

ANALIZA PODATKOV O PROMETNIH TOKOVIH IN MOBILNOSTI

AJDA PRETNAR, TOMAŽ HOČEVAR IN TOMAŽ CURK

Univerza v Ljubljani, Fakulteta za računalništvo in informatiko, Ljubljana, Slovenija.
E-pošta: ajda.pretnar@fri.uni-lj.si, tomaz.hocevar@fri.uni-lj.si, tomaz.curk@fri.uni-lj.si

Povzetek Podatki o prometu nam lahko pomagajo odgovoriti na več vprašanj o mobilnosti ljudi. Opaženi vzorci razkrivajo, kako narod vozi, kaj počne med tednom in med konci tedna ter kako se navade ljudi spreminjajo čez leto. Tovrstne informacije nam pomagajo razumeti sedanje in prihodnje vedenje turistov ter omogočajo prilagoditev in vpliv na promet. Analizirali smo javno dostopen nabor podatkov števecov cestnega prometa v Sloveniji. Razvili smo računske metode za iskanje zanimivih vzorcev v prometu. Lokacije števecov prometa smo gručili glede na podobnosti v opazovanih prometnih profilih. Takšna avtomatizirana kvantitativna analiza velike količine podatkov je dragoceno orodje za odkrivanje zanimivih lastnosti v prometu. Odprti podatki števecov prometa v realnem času nudijo še globlji vpogled v mobilnost. Na podlagi opaženih korelacij med bližnjimi števci prometa smo ustvarili model cestnega omrežja. Z modelom smo analizirali prometni tok po cestnem omrežju in razvili metodo za štetje prometa v določeno izbrano regijo in iz nje.

Ključne besede:

promet,
števci,
profil,
graf,
tok

1 Uvod

Analiza prometa je tipično umeščena na področje fizike, kjer skušajo raziskovalci razumeti obnašanje voznikov in dinamiko stanja v prometu s teorijami dinamike tekočin in čakalnih vrst (Gerlough and Huber 1976). To vključuje tudi izgradnjo primernih modelov in načinov za simulacijo prometa ob različnih eksperimentalnih pogojih (Treiber and Kesting 2013). Ker je promet tako pomemben vidik vsakdanjega življenja, je preučevanje prometa pravzaprav svoje raziskovalno področje.

V našem delu smo se osredotočili na iskanje prometnih zanimivosti in zakonitosti na osnovi podatkov o prevoznosti točk, kjer so nameščeni števcji prometa. S primerjavo prevoznosti števcjev v različnih obdobjih smo odkrili določene zanimivosti, ki so lahko izhodišče za antropološke in druge raziskave in aplikacije. Zgradili smo tudi model omrežja števcjev, ki omogoča nadaljnje analize prometnih tokov. V modelu so glavne entitete števcji, ki so povezani z usmerjenimi povezavami glede na prehajanje vozil med njimi. Tako lahko opazujemo tok prometa, ki vstopi v državo na izbranem mejnem prehodu in se širi preko drugih števcjev prometa po Sloveniji, ali pa se osredotočimo na neko regijo (pokrajino, občino) in opazujemo, kako se obnaša tok prometa v oz. iz regije.

2 Podatki

Podatki o prevoznosti števcjev prometa so dostopni na portalu Prometno-informacijskega centra za državne ceste (<http://promet.si/>). Vrednosti števcjev prometa se posodabljaajo na približno 5 minut, zato smo s poizvedbami zbirali podatke z enako časovno periodo. Zbiranje podatkov je potekalo od avgusta 2019. V dveh letih se je tako nabralo približno 200.000 meritev prevoznosti vseh števcjev. V vsaki meritvi se nahajajo podatki o približno 1300 števcjih, kar skupaj znaša 260 milijonov meritev. Način pridobivanja in vsebina razpoložljivih podatkov sta natančneje opisana v dokumentaciji portala¹.

