

Utasszámlálás a városi közösségi közlekedésben: mire lehet alkalmas több adat és a „free WiFi”?

Megyeri István - Dr Csirik János - Dr Majó-Petri Zoltán

Szegedi Tudományegyetem Informatikai Intézet, Üzleti Tudományok Intézet
telefon: 62/546 396, fax: 62/ 546 397

e-mail: imegyeri@inf.u-szeged.hu, csirik@inf.u-szeged.hu, majoz@eco.u-szeged.hu

Kivonat: A városi közösségi közlekedés szakembereinek körében évtizedek óta téma a járművek kihasználtságának mérése. A manuális, azaz a megfigyeléses forgalomszámlálás mellett egyre több közlekedési társaság alkalmaz automatikus módszereket azért, hogy folyamatosan adatokat kaphasson az utasszámról. A technikai fejlődésének köszönhetően a fel és leszállások számlálására és a helymeghatározáshoz egyre megbízhatóbb automatikus adatgyűjtő rendszerek állnak rendelkezésünkre. Ez megnyitotta az utasszámlálás piacát az adatfeldolgozással foglalkozó szervezetek és szakemberek előtt, és kitágította a vizsgálható kérdések halmazát: meg tudjuk-e mérni az utazások célját és hosszát? Tudjuk-e követni az átszállási mintázatokat? Cikkünkben bemutatjuk az utasszámlálásra rendelkezésre álló technológiákat kezdve a manuális utasszámlálástól az infravörös érzékelőn keresztül az utasmédián át egészen a wifi rendszerekig. Ezek közvetlenül vagy közvetve mind alkalmasak úgynevezett keresztmetszeti mérésekre, a járművek terheltségének kielégítő pontosságú becslésére, de álláspontunk szerint jelenleg csak az utasmédia és a WiFi lehet alkalmas célforgalmi mátrixok készítésére. Hipotézisünk szerint a járművekre szerelt „free WiFi” routereken futó úgynevezett „probe request” küldő és „probe response” címzettként szereplő egyedi MAC címek gyűjtésével és leszűrésével következtethetünk a járművön tartózkodó utasok számára, a fel és leszállások helyére és időpontjára, amennyiben azt összekapcsoljuk a jármű által gyűjtött ATM adatokkal. Kísérletünkben a WiFi mérés jóságának tesztelésére egy ugyanabban az időben végzett manuális mérés adatait használtuk kontrollnak. A kísérletünk eredményei azt mutatják, hogy magas korreláció van a gyűjtött adatok és a manuális forgalomszámlálás között. A WiFi adatok alapján kapott becslés általában alacsonyabb értéket ad az utasok valós számához képest, aminek alapvető magyarázó változója, hogy nem minden utas rendelkezik, vagy kapcsolja be a WiFi képes eszközt. A WiFi alapú mérési módszernek további fejlesztési és kutatási kérdése a zajok kiszűrése, és a személyes adatok védelmének biztosítása.

Kulcsszavak: *automatikus utasszámlálás, célforgalmi mátrix,*

Bevezetés

A közösségi közlekedés szakembereinek körében évtizedek óta téma a vonalak, járművek kihasználtságának mérése. A manuális, azaz a megfigyeléses forgalomszámlálás mellett/helyett egyre több közlekedési társaság alkalmaz automatikus módszereket azért, hogy folyamatosan adatokat kaphasson az utasszámról.

Az utasszám egyre pontosabb ismeretében az „okos városok” számára kulcsfontosságú mutatók képezhetők a költségek vonalakra, kilométerekre történő vetítésében, de a közlekedési eszköz használatával megspórolt szennyező anyag kibocsátás kapcsán is.

A technikai fejlődésének köszönhetően a fel és leszállások számlálására és a helymeghatározáshoz egyre megbízhatóbb automatikus adatgyűjtő rendszerek állnak rendelkezésünkre. Ez megnyitotta az utasszámlálás piacát az adatfeldolgozással foglalkozó szervezetek és szakemberek előtt, és kitágította a vizsgálható kérdések halmazát: tudjuk-e mérni az utazások célját és hosszát?

Az első nagy adatmennyiséggel dolgozó utasforgalmi mérések elsősorban keresztmetszeti számlálásokra épültek, melynek során egy adott keresztmetszetben (megállóban) számláljuk az áthaladó forgalom nagyságát (fel és leszállások számát), esetleg rögzítjük az áthaladó utasok típusát (kor, nem, jegy, bérlet, stb.). Ez pontos képet ad az útszakaszon áthaladó forgalom nagyságának, időbeni ingadozásának és összetételének, irány szerinti megoszlásának meghatározása, de nem mond semmit az utazások hosszára, céljára.

Emiatt egyre nagyobb igény mutatkozott úgynevezett célforgalmi forgalomfelvételekre ahol azt szeretnénk megállapítani, honnan hová tartanak az utasok. Itt a csomópontok esetében arra vagyunk

kíváncsiak, hogy az egyes ágak (vonalak) között milyen a forgalomáramlás. A célforgalmi vizsgálat eredménye a célforgalmi mátrix.

A körzetek/megállók alkotják a célforgalmi mátrix sorait és oszlopait, vagyis adott területre vonatkozó célforgalmi mátrix annyi sorból és oszlopból áll, ahány körzetre osztottuk a területet.

A célforgalmi forgalomfelvétel esetén érdemes lehet különbséget tenni az utazások (unlinked trips) és összekapcsolt utazások (linked trips) között. Utazások száma alatt értjük azon utasok számát akik tömegközlekedési eszközt használnak. Az utasok számlálásra kerülnek minden alkalommal amikor felszállnak egy járműre. Nem számít hány járművet használtak a kiindulási ponttól a célig. Például egy utas aki felszáll egy járatra, majd a leszállást követően azonnal átszáll egy másikra két utat tett meg.

