., Axhausen, K.W., 2015. Comparison of travel diaries generated from smartphone

- data and dedicated GPS devices. *Transportation Research Procedia* 11, 227 – 241. URL: <http://www.sciencedirect.com/science/article/pii/S2352146515003117>, doi:<https://doi.org/10.1016/j.trpro.2015.12.020>. transport Survey Methods: Embracing Behavioural and Technological Changes Selected contributions from the 10th International Conference on Transport Survey Methods 16-21 November 2014, Leura, Australia.
- Montoya, D., Abiteboul, S., Senellart, P., 2015. Hup-me: inferring and reconciling a timeline of user activity from rich smartphone data, pp. 1–4. doi:[10.1145/2820783.2820852](https://doi.org/10.1145/2820783.2820852).
- Mun, M., Estrin, D., Burke, J., Hansen, M., 2008. Parsimonious mobility classification using GSM and WiFi traces .
- Nick, T., Coersmeier, E., Geldmacher, J., Goetze, J., 2010. Classifying means of transportation using mobile sensor data, in: *The 2010 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–6. doi:[10.1109/IJCNN.2010.5596549](https://doi.org/10.1109/IJCNN.2010.5596549).
- Nitsche, P., Widhalm, P., Breuss, S., Brändle, N., Maurer, P., 2014. Supporting large-scale travel surveys with smartphones – a practical approach. *Transportation Research Part C: Emerging Technologies* 43, 212 – 221. doi:<https://doi.org/10.1016/j.trc.2013.11.005>. special Issue with Selected Papers from Transport Research Arena.
- Nour, A., Hellinga, B., Casello, J., 2016. Classification of automobile and transit trips from smartphone data: Enhancing accuracy using spatial statistics and GIS. *Journal of Transport Geography* 51, 36 – 44. URL: <http://www.sciencedirect.com/science/article/pii/S0966692315002070>, doi:<https://doi.org/10.1016/j.jtrangeo.2015.11.005>.
- Oudre, L., Vayatis, N., 2018. A review of change point detection methods .
- Parkka, J., Ermes, M., Korpipaa, P., Mantyjarvi, J., Peltola, J., Korhonen, I., 2006. Activity classification using realistic data from wearable sensors. *IEEE Transactions on Information Technology in Biomedicine* 10, 119–128. doi:[10.1109/TITB.2005.856863](https://doi.org/10.1109/TITB.2005.856863).
- Pattara-atikom, W., Peachavanish, R., 2007. Estimating road traffic congestion from cell dwell time using neural network, pp. 1 – 6. doi:[10.1109/ITST.2007.4295824](https://doi.org/10.1109/ITST.2007.4295824).
- Pereira, F., Carrion, C., Zhao, F., Cottrill, C.D., Zegras, C., Ben-Akiva, M., 2013. The future mobility survey: Overview and preliminary evaluation, in: *Proceedings of the Eastern Asia Society for Transportation Studies*.
- Prelicean, A.C., Gidofalvi, G., Susilo, Y.O., 2014. Mobility collector. *Journal of Location Based Services* 8, 229–255. URL: <https://doi.org/10.1080/17489725.2014.973917>, doi:[10.1080/17489725.2014.973917](https://doi.org/10.1080/17489725.2014.973917), arXiv:<https://doi.org/10.1080/17489725.2014.973917>.
- Prelicean, A.C., Gidofalvi, G., Susilo, Y.O., 2018. Meili: A travel diary collection, annotation and automation system. *Computers, Environment and Urban Systems* 70, 24 – 34. URL: <http://www.sciencedirect.com/science/article/pii/S0198971517305240>, doi:<https://doi.org/10.1016/j.compenvurbsys.2018.01.011>.
- Rasmussen, T.K., Ingvarsson, J.B., Halldorsdottir, K., Nielsen, O.A., 2015. Improved methods to deduct trip legs and mode from travel surveys using wearable GPS devices: A case study from the Greater Copenhagen area. *Computers, Environment and Urban Systems* 54, 301–313. doi:[10.1016/j.compenvurbsys.2015.04.001](https://doi.org/10.1016/j.compenvurbsys.2015.04.001).
- Reddy, S., Burke, J., Estrin, D., Hansen, M., Srivastava, M., 2008. Determining transportation mode on mobile phones, in: *2008 12th IEEE International Symposium on Wearable Computers*, pp. 25–28. doi:[10.1109/ISWC.2008.4911579](https://doi.org/10.1109/ISWC.2008.4911579).