¹ https://promet.si/portal/res/doc/B2B_full.pdf

```
"title": "AC-A1, Zadobrova AC : smer LJvzh (p)",  
"id": "1005-12",  
"updated": "2020-12-06T11:45:00Z",  
"summary": "AC-A1, Zadobrova AC : smer LJvzh (p) - Normalen promet  
          (36 vozil/h, povp. hitrost=127km/h, povp. razmik=116.7s,  
          zasedenost=0.2%)",
```

Posamezna meritev števca je sestavljena iz več atributov. Prvi del atributov predstavlja povzetek z opisom števca, identifikacijsko številko, časom in povzetkom meritev.

```
"stevci_lokacija": "1005",  
"stevci_lokacijaOpis": "Zadobrova AC",  
"stevci_cestaOpis": "AC-A1",  
"stevci_odsek": "0047",  
"stevci_stacionaza": "2.421",  
"stevci_smer": "12",  
"stevci_smerOpis": "smer LJvzh",  
"stevci_pasOpis": "(p)",  
"stevci_regija": "ACB LJ",  
"stevci_geoX": "467703",  
"stevci_geoY": "103025",  
"stevci_vmax": "100",
```

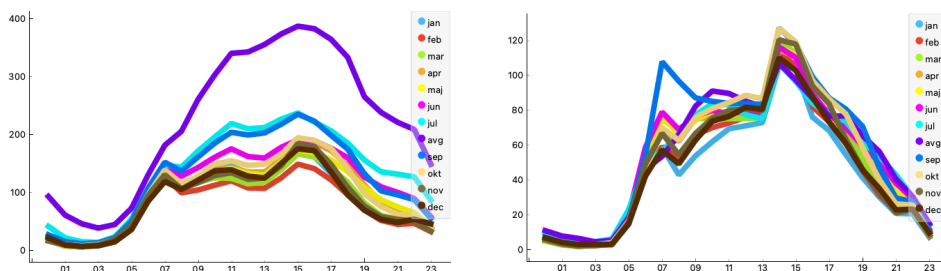
Drugi del atributov opisuje, kje se števec nahaja. Ti podatki so za isti števec v različnih poizvedbah seveda enaki. Vsak števec se nahaja na neki lokaciji, ki pripada cestnemu odseku, kot so določeni s strani Direkcije Republike Slovenije za infrastrukturo . Na vsaki lokaciji se tipično nahajajo števci v obe smeri in na vseh pasovih. Navedena je tudi omejitev hitrosti na lokaciji merjenja. Natančne koordinate števca so podane v koordinatnem sistemu MGI 1901 (EPSG:3912).

```
"stevci_datum": "06/12/2020",  
"stevci_ura": "12:45:00",  
"stevci_stev": "36",  
"stevci_hit": "127",  
"stevci_gap": "116.7",  
"stevci_occ": "2",  
"stevci_stat": "1",  
"stevci_statOpis": "Normalen promet"
```

Zadnji del atributov opisuje konkretne meritve. Navedena sta datum in čas, prevoznost v številu vozil na uro, povprečna hitrost in razmik, zasedenost ceste v odstotkih in ocenjen status prometa.

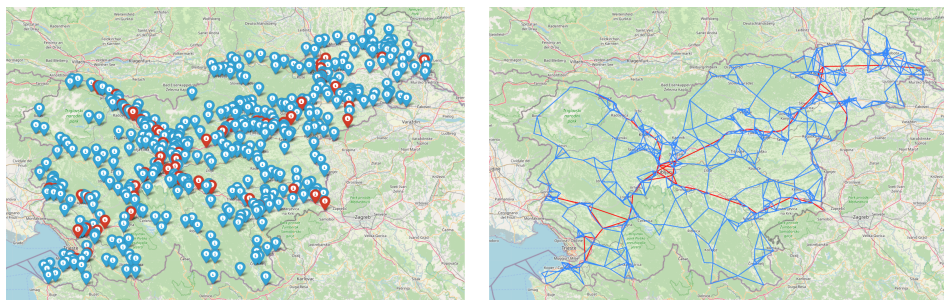
3 Profili prometa

Količino prometa lahko agregiramo in opazujemo v različnih časovnih enotah. Za posamezen števec lahko sestavimo letni profil (po mesecih), tedenski (po dnevih) ali dnevni (po urah). Večina analiz je bila opravljena na dnevni profilih, pri čemer za posamezni števec izračunamo povprečno prevoznost v vsaki uri, povprečje pa lahko računamo v različnih časovnih intervalih, npr. mesecih. S tako vizualizacijo profilov prometa lahko primerjamo, kako se čez dan po urah spreminja prevoznost števnega mesta v različnih mesecih.