Az utazások száma az összes felszállásra utal, míg az összekapcsolt utazások száma az utasok számának meghatározására használt és a teljes utak számát méri a kiindulási ponttól a célig (átszállásokat is beleértve).

Az utasszámlálás célja még tovább finomítható: nem csak a forgalom mérése lehet a fókuszban, hanem a megállóban várakozó utasok számának becslése is [1]. Ez alapján az utasok várakozási ideje is meghatározható, ami segíthet a város közössége szempontjából optimális menetrend kialakításában.

1. A mérési módszerek fejlődése: a manuálistól az automata adatgyűjtésig

A manuális utasszámlálást – habár még a mai napig használatos módszer kérdezőbiztosokkal, diákmunkásokkal az elmúlt évtizedben az automatikus utas számlálás (automatic passenger counting - APC) váltotta fel, mely minden esetben valamilyen szenzor/érzékelő mért adatai alapján történik. Ezek pontosságok a ténylegesen jelen levő utasok számára vetítve 5 és 10% közötti hibát tartalmazhatnak.

1.1 Infravörös érzékelők (IR)

A piacon az egyik leggyakrabban előforduló megoldás az infravörös (Infra red - IR) érzékelők. Két fő kategóriába sorolhatók:

- aktív: egy adó és egy vevő egységből áll, a közöttük levő nyalábot megszakítva aktiválódik
- passzív: legyezőszerű nyalábokat bocsát ki, mellyel a testhőmérsékletet érzékelik.

Az első esetben, több vevő és adó egységet szerelnek fel egymással párhuzamosan, általában kettőt az ajtóban úgy, hogy a fel és leszálláskor az utasok megszakítsák a nyalábot. Ez a gyenge pontja is ennek a megoldásnak, mivel több szenzort is fel kell felszerelni ajtónként, ahhoz hogy az irányokat is detektálni lehessen. A passzív érzékelők eredetileg csak észlelésre lettek tervezve, azonban adaptálhatók automatikus számlálásra is. Ezen megoldás esetén szintén több érzékelőre van szükség az irányok detektálására. Nagy érzékelési terület esetén a detektálás pontossága csökken. A legpontosabb műszaki megoldások mindkét szenzort egyszerre használják.

Ugyan bármilyen járműre felszerelhető, nincs mozgó alkatrész, könnyen cserélhető, de ajánlott az érzékelők napi tisztítása, és mivel az utasok számára látható, a vandalizmus vagy a szenzor korrumpálása sem kizárt. Hátránya még, hogy érzékeny lehet az elektrosztatikus sugárzásra. Általánosságban elmondható, hogy megközelítőleg 90% pontosságú adatsorokat szolgáltat, és általában a városban közlekedő járművek 10-30%-a van szenzorokkal felszerelve (a járművek rotálva mennek a különböző útvonalakon).

1.2 Taposó lépcső (Treadle mat)

A taposó lépcső a jármű ajtajában általában az első lépcsőre van szerelve. Rögzíti az utasokat, amint rálépnek a lépcsőre. Bizonyos küszöbértéket meghaladó terhelés esetén számlálásra kerül. Az IR érzékelőkhöz hasonlóan, több piaci megoldás is létezik.

A megoldás gyengeségei közé tartozik, hogy utólagos beszerelése körülményes. Megfelelő beszerelés esetén pontosabb eredményt ad, mint az IR szenzorok (akár 95%), és nem szükséges napi karbantartás, valamint nem látható az utasok számára. Általánosságban a beszerelése összetettebb feladat, cseréje nehézkes. A gyors, tömeges több soros utasáramlásra érzékenyek.

1.3 Jármű diagnosztikai adatok (WIM rendszer, OBU)

A tömegközlekedési eszközök fekete dobozai különböző adatokat gyűjtenek a járművekről, melyből meghatározható a jármű össztömege, a tengelyek terhelése [2]. A módszer lényege, hogy először végzünk üres járatos mérést (adatgyűjtés a járművel egy megadott vonal mentén), majd megadott terheléssel egy referencia mérést. Ezen méréseket felhasználva ellenőrizhetjük a jármű össztömeg változását az üres járatos méréshez képest. Az éles mérések során az utasok össztömegét megkaphatjuk a jármű teljes tömegéből a fentiek ismeretében. Ezen adatot felhasználva, becsülhető a járművön levő utasok száma.

Az utasok össztömegbecslése történhet a fogyasztás, jármű össztömeg, egyéb jármű diagnosztikai adatsorból vagy ezek kombinációjából.

A módszer előnye, hogy semmilyen „ajtószenzor” beszerelését nem igényli, a már meglévő adatokból dolgozik. A használt érzékelők és műszerek jóságát pedig a rendszeres műszaki ellenőrzések garantálják. A megoldás gyengeségei közé tartozik, hogy csak nagy utasszám esetén ad megbízható eredményt. További probléma, hogy útvonalanként és járműveként több kalibráció szükséges a becsléshez szükséges paraméterek meghatározására. Ehhez hasonlóan utasszám becslésre a „weigh in motion” rendszereket is használnak, ahol úttestbe vagy a sínekbe építenek be szenzorokat, melyen a jármű áthaladtakor rögzítésre kerül a tengelyterhelés és a jármű tömege.

1.4 Képfelvétel (járműkamerák)

Képfelvétel alapján történő APC-t két főbb csoportra oszthatjuk a használt kamera típusa szerint: az egyik amit az ajtó fölé helyeznek fel, és felülnézeti képet ad a ki és belépő utasokról, a másik az egykamerás, ami a forgalom szempontjából vizsgált területre irányul.

