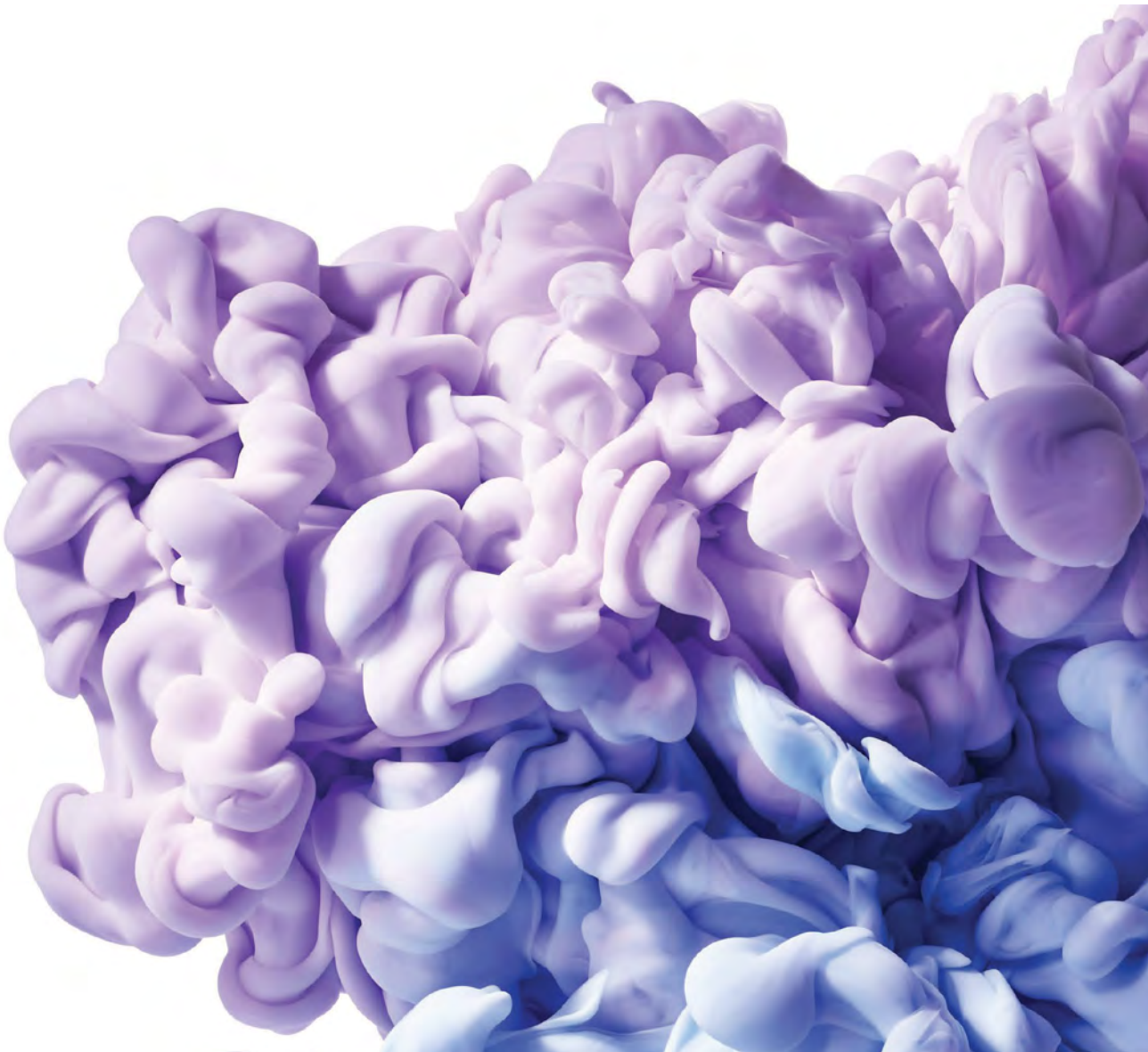
marketlocator

Analýza lokalizačných dát mobilných operátorov v Bratislavskom a Trnavskom kraji



marketlocator

Analýza lokalizačných dát mobilných operátorov v Bratislavskom a Trnavskom kraji pre potreby Bratislavskej integrovanej dopravy