Slika 1: Leva slika prikazuje profile števca med Kozino in mejnim preходом Kozina, kjer je viden izrazit porast prometa v avgustu. Desna slika prikazuje števec med Dolenjim Lakošem in Lendavo ter prikazuje izrazit porast v jutranjih urah septembra, ko se v Lendavi odvija Festival Vinarium s trgatvijo.

Da bi se izognili dolgotrajnemu ročnemu primerjanju različnih profilov in iskanju zanimivih oz. informativnih vizualizacij, smo uporabili več različnih metrik za rangiranje vizualizacij (z-score, koeficient variacije, ...) (Holmes, Illowsky, and Dean 2017). Te metrike se razlikujejo glede na lastnost, ki jo želimo zajeti. Lahko nas



Slika 2: Lokacije glavnih števecov in povezave med njimi v modelu omrežja. Z rdečo so označeni števeci na avtocestah in hitrih cestah.

zanimajo vizualizacije, na katerih profil močno odstopa v posamezni uri, ali pa smo bolj osredotočeni na vizualizacije, kjer profil za določen mesec čez cel dan konsistentno odstopa od povprečja ostalih. Slika 1 prikazuje oba primera.

Profile števecov lahko primerjamo s Spearmanovim koeficientom korelacije (Spearman 1904) in jih na podlagi izračunane podobnosti s hierarhičnim gručenjem razvrstimo v skupine (Kaufman and Rousseeuw 2009). S takim postopkom smo izluščili skupine števecov, ki imajo izrazito povečanje prometa samo v jutranjih prometnih konici, samo v popoldanski prometni konici ali pa v obeh.

4 Modeliranje cestnega omrežja

Analiza prometa na posameznih števcih nam ponuja točkoven vpogled v prometno dogajanje. Za analizo prometnih tokov pa potrebujemo povezljivost med merilnimi mesti oz. števcami. To omrežje števecov smo modelirali z grafom, kjer vozlišča predstavljajo števecove, povezave pa neposredne poti med njimi (Slika 2) (Brandes 2005).

Izgradnja omrežja je v celoti avtomatizirana. Postopek najprej identificira glavne števecove, ki enakomerno pokrivajo celotno cestno omrežje. Algoritem nato vzpostavi povezave med bližnjimi števcami, na koncu pa odstrani odvečne povezave zaradi tranzitivnosti.

Povzeti postopek vsebuje precej podrobnosti, ki so ključne za uspešno izvedbo. Določanje povezljivosti med števcami temelji na razdaljah in časih potovanja po cestah med števci. V ta namen smo uporabili storitev Openrouteservice². Pridobili smo tudi podatke o poteku cestnih odsekov, s katerimi smo si pomagali pri izgradnji modela. Pomemben del postopka predstavljajo tudi Pearsonovi korelacijski koeficienti med števci. Če meritve na dveh števcih primerno zamaknemo glede na čas potovanja med njima, lahko primerjamo, kako dobro so meritve obeh števcov korelirane. Če se promet večja in manjša sorazmerno na obeh števcih, lahko sklepamo, da sta števca povezana. Te koeficiente uporabljamo pri določanju povezav in uteži na njih.

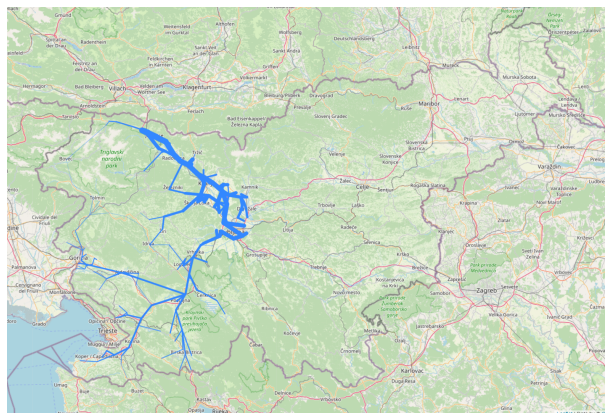
5 Prometni tokovi

Zgrajeno omrežje števcov smo uporabili za modeliranje prometnega toka od mejnih prehodov po Sloveniji. Slika 3 prikazuje oceno, kako se razporedi tok vozil, ki vstopi v Slovenijo na mejnem prehodu Karavanke. Večina vozil potuje proti jugu mimo Ljubljane, nato pa nadaljujejo proti Primorski. Izračun temelji na omrežju z utežmi povezav, ki so bile določene na podlagi meritev števcov v mesecu avgustu.