Az első megközelítésre mind akadémiai [3] [4] és ipari megoldás is létezik. A detektálás során felülnézeti képen az utasok alakját azonosítják szétválasztva a csomagoktól vagy állatoktól; a mozgás irány is meghatározásra kerül, mely alapján azonosítható a le és felszállás. Ezen eszközök általában rendelkeznek integrált LED-el, mellyel üzemelhetnek akár rossz fényviszonyok között is. Egy küszöbérték megadásával akár az utasok magassága alapján különbséget tehet felnőttek és gyerekek között is.

A második megközelítés nem kifejezetten forgalomszámlálásra használatos, de léteznek akadémia megoldások [5] [6]. A detektáláshoz szükséges egy nagy mintás adatbázis gyűjtése, mely tartalmaz különböző helyzetekben képet az utasokról és ezen képekhez annotáció (emberek bejelölése). Ezen adatok alapján fejleszthető egy gépi tanuló modell vagy képfeldolgozási módszer, amely képes, az utasok számát megbecsülni. Az ilyen rendszerek képesek a taposó lépcsőkhöz hasonló pontosságra.

1.5 Intelligens utasmédia (smartcard, e-ticketing)

Amennyiben rendelkezésre áll intelligens utasmédia, és teljesítmény alapú árpolitika, kialakíthatók jegy és/vagy bérlet alapú forgalom számlálási rendszerek is. Ennek során az elektronikus jegykezelőkben rögzített adatok alapján határozzák meg az állomásokon fel és leszálló utasok számát, úgynevezett check in - check out rendszerekben.

Ezzel is vannak mérési problémák, pl.: ha valamilyen oknál fogva a felszállási/leszállási pont ismeretlen. Néhány rendszerben (pl.: London), az utasoknak érvényesíteni kell a fel és leszálláskor egyaránt az utazási kártyájukat. Az ilyen rendszerek leggyakrabban ún. smartcard-ot használnak, amely hasonló elven működik, mint a bankkártya paypass funkciója: érvényesítéskor egy megadott érzékelőhöz kell tartani a kártyát. A felszállás során a smart card-ok egy egyéni azonosítót szolgáltatnak, amellyel azonosítható és követhető az utas. Általában ezen információ elégséges a célforgalmi mátrix kialakításához.

A rendszer hátránya, hogy nem minden esetben oldható meg, hogy a városi közlekedésben minden utas külön-külön érvényesítve szálljon fel a járműre: ez a megoldás az utasok kooperációját igényli, mivel a kártyákat egy leolvasó antenna 5-10 cm-es közelébe kell tartani. Ahhoz, hogy az utasok a leszálláskor is érvényesítsék a kártyájukat, csak teljesítményt/kilométert/megállót is figyelembe vevő rugalmas jegyrendszer esetében lehet elvárni.

A smart card előnyös tulajdonsága, hogy rendszer szinten azonosításra kerül minden utas, így közvetlenül használható célforgalmi mátrix becsléséhez.

1.6 Wifi adatok

Manapság a legtöbb embernél van legalább egy wifi képes okos eszköz (telefon, tablet, óra, laptop). Ezen eszközök végig kísérik bennünket a hétköznapi rutinok során, a közösségi közlekedés használatakor is. A hálózati adatforgalom csökkentése érdekében ezek az eszközök alapértelmezetten úgy vannak konfigurálva, hogy rendszeresen pásztázzák a környezetüket lehetséges wifi hozzáférési pontokat keresve. A pásztázás során az eszközök ún. „probe request”-et küldenek, mely tartalmaz egy egyedi azonosítót (MAC adress: media access control address) az eszközre nézve. Ezen kéréseket rögzítve élhetünk forgalombecsléssel mind a megállók, és a járművek tekintetében.

A nemzetközi gyakorlatban már évek óta próbálkoznak ilyen jellegű adatgyűjtéssel a megállóban és a járművön is, azonban kevés az adat, nem tudunk egzakt korrelációt felállítani [7].

Léteznek nagyvárosi próbák is: Londonban 2016-ban végeztek pilot mérést egy hónapon keresztül. Itt metró vonalak mentén az összes megállóba telepítettek wifi hozzáférési pontot, melyek gyűjtve az utasok eszközeiből érkező kéréseket, meghatározható a kiindulási és a célpont közötti pontos útvonal is. A pilot mérés során gyűjtött adatokból kiderül, hogy két megálló között több fajta útvonalat is használnak az utasok, és az is kiderült, hogy nem a legrövidebb a leggyakoribb. Következő lépésként említik a végleges forma kidolgozását.

A módszer egy lehetséges gyenge pontja, hogy az eszközöket mac cím alapján azonosítja, ami ellen léteznek védekezési módszerek. Mindezek ellenére általános elmondható, hogy a wifi adatok alapján a fel és leszállási helyek mellett a konkrét útvonal is meghatározható, és a kérések gyakorisága alapján akár a felhasználói viselkedés is elemezhető [8].

1.7 Mivel lehet több egy wifi alapú mérés?

A bemutatott különböző műszaki megoldások ugyan egyaránt monitorozzák az utasszámot, de az adatok felhasználhatóságát tekintve nagyon eltérő kimeneteket kapunk, ha a városi közösségi közlekedésben érdekes dimenziók mentén értékeljük a lehetséges megoldásokat, és nem csak a fel és leszállások számára vagyunk kíváncsiak. Az alábbi táblázat tartalmazza a lehetséges kimeneteket.