(Vybrané výstupy predmetu zmluvy)

Úvod 2
Sieťová lokalizácia mobilného zariadenia 2
Využitie lokalizačných dát mobilných operátorov v dopravných priestorových analýzach 5
Metodika 6
Základná charakteristika dátovej vzorky 6
Limity dátovej vzorky 8
Výstupy 12
Tabuľkové a grafické výstupy 14
Extrahovanie spacej a pracovnej lokality 14
Extrahovanie polohy SIM karty v časových rezoch 17
Mapové výstupy 20
Zoznam použitej literatúry 27

Úvod

Poznanie toho, kde ľudia žijú a pracujú, ako aj ich každodenného pohybu v priestore sú kľúčovými vstupmi pre pre formulovanie rozmanitých potrieb v rámci územného plánovania, verejnej správy či managementu dopravy. Kým v minulosti sme sa mohli v analýzach oprieť o údaje z národného cenzu, dnes sa tieto tradičné údajové zdroje javia ako nedostatočné pre zachytenie dennej priestorovej dynamiky obyvateľstva. V súčasnosti sa pracovné či privátne aktivity obyvateľov sa neviažu na jedno miesto tak pevne, ako tomu bývalo v minulosti a miesto trvalého pobytu, miesto bývania a miesto práce sa čoraz častejšie od seba odlišujú. Dôvodom je predovšetkým rastúca fluktuácia a mobilita obyvateľstva. Pohyby obyvateľstva sa stávajú pestrejšími, nadobúdajú nerutinný a nepravidelný charakter. V tomto kontexte sa potreba aktuálnych, presných a spoľahlivých údajov o priestorovej mobilite javí ako nanajvýš potrebná.

Využitie informačno-komunikačných technológií – najmä mobilnej komunikácie vytvára nové príležitosti pre štúdium priestorovej mobility. Vďaka vysokej penetrácii mobilných telefónov v populácii a schopnosti sledovať ich pohyb na úrovni antén mobilnej siete môžeme prekonať viaceré limity, ktoré sa spájajú s tradičnými prístupmi založenými na populačných cenzoch, najmä čo sa týka frekvencie zisťovania populačných dát, rýchlosti ich spracovania a v neposlednom rade i ochoty obyvateľov poskytovať presné údaje v rámci cenzov. Bezprecedentný rozsah pokrytia populácie prostredníctvom mobilnej siete vytvára unikátny predpoklad pre komplexný prístup sledovania miest a regiónov, čo je principiálne nemožné (a rovnako aj finančné náročné) s využitím tradičných foriem získavania dát.

Sieťová lokalizácia mobilného zariadenia

Princíp sieťovej lokalizácie v mobilnej sieti je prirodzenou nadstavbou základných vlastností mobilnej komunikácie. Z priestorového pohľadu môžeme územie rozdeliť do buniek (*cells* – odtiaľ *cellular network*, resp. *cell phone*), ktoré obsluhujú jednotlivé antény mobilnej siete (*Base Transceiver Station – BTS*). Určenie polohy mobilného zariadenia v sieťovej lokalizácii prebieha nasledovne. Zariadenie s aktívnou SIM kartou sa pripája na BTS stanicu. Spravidla ide o najbližší vysielač, nemusí to však byť pravidlom. Niektoré BTS stanice totiž slúžia na prenos dát (napr. 4G bunky), iné na prenos hovorov (2G/3G). Zároveň sieť autonómne distribuuje signál pripojeným zariadeniam podľa svojho vyťaženia. Každá bunka distribuuje signál rôznej intenzity rôznymi smermi. Plochu, na ktorú žiari nazývame vyžarovací polygón, a vždy ide o nepravidelný útvar, ktorého tvar je výsledkom smerovania a intenzity signálu, ako aj charakteru prírodných podmienok (najmä reliéfu, zástavby). Keďže výpočty s

nepravidelnými polygónmi sú pomerne náročné, polygóny podliehajú vyhladzovaniu a matematickému zjednodušovaniu, aby s nimi bolo možné jednoduchšie pracovať.