Tok prometa modeliramo z drevesom najkrajših poti v grafu. Posamezne povezave so utežene glede na korelacijo med števci. Pot čez zaporedje povezav pa ima dolžino, ki je enaka produktu korelacij na povezavah. Iskanje drevesa najkrajših poti je klasičen problem v teoriji grafov (Cormen et al. 2009).

Očitna pomanjkljivost postopka je, da je precej odvisen od dolžine poti. Več povezav prečka pot, manjša bo njena dolžina (produkt korelacij na poti), zato postopek daje prednost kratkim potem. Druga težava izhaja iz rekonstrukcije omrežja, v kateri lahko pride do manjših napak pri povezovanju števcov. To je razvidno na ljubljanski obvoznici, kjer bi pričakovali povezavo po zahodnem delu. Kljub temu so korelacije po ostalih povezavah na obvoznici dovolj visoke, da to ne onemogoči toka proti Primorski.

² <https://openrouteservice.org/>



Slika 3: Tok prometa z mejnega prehoda Karavanke.

Z uporabo modela omrežja števecv lahko ocenjujemo tudi promet v regijo in iz regije. Meje regije lahko definiramo s poljubnim večkotnikom. Povezave, ki sekajo to mejo prispevajo k spremembi toka v regijo. Če se števec A nahaja izven regije, števec B pa znotraj, prispeva povezava $A \rightarrow B$ k povečanju prometa v regijo. Ali se točka nahaja v ali izven večkotnika je klasičen problem računske geometrije (Preparata and Shamos 2012). Ker poznamo količino prometa samo na števcih oz. vozliščih, moramo oceniti promet po povezavi glede na vzpostavljene uteži. Na podlagi množice števecv (s pripadajočimi utežmi), ki prispevajo k povečanju števila vozil v regiji in tistih, ki prispevajo k zmanjšanju, lahko opazujemo spremembe toka prometa v regijo v izbranem časovnem obdobju.

Opomba

Objavljeni znanstveni prispevek na konferenci.

Zaključek

Opravljeni analize profilov prometne obremenitve in izgradnja modela sta le začetna koraka v smeri analize prometa. Razvili smo prototip spletne strani, ki omogoča enostavno analizo zbranih podatkov in uporabo zgrajenega modela za opazovanje prometnih tokov. Ti se bodo lahko uporabljali za ugotavljanje značilnosti, napovedovanje prihodnjih prometnih obremenitev in v zadnjem koraku za oblikovanje ukrepov za načrtovanje infrastrukture ali razbremenitev oz. bolj enakomerno porazdelitev prometa.

Zahvala

Raziskava, predstavljena v tem članku, je bila izvedena v okviru raziskovalno-razvojnega projekta TRL3-6 Turizem 4.0 - Obogatena turistična izkušnja (OP20.03536), ki ga sofinancira Ministrstvo za izobraževanje, znanost in šport Republike Slovenije in Evropski sklad za regionalni razvoj Evropske unije.

Literatura

- Brandes, Ulrik. 2005. *Network Analysis: Methodological Foundations*. Vol. 3418. Springer Science & Business Media.
- Cormen, Thomas H, Charles E Leiserson, Ronald L Rivest, and Clifford Stein. 2009. "Introduction to Algorithms." In. MIT press.
- Gerlough, David L, and Matthew J Huber. 1976. "Traffic Flow Theory."
- Holmes, Alexander, Barbara Illowsky, and Susan Dean. 2017. *Introductory Business Statistics*. Rice University.
- Kaufman, Leonard, and Peter J Rousseeuw. 2009. *Finding Groups in Data: An Introduction to Cluster Analysis*. Vol. 344. John Wiley & Sons.
- Preparata, Franco P, and Michael I Shamos. 2012. "Computational Geometry: An Introduction." In. Springer Science & Business Media.
- Spearman, Charles. 1904. "The Proof and Measurement of Association Between Two Things." *American Journal of Psychology* 15 (1): 72–101. <https://doi.org/10.2307/1412159>.
- Treiber, Martin, and Arne Kesting. 2013. "Traffic Flow Dynamics." *Traffic Flow Dynamics: Data, Models and Simulation*, Springer-Verlag Berlin Heidelberg.