1. táblázat: Utasszám becslésre alkalmazható műszaki megoldások adattartalma

értékelési szempontok	infra érzékelők	taposó lépcsők	jármű diagnosztika	kamera	utasmédia	WiFi
fel és leszállások száma	közvetlen adatsorok	közvetlen adatsorok	származtatott adatok	közvetlen adatsorok	közvetlen adatsorok	származtatott adatok
utasok száma a járművön	származtatott adatok	származtatott adatok	származtatott adatok	közvetlen adatsorok	származtatott adatok	közvetlen adatsorok
célforgalmi mátrix (unlinked trips)	nincs	nincs	nincs	származtatott adatok	származtatott adatok	származtatott adatok
célforgalmi mátrix (linked trips)	nincs	nincs	nincs	nincs	származtatott adatok	származtatott adatok
az utazás hossza	nincs	nincs	nincs	nincs	származtatott adatok	közvetlen adatsorok
megállóban várakozók száma	nincs	nincs	nincs	nincs	nincs	közvetlen adatsorok

A táblázat jól szemlélteti, hogy a wifi alapú utasszám mérések minden esetben legalább közvetlen adattal szolgálhatnak olyan dimenziókban, amikor nem csak a jármű telítettségére, hanem az utazások mintázatára is kíváncsiak vagyunk: legyen ez az utazások hossza, vagy a leggyakoribb útvonalak feltérképezése, és annak időbeli változása és lefolyása.

2. Automatikus utasszámlálási modell és mérési adatok Szegeden

A legtöbb utazási társaság biztonsági okokból és a járműflotta menedzselése céljából minden járművön ATM adatrögzítőt vagy másnéven tachográfot üzemeltet. A tachográfok számos szenzor adatot rögzítenek változó gyakorisággal: a rögzített adatok között szerepelnek gps koordináták, időbélyegek, ajtónyitási és tengelyterhelési adatok is.

Az okoseszközök vezeték nélküli kommunikáció révén lehetőséget nyújtanak, hogy bárhol és bármikor internethozzáférést létesítsünk. A kommunikáció két módon jöhet létre mobilinternet (GSM) vagy 802.11 (WiFi) protokollon keresztül. Minden 802.11 szabványt támogató eszközhöz tartozik egy MAC-cím, ami egy hexadecimális sorozat, amellyel még a gyártás során látják el. A MAC cím egy globálisan egyedi azonosító, amely alapvető része a WiFi kommunikációnak. Minden az eszköznek küldött vagy az eszköz által kibocsátott üzenet tartalmazza ezt az azonosítót.

A „probe request” egy különleges Wifi csomag, amelyet az eszközök a környezetükben levő hozzáférési pontok felderítésére használnak. A kérésre küldött válasz „probe response” segítségével kapnak információkat a környezetükben levő hálózatok képességeiről és beállításairól. Mivel a „probe” kéréseket az eszközök rendszeres időközönként küldik, így lehetőséget nyújtanak az utasok WiFi képes eszközeinek dedikálására.

Az automatikus utasszámlálási módszerek jóságát jellemezhetjük, ha különböző forrásokból érkező forgalmi adatokkal összevetjük és a hasonlóságot számszerűsítjük pl.: négyzetes hibát vagy korrelációt mérve. Ilyen esetben referencia adat lehet, közlekedési társaságok által meghatározott időnként végzett kézi forgalomszámlálás. A kézi számlálás során általában a járművön minden ajtónál ül egy adatrögzítő aki minden megállóban feljegyzi a fel és leszállók számát.

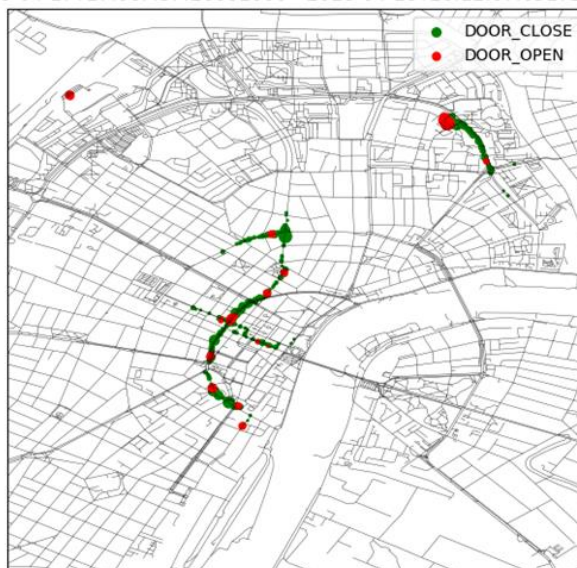
2.1 A felhasznált adatforrások

A nyers ATM adatok megközelítőleg havonta kerülnek mentésre PCM formátumban. A PCM formátum az FDS programmal alakítható át CSV formátumra, amelyen a későbbi elemzéseket végeztük.

A Wifi csomagok rögzítésére egy TP-Link TL-WR941ND router-t használtunk monitor mode-ban, amely OpenWrt és tcpdump segítségével rögzítette a csomagokat időbélyeggel ellátva.

Az ábrán látható egy adott mac címhez tartozó WiFi jelsorozat térképen megjelenítve. A gps koordinátákat az idő alapján összekapcsolt ATM adatok szolgáltattak, a használt időintervallumot önkényesen a szemléltetés érdekében választottuk.

2018-04-17T17:06:45.426881000 - 2018-04-18T20:12:07.651724000



1. ábra: WiFi adatokra építő egyedi utazás megjelenítése ajtónyitásokkal, térképen

Az SZKT által évente végzett utasszámlálás során a járatokon minden ajtónál számlálók jegyzik fel az aktuális megállóban fel és leszálló utasok számát. 2018-ban az utasszámlálás két napon történt: április 14-ikén (szombat) és április 18-ikán (szerda).