Lokalizačné údaje, ktoré môžeme získať prostredníctvom sieťovej lokalizácie, môžeme principiálne rozdeliť na aktívne a pasívne (Ahas et al. 2007a). Kým pri pasívnom type ide o využitie existujúcich záznamov v rámci systému mobilného operátora, pri aktívnom type sa záznam vytvára na základe konkrétneho dopytu a s využitím špecializovaného softvéru. Osobitne môžeme uviesť ešte záznamy, ktoré sa zhromažďujú za jednotlivé prenosové časti mobilnej siete – štatistiku antén (Novák 2010).

Štatistika antén: Prevádzka mobilnej siete si vyžaduje sledovanie vyťaženosti jednotlivých prenosových staníc (BTS) mobilnej siete. K zaznamenanému počtu a dĺžke hovorov (či iných dátových prenosov) tak vieme priradiť priestorový atribút – polohu danej antény BTS. To však v praxi nemusí byť jednoduché, keďže rozmiestnenie prenosových staníc môže byť súčasťou neverejných strategických informácií mobilného operátora. V takom prípade je možné bodový charakter údajov transformovať do fiktívnych polygónov prenosovej siete (napr. Thiessenove polygóny) alebo pristúpiť k agregácii údajov do iných priestorových jednotiek (napr. administratívneho členenia). Možnosti využitia štatistiky antén sú napríklad vo výskume dennej dynamiky mesta, ako perspektívne sa javí aj prepojenie týchto lokalizačných dát s údajmi zo sčítania obyvateľstva (charakter zástavby, socio-demografické charakteristiky), či s databázou funkčného využitia zeme (Novák 2010).

Pasívna lokalizácia: Obdobne ako v prípade štatistiky antén ide o informácie, ktoré v realite už existujú. Informácie o činnosti telefónu slúžia ako podklad pre vyúčtovanie hovorov, SMS a pod. Vzhľadom na skutočnosť, že k údajom o aktivite mobilného zariadenia vieme priradiť nielen polohu, ale aj niektoré základné informácie o užívateľovi (vek, pohlavie, či fakturačná adresa), získavame bohatú databázu, ktorej využitie (pri zachovaní anonymity užívateľov, resp. pri čiastočnom agregovaní údajov) prináša nesmierne cenný nástroj pre územné plánovanie a manažment mesta. Príkladom môže byť využitie dennej a nočnej lokalizácie užívateľov, ktoré vytvárajú základnú kostru ich každodenných aktivít (domov – práca). Údaje o koncentrácii dennej a nočnej lokalizácii nám umožňujú nielen spresniť priestorovú distribúciu obyvateľstva (napr. koľko ľudí býva v danom meste / v blízkosti nákupného centra a pod.), ale aj extrahovať údaje o predpokladanej priestorovej mobilite (napr. koľko ľudí dochádza do mesta / do nákupného centra). Pasívne lokalizačné údaje sa tak stávajú cenným zdrojom informácii pre O-D dopravné modely, ale aj pre monitoring aktuálnej dopravnej situácie. Pohyb mobilného zariadenia medzi jednotlivými bunkami siete umožňuje sledovať rýchlosť a intenzitu dopravy a v súčasnosti je už bežnou súčasťou aplikácií poskytujúcich navigáciu. Pasívne lokalizačné údaje však ponúkajú oveľa väčšiu škálu využitia. Mobilný telefón dnes používame nielen na komunikáciu, ale aj nakupovanie, zábavu, či športové aktivity. Využitím niektorých doplnkových údajov, ktoré

zaznamenávajú operátori vo svojich databázach, tak môžeme sledovať aj oveľa špecifickejšie správanie ľudí, ako napr. využitie verejnej dopravy, parkovanie, návštevnosť parkov a pod.