- Reddy, S., Mun, M., Burke, J., Estrin, D., Hansen, M., Srivastava, M., 2010. Using mobile phones to determine transportation modes. *ACM Trans. Sen. Netw.* 6, 13:1–13:27. URL: <http://doi.acm.org/10.1145/1689239.1689243>, doi:[10.1145/1689239.1689243](https://doi.org/10.1145/1689239.1689243).
- Rožić, L., Fosin, J., Carić, T., 2015. Solving the time dependent vehicle routing problem using real-world speed profiles, in: *Central European Conference on Information and Intelligent Systems (CECIIS)*, pp. 193–200.
- Schapiro, R.E., 1990. The strength of weak learnability. *Machine Learning* 5, 197–227. URL: <https://doi.org/10.1007/BF00116037>, doi:[10.1007/BF00116037](https://doi.org/10.1007/BF00116037).
- Schöenfelder, S., Axhausen, K., Antille, N., Bierlaire, M., 2002. Exploring the potentials of automatically collected GPS data for travel behaviour analysis - a Swedish data source 13.
- Schuessler, N., Axhausen, K., 2009a. Processing raw data from global positioning systems without additional information. *Transportation Research Record* 2105, 28–36. doi:[10.3141/2105-04](https://doi.org/10.3141/2105-04).
- Schuessler, N., Axhausen, K., 2009b. Processing raw data from global positioning systems without additional information. *Transportation Research Record* 2105, 28–36. doi:[10.3141/2105-04](https://doi.org/10.3141/2105-04).
- Schuessler, N., Montini, L., Dobler, C., 2011. Improving post-processing routines for GPS observations using prompted-recall data, *International Steering Committee for Travel Survey Conferences (ISCTSC)*. doi:[10.3929/ethz-a-006689506](https://doi.org/10.3929/ethz-a-006689506). 9th International Conference on Transport Survey Methods; Conference Location: Termas de Puyehue, Chile; Conference Date: November 14-18, 2011.
- Shafique, M., Hato, E., 2015a. Formation of training and testing datasets, for transportation mode identification. *Journal of Traffic and Logistics Engineering* 3, 77–80. doi:[10.12720/jtle.3.1.77-80](https://doi.org/10.12720/jtle.3.1.77-80).
- Shafique, M.A., Hato, E., 2015b. Use of acceleration data for transportation mode prediction. *Transportation* 42, pp. 163–188.
- Shen, L., Stopher, P.R., 2014. Review of GPS travel survey and GPS data-processing methods. *Transport Reviews* 34, 316–334. URL: <https://doi.org/10.1080/01441647.2014.903530>, doi:[10.1080/01441647.2014.903530](https://doi.org/10.1080/01441647.2014.903530), arXiv:<https://doi.org/10.1080/01441647.2014.903530>.
- Siirtola, P., Rönning, J., 2012. Recognizing human activities user-independently on smartphones based on accelerometer data. *International Journal of Interactive Multimedia and Artificial Intelligence* 1, 38–45. URL: [http://www.ijimai.org/journal/sites/default/files/IJIMAI20121\\_5\\_5.pdf](http://www.ijimai.org/journal/sites/default/files/IJIMAI20121_5_5.pdf), doi:[10.9781/ijimai.2012.155](https://doi.org/10.9781/ijimai.2012.155).
- Sohn, T., Varshavsky, A., LaMarca, A., Y. Chen, M., Choudhury, T., Smith, I., Consolvo, S., Hightower, J., Griswold, W., de Lara, E., 2006. Mobility detection using everyday GSM traces, pp. 212–224. doi:[10.1007/11853565\\_13](https://doi.org/10.1007/11853565_13).
- Sonderen, T., 2016. Detection of transportation mode solely using smartphones.
- Steenbruggen, J., Borzacchiello, M.T., Nijkamp, P., Scholten, H., 2011. Mobile phone data from GSM networks for traffic parameter and urban spatial pattern assessment: A review of applications and opportunities. *Geojournal* 78, 1–21. doi:[10.1007/s10708-011-9413-y](https://doi.org/10.1007/s10708-011-9413-y).
- Stenneth, L., Wolfson, O., Yu, P.S., Xu, B., 2011. Transportation mode detection using mobile phones and GIS information, in: *Proceedings*

- of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, ACM, New York, NY, USA. pp. 54–63. URL: <http://doi.acm.org/10.1145/2093973.2093982>, doi:10.1145/2093973.2093982.
- Su, X., Caceres, H., Tong, H., He, Q., 2016. Online travel mode identification using smartphones with battery saving considerations. *IEEE Transactions on Intelligent Transportation Systems* 17, 2921–2934. doi:10.1109/TITS.2016.2530999.