Április 18-ikán az egyik trolibuszon történt a WiFi adatgyűjtés, a becslések során kontroll adatként használtuk a feljegyzett utasszámokat. A későbbiben bemutatott hibaérték a manuális feljegyzés és az automatizált módszerek által kapott eredmény közötti hibát jelenti.

A mérési keretrendszer komponenseit időbélyeg alapján kapcsoltuk össze. Az automatikus utasszámlálási módszerek mindegyike a szenzorok által rögzített időbélyegeknél köszönhetően a pontos időpontra nézve tud becslést adni a terheltséggel vagy fel-leszállók számát illetően. A manuális forgalomszámlálási adatok nem tartalmaznak primer időbélyeget, csupán a végállomási indulást. Az összehasonlíthatóság végett ezért szükség van a szenzor adatokból meghatározott ajtó nyitások/zárások és a forgalomszámlálási adatokban szereplő megállók összerendelésére.

Ismerve a járművek által érintett megállók gps koordinátáit és az ajtónyitáshoz tartozó gps koordinátát a koordináták távolsága alapján összerendelhető az ajtó nyitás a megálló nevével. Majd a megálló név alapján összekapcsolhatók a forgalomszámlálási adatok a jármű megállóhoz érkezésének és elhagyásának pontos idejével.

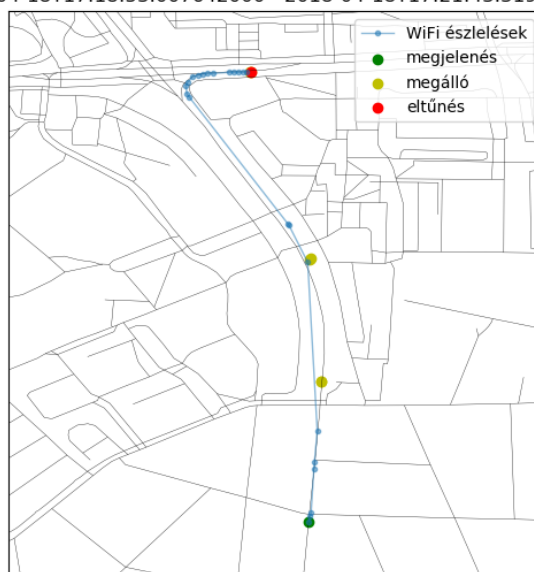
Amennyiben egy megállóban egy sofőr többször is ajtót nyitott, úgy az első ajtónyitás és az utolsó ajtózáras ideje került rögzítésre. Ha a megállóban nem történt ajtónyitás, akkor a forgalomszámlálási adatokban nem került pontos idő hozzárendelésre.

2.2 Becslés a WiFi adatok alapján

A WiFi adatok feldolgozása során az első lépés az észlelt eszközök MAC címeinek meghatározása volt. Második lépésként meg kell határozni minden egyes MAC címhez az általa kibocsátott vagy számára küldött csomagokat. Azonban egy MAC-címhez tartozó eszköz a rögzítés időszakában akár több különböző időpillanatban is jelen lehetett a router környezetében. Emiatt a MAC címekhez tartozó jelsorozatok további darabokra kellett szétválasztani. Az észlelt darabok azonban nem csak az utasokhoz tartozhatnak, hanem bármely a router hatósugarán belül eső okoseszközből érkezhettek (mint gyalogos, megállóban várakozó, autós, más tömegközlekedési eszközön levő utasok, stb.). Így ezen darabok közül kiszűrve az utazásokhoz tartozó jeleket kaphatjuk meg a releváns WiFi adatokat. Ezen WiFi csomagokhoz meghatározva a megjelenéséhez és az eltűnéséhez időben legközelebbi ajtónyitást becslés adható az utas fel- és leszállási helyére.

A hatósugaron belül eső eszközök MAC címét a „probe request” küldő és „probe response” címzettként szereplő egyedi MAC címek leszűrésével határoztuk meg. Majd minden azonosítóhoz hozzárendelésre kerültek azon WiFi csomagok, amelyekben akár küldőként vagy címzettként szerepelt.

2018-04-18T17:18:53.007642000 - 2018-04-18T17:21:43.319078000



2. ábra: Példa utazáshoz tartozó WiFi jelsorozat térképen megjelenítve

Az így kapott jelsorozatokhoz leíró statisztikákat nyertünk ki, mely segítségével meghatározhatók azon eszközök, amelyek több különböző időpontban is jelen voltak a router környezetében. Ilyen leíró jellemző pl.: a MAC címhez tartozó észlelés teljes időtartama vagy a jelsorozat ideje alatt a jármű által megtett km-

ek száma. A megtett km-ek meghatározásához, a WiFi adatokat idő alapján összekapcsoltuk az ATM által rögzített adatokkal.

Azon MAC címek, amelyek jellemzőik alapján több mint 30 percet voltak jelen, vagy legalább 8 km-t tett meg a jármű, szétválasztásra kerültek minden egyes 30 percet meghaladó kimaradás után. Az így kapott WiFi észlelésekhez ismét meghatározásra kerültek a leíró jellemzők, azonban bővült a lista újabb jellemzőkkel mint:

- különböző gps koordináták száma (n_unique_gps)
- jelsorozatokhoz tartozó gps koordináták távolságának (a távolságot a Harvesine formula segítségével számítottuk) összege (gps_dist_diff_sum)
- az észlelések száma (length)
- a MAC címhez tartozó local bit értéke(local_bit)
- a gps-ből számolt és a jármű által megtett távolság hányadosa (dist_ratio)
- az adott időszakaszban történt ajtónyitások száma (nr_intersect)
- a küldött probe request-ek száma (n_probe)

Az ajtónyitások számának meghatározásához, minden egyes jelsorozathoz hozzárendelésre került az első és utolsó észleléshez legközelebbi ajtó nyitás és zárás ideje (start_door_open_time, start_door_close_time, end_door_open_time, end_door_close_time) a korábban leírt ATM adatokból kinyert információk alapján.