- Sun, J., Zhang, C., Zhang, L., Chen, F., Peng, Z.R., 2014. Urban travel behavior analyses and route prediction based on floating car data. *Transportation Letters* 6, 118–125. doi:10.1179/1942787514Y.0000000017.
- Thomas, T., Puello, L.L.P., Geurs, K., 2018. Intrapersonal mode choice variation: Evidence from a four-week smartphone-based travel survey in the Netherlands. *Journal of Transport Geography* URL: <http://www.sciencedirect.com/science/article/pii/S0966692317301096>, doi:<https://doi.org/10.1016/j.jtrangeo.2018.06.021>.
- Tsui, S., Shalaby, A., 2006. Enhanced system for link and mode identification for personal travel surveys based on global positioning systems. *Transportation Research Record* 1972, 38–45. doi:10.3141/1972-07.
- Wang, B., Gao, L., Juan, Z., 2018. Travel mode detection using GPS data and socioeconomic attributes based on a random forest classifier. *IEEE Transactions on Intelligent Transportation Systems* 19, 1547–1558. doi:10.1109/TITS.2017.2723523.
- Wang, H., LIU, G., DUAN, J., Zhang, L., 2017. Detecting transportation modes using deep neural network. *IEICE Transactions on Information and Systems* E100.D, 1132–1135. doi:10.1587/transinf.2016EDL8252.
- Wang, L., Gjoreski, H., Ciliberto, M., Mekki, S., Valentin, S., Roggen, D., 2019. Enabling Reproducible Research in Sensor-Based Transportation Mode Recognition With the Sussex-Huawei Dataset. *IEEE Access* 7, 10870–10891. doi:10.1109/ACCESS.2019.2890793.
- Wang, S., Chen, C., Ma, J., 2010. Accelerometer based transportation mode recognition on mobile phones, in: 2010 Asia-Pacific Conference on Wearable Computing Systems, pp. 44–46. doi:10.1109/APWCS.2010.18.
- Wolf, J., 2000. Using GPS data loggers to replace travel diaries in the collection of travel data. Ph.D. thesis.
- Xiao, G., Juan, Z., Gao, J., 2015. Inferring Trip Ends from GPS Data Based on Smartphones in Shanghai.
- Xiao, Z., Wang, Y., Fu, K., Wu, F., 2017. Identifying different transportation modes from trajectory data using tree-based ensemble classifiers. *ISPRS International Journal of Geo-Information* 6, 57. doi:10.3390/ijgi6020057.
- Yang, F., Yao, Z., Jin, J., 2015. GPS and acceleration data in multimode trip data recognition based on wavelet transform modulus maximum algorithm. *Transportation Research Record* 2526, 90–98. doi:10.3141/2526-10.
- Yang, S., Wu, Y.J., 2017. Travel mode identification using bluetooth technology. *Journal of Intelligent Transportation Systems* doi:10.1080/15472450.2017.1384698.
- Yu, M.C., Yu, T., Wang, S.C., Lin, C.J., Chang, E.Y., 2014. Big data small footprint: The design of a low-power classifier for detecting transportation modes. *Proc. VLDB Endow.* 7, 1429–1440. URL: <http://dx.doi.org/10.14778/2733004.2733015>, doi:10.14778/2733004.2733015.
- Zhang, L., Dalyot, S., Eggert, D., Sester, M., 2012. Multi-stage approach to travel-mode segmentation and classification of GPS traces. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XXXVIII-4/W25*. doi:10.5194/isprsarchives-XXXVIII-4-W25-87-2011.
- Zheng, Y., Chen, Y., Li, Q., Xie, X., Ma, W.Y., 2010a. Understanding transportation modes based on GPS data for web applications. *ACM Trans. Web* 4, 1:1–1:36. URL: <http://doi.acm.org/10.1145/1658373.1658374>, doi:10.1145/1658373.1658374.
- Zheng, Y., Xie, X., 2008. Learning transportation mode from raw GPS data for geographic application on the web.
- Zheng, Y., Xie, X., Ma, W.Y., 2008. Understanding mobility based on GPS data. URL: <https://www.microsoft.com/en-us/research/publication/understanding-mobility-based-on-gps-data/>.
- Zheng, Y., Xie, X., Ma, W.Y., 2010b. Geolife: A collaborative social networking service among user, location and trajectory. *IEEE Data(base) Engineering Bulletin* URL: <https://www.microsoft.com/en-us/research/publication/geolife-a-collaborative-social-networking-service-among-user-location-and-trajectory/>.