A bővített jellemzők listájával egyszerű szabályok alkalmazásával kiszűrhetjük a nem utazásokhoz tartozó WiFi jeleket.

2.3 Mérési eredmények

A kidolgozott módszerek minőségének mérése céljából összehasonlítottuk a becslést az értékeket a manuális forgalomszámlálási adatokkal, az összehasonlításhoz két metrikát használtunk; az átlagos négyzetes hiba négyzetgyökét (root mean squared error, RMSE) és a Pearson korrelációt a becslést és a forgalomszámlálási adatok között. A két metrika alapvetően különböző szempontokból vizsgálja a módszerek jószágát: míg a korreláció azt jelzi, hogy mennyire sikerült a trendet lekövetnie a becslésnek, addig a RMSE arról ad információt, hogy mekkora az átlagos eltérés a két görbe között.

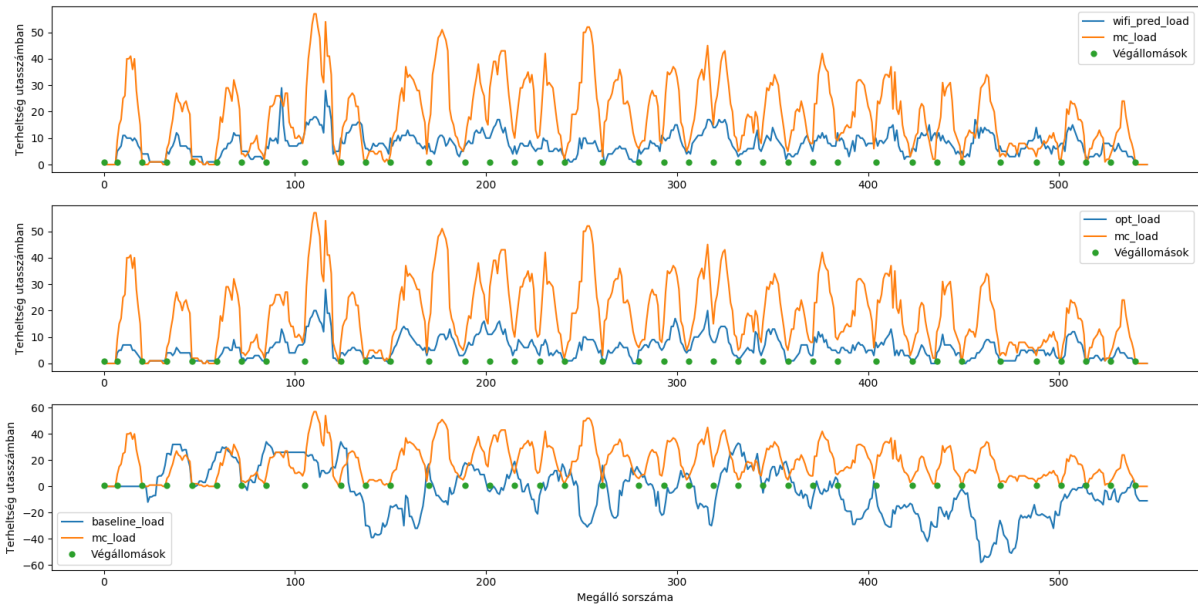
Az elért eredményeink az alábbi táblázatokban láthatóak. A Baseline egy referencia megoldás, mely minden megállóra az utoljára látott fel és leszállók számát jósolja (a megálló első előfordulásánál nulla). A várakozásunk az, hogy a különböző adatforrásokból kinyert becslések teljesítménye felülmúlja a Baseline eredményét. A WiFi threshold a „Becslés WiFi adatok alapján” alfejezetben bemutatott megoldás értékei, amik elemszámukból adódóan felső becslésként szolgálnak.

A WiFi opt egy random search algoritmus melynek célfüggvénye a becslést és a forgalomszámláláson rögzített fel és leszállók számának négyzetes hibája. Kezdetben minden wifi jelsorozatot utasnak számít, majd a jelsorozatot véletlen átbillentve utas/nem utas állapotba törekszik a négyzetes hibát csökkenteni. Ha egy jelsorozat ugyanakkora hibát eredményez utas és nem utasként is akkor, a nem utas állapotot preferálja (hiányos wifi jelsorozat esetén rosszul hozzárendelt megállókat eldobja). A WiFi opt a WiFi adatok alapján elérhető hiba alsó korlátot hivatott demonstrálni.

2. táblázat: Különböző módszerekkel elért eredmények összehasonlítása

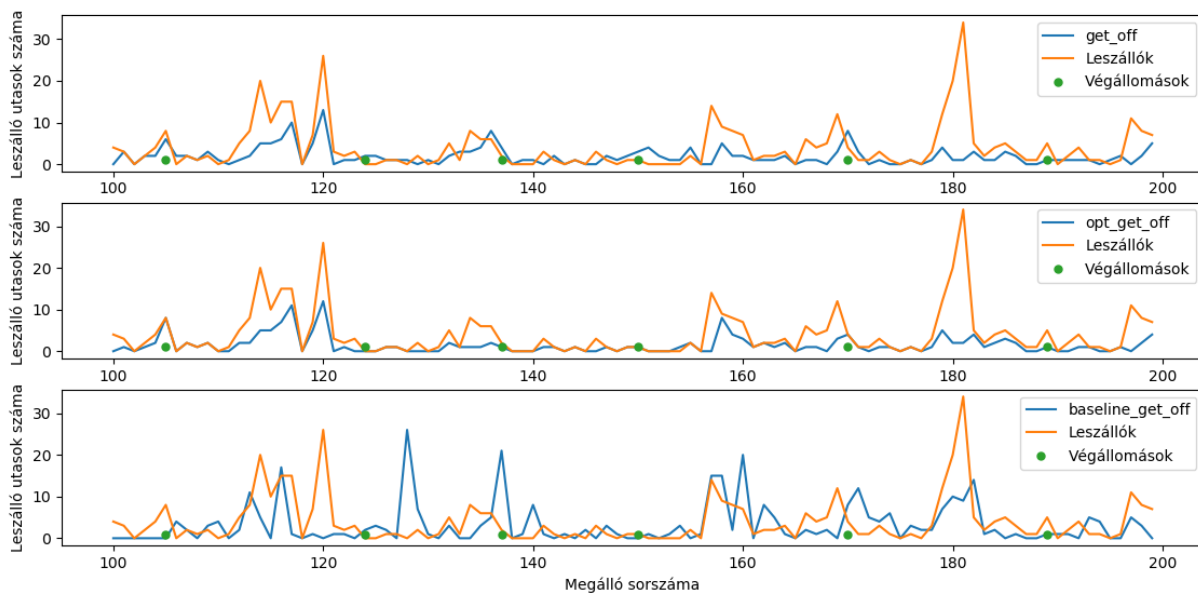
	felszállók RMSE	leszállók RMSE	felszállók Pearson	leszállók Pearson	terheltség Pearson	terheltség RMSE
Baseline	5.2	5.3	0.33	0.26	0.00249	31.53
WiFi threshold	4.09	4.32	0.556	0.406	0.667	14.96
WiFi opt	3.8	4.01	0.761	0.653	0.814	15.88

A különböző módszerekkel kapott becslést az alábbi grafikonok szemléltetik terheltségre, fel és leszállásra (minden esetben az mc_load jelölésű narancssárga a manuális mérés grafikonja, az opt_load a WiFi mérés eredménye).

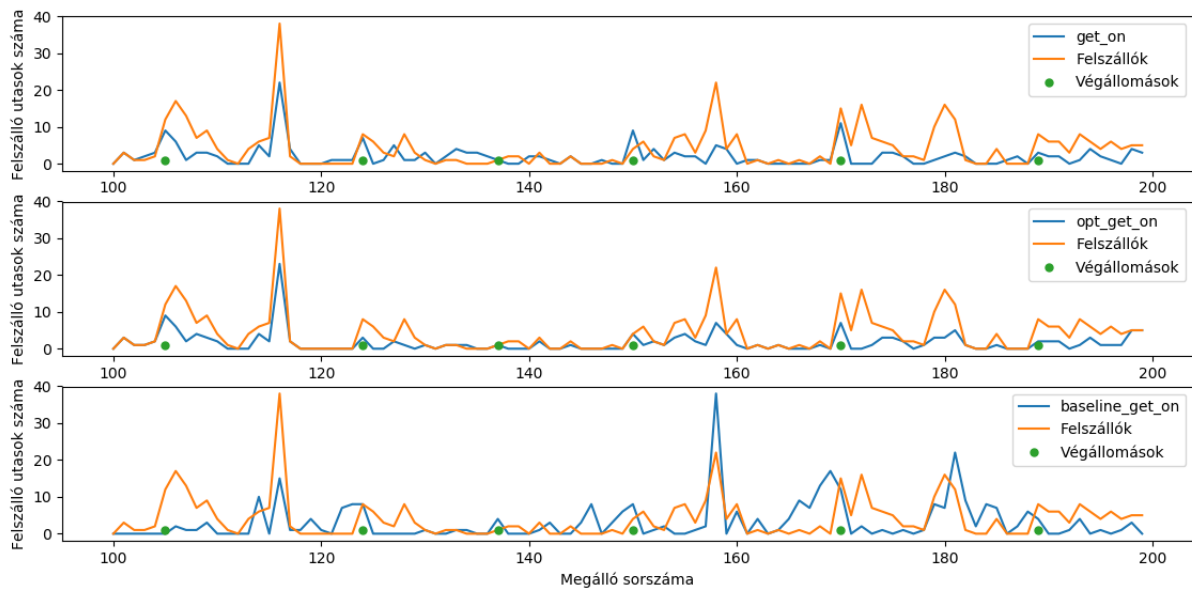


3. ábra: WiFi adatok alapján különböző módszerek által becsült terheltség összehasonlítása

Az ábrán látható, hogy a jármű terheltségi változásait a WiFi opt-load becslés jól leköveti, de módszeresen alábecsüli.



4. ábra: WiFi adatok alapján különböző módszerek által becsült leszállók részösszehasonlítása



5. ábra: WiFi adatok alapján különböző módszerek által becsült felszállók részösszehasonlítása

Összességben a grafikonok szabad szemmel is jól szemléltetik a becslés és a mérési adatok előzőekben már bemutatott korrelációját.

2.4 Az eredmények jósága és az adatvédelem

A bemutatott módszer olyan automatikus utasszámlálási megoldás, ami a járműveken elérhető szenzor adatok és WiFi adatok alapján becsüli az utasok számát. A módszer eredményeit összehasonlítottuk az adatrögzítés idejében az SZKT által gyűjtött utasszámlálási adatokkal. Az összehasonlítás során két metrika alapján végeztük a kiértékelést: négyzetes átlagos hiba négyzetgyöke (rmse) és Pearson korreláció. A kiértékelést az utasszámlálók által rögzített megállónkénti fel és leszállók, illetve terheltségre is elvégeztük.

A kísérletünk eredményei azt mutatják, hogy magas korreláció van a gyűjtött adatok és az SZKT által végzett manuális forgalomszámlálás között.

A WiFi adatok előnye, hogy a terheltségbecslés mellett lehetőséget ad az utasok pontosabb nyomkövetésére, mint a fel és leszállási megálló meghatározása. A WiFi adatok alapján kapott becslés általában alacsonyabb értéket ad az utasok valós számához képest, hiszen nem minden utas rendelkezik (vagy kapcsolja be) a WiFi képes eszközt.

A bemutatott megoldások mindegyike felülmúlja a Baseline megoldást, ahogy azt vártuk. Azonban a WiFi opt eredményei alapján azt is láthatjuk, hogy a WiFi threshold módszer még tovább javítható.

A wifi alapján történő monitorozási módszerek magukban hordozzák az utasok azonosítási lehetőségeit, amely esetlegesen személyiségi jogokat sérthet. Ezért fontos, hogy az utasokat biztosítsuk a jogaik védeltségéről és ismerjük a módját, hogyan kerülhetik el a résztvételt (pl.: wifi esetén repülőgép üzemmód bekapcsolása mint Londonban) ha esetleg ez sértené a személyes jogaikat. Mindezeket túl az utazók bizalmatlanságát és elégedetlenségét megelőzendő, az utasokat értesíteni kell, hogy pontosan milyen adat kerül gyűjtésre és milyen célból.

Az adatok közzététele előtt, a külső fél által történő azonosítás megelőzése érdekében szükséges egy anonimizációs lépés is a folyamat során. Ilyenkor az identifikációra alkalmas információ (pl.: mac cím) eltávolításra kerül.

Konklúzió

A cikket a Bevezetés és a Konklúzió rövid, legfeljebb 1 oldalas fejezetei keretezik. Ezeknek más címet nem kell adni, és nem is sorszámozódnak. Álláspontunk szerint a bemutatott módszerek között nincs olyan módszer, amely minden tekintetben domináns lehetne a többinél, sőt amennyiben lehetséges, érdemes több módszer együttes alkalmazása a kontroll adatok és a mérések optimalizálása érdekében. Így

a jövőben lehetőség nyílik a két vagy több módszer ellenőrzésére, főleg ha különböző szinteken (Linked vagy Unlinked trip) érkezik adat a közösségi közlekedési eszközök használatáról.

Az infravörös és a taposó lépcső lényegében hasonló lehetőségeket rejtenek magukban közel azonos pontosság mellett. A jármű dinamikai adatok és a kamerás képfelvételek használata adódik a fekete dobozokból gyűjtött adatkörökből, de ezektől a megoldásoktól sem várhatunk 99% feletti pontosságot, és nem alkalmasak „utazási hosszak” detektálására.

Az utasmédia (például smart card) és a „free WiFi” alkalmas lehet a célforgalmi mátrixok elkészítésére. A cikkben bemutatott WiFi ígéretes adatforrásnak tűnik, főleg ha minden járművön mint kényelmi szolgáltatás elérhető, azonban kiemelendő a mac cím randomizációs problémája, mely gátat jelenthet az alkalmazásában, és a GDPR megjelenésével a jövőben több személyiségi joggal kapcsolatos kérdést is tisztázni érdemes.

Irodalomjegyzék

- [1] Oransirikul, T., Nishide, R., Piumarta, I., & Takada, H. (2014). Measuring bus passenger load by monitoring Wi-Fi transmissions from mobile devices. *Procedia Technology*, 18, 120–125. <https://doi.org/10.1016/j.protcy.2014.11.023>
- [2] Kovács, R., Nádai, L., & Horváth, G. (2009). Concept validation of an automatic passenger counting system for trams. *Proceedings - 2009 5th International Symposium on Applied Computational Intelligence and Informatics, SACI 2009*, (4), 211–215. <https://doi.org/10.1109/SACI.2009.5136243>
- [3] Paulius, L., Rimvydas, S., Vygandas, V., & Rytis, M. (2013). Application of Computer Vision Systems for Passenger Counting in Public Transport. *Elektronika Ir Elektrotechnika*, 19(3), 1392–1215. <https://doi.org/10.5755/j01.eee.19.3.1232>
- [4] Yahiaoui, T., Khoudour, L., & Meurie, C. (2010). Real-time passenger counting in buses using dense stereovision. *Journal of Electronic Imaging* 19. Retrieved from https://www.researchgate.net/profile/Louahdi_Khoudour/publication/238982016_Real-time_passenger_counting_in_buses_using_dense_stereovision/links/56b85ef508ae3c1b79b2496f.pdf
- [5] Lempitsky, V., & Zisserman, A. (2010). Learning To Count Objects in Images. *Neural Information Processing Systems Conference*. Retrieved from <http://papers.nips.cc/paper/4043-learning-to-count-objects-in-images.pdf>
- [6] García, J., Gardel, A., Bravo, I., Lázaro, J. L., Martínez, M., & Rodríguez, D. (2013). Directional People Counter Based on Head Tracking. *IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS*, 60(9). <https://doi.org/10.1109/TIE.2012.2206330>
- [7] Oransirikul, T., Nishide, R., Piumarta, I., & Takada, H. (2014). Measuring bus passenger load by monitoring Wi-Fi transmissions from mobile devices. *Procedia Technology*, 18, 120–125. <https://doi.org/10.1016/j.protcy.2014.11.023>
- [8] Jamil, S., Khan, S., Basalamah, A., & Lbath, A. (2016). Classifying smartphone screen ON/OFF state based on wifi probe patterns. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing Adjunct - UbiComp '16* (pp. 301–304). New York, New York, USA: ACM Press. <https://doi.org/10.1145/2968219.2971377>