Pasívne lokalizačné údaje mobilného zariadenia pri sieťovej lokalizácii vznikajú dvoma základnými spôsobmi:

1. **Pingové alebo signalizačné dáta** (*ping data*) sú automaticky generované záznamy, ktoré produkuje mobilná sieť pri pravidelných kontrolách pripojených zariadení. Jednotlivé BTS stanice v pravidelných intervaloch (definovaných oblasťou/operátorom) vysielajú signál na všetky dostupné a prihlásené zariadenia. Na základe ich polohy a vyťaženosti siete sa určí, ktorá BTS stanica bude konkrétnemu zariadeniu poskytovať signál, čím sa zabezpečuje čo najlepšie pokrytie pre každé zariadenie v dosahu. Prednosť má pritom pripojenie k anténe s novšou technológiou prenosu pred intenzitou signálu. Tieto dáta vznikajú bez interakcie užívateľa so zariadením. SIM karta vygeneruje záznam pri každej zmene BTS stanice, čím sa vygenerujú desiatky až stovky záznamov denne. Vzhľadom na veľký počet takýchto záznamov je ich spracovanie pomerne náročné. Napriek tomu ich význam v priestorových analýzach neustále narastá. Dôvodom sú zmeny vo využívaní mobilného telefónu. Kým v minulosti sme ho využívali najmä na volania a SMS správy, v súčasnosti toto využitie ustupuje v prospech dátových služieb (sociálne siete, instant chat a pod.) a pasívnemu využívaniu telefónu (napr. rôzne notifikácie v aplikáciách si nevyžadujú aktívnu interakciu).

2. **CDR dáta** (*Call-detail-record*) sú záznamy, ktoré vznikajú pri akejkoľvek aktivite mobilného telefónu (SIM karty) s mobilnou sieťou. Za takéto aktivity považujeme uskutočnený / prijatý (neprijatý) hovor, doručenie / zaslanie SMS, či dátový prenos. CDR záznam teda nevzniká automaticky, ale výlučne pri aktivite telefónu (SIM karty). Takýchto záznamov preto existuje podstatne menej, než pri signalizačných dátach. Počas bežného dňa SIM-karta vygeneruje rádovo desiatky CDR záznamov.

Aktívna lokalizácia: Kým pri pasívnej lokalizácii vychádzame z databázovej infraštruktúry prevádzkovateľa mobilnej siete (ide o anonymizované, či agregované dáta), pri aktívnej lokalizácii vznikajú záznamy o polohe individuálneho mobilného zariadenia na základe cieleného lokalizačného dopytu a s využitím špecializovaného softvéru. Nevyhnutnou podmienkou záznamu je informovaný súhlas užívateľa mobilného zariadenia a presne stanovené podmienky charakteru a dĺžky záznamu. Právne a etické zásady spojené s aktívnou lokalizáciou diskutuje Ahas et al. (2007b), či Novák (2010), ktorí priniesli aj inšpiratívne pilotné sondy využívajúce aktívnu mobilnú lokalizáciu. Ficek et al. (2013) sa zaoberajú využitím efektívneho zisťovania polohy mobilného užívateľa využitím textových správ (SMS). Tento spôsob prekonáva základný nedostatok prevádzkových údajov mobilných sietí – záznamov o aktivite mobilného zariadenia – ktorých časové rozostupy (frekvencia záznamov) neumožňujú presné určenie polohy užívateľa počas dňa.

Priestorová presnosť určovania polohy mobilného zariadenia je závislá od technických parametrov mobilnej infraštruktúry, predovšetkým od granularity buniek mobilnej siete (hustota rozmiestnenia BTS staníc v priestore) a ich technológie (2G/3G/4G). Vo všeobecnosti je lokalizácia najpresnejšia v mestách, ktoré sú pokryté zväčša BTS stanicami s technológiou 4G, pričom hustota lokalizácie BTS staníc je vysoká.

Využitie lokalizačných dát mobilných operátorov v dopravných priestorových analýzach

Súčasná spoločnosť je viac ako kedykoľvek predtým vytváraná tokmi ľudí, tovarov a informácií, avšak dáta, ktoré tieto toky zaznamenávajú, sú pomerne náročné na spracovanie využitím tradičných prístupov spoločenského výskumu. Výsledkom adaptácie výskumných metód na nové podmienky je predovšetkým využitie informačno-komunikačných technológií, ktoré prinášajú nielen zmenu paradigmy priestorového správania, ale aj prístupu k jeho sledovaniu. Osobitné postavenie v rámci priestorových analýz má využitie mobilnej komunikácie. Mobilný telefón sa stal neoddeliteľnou súčasťou každodenného života a unikátnym zdrojom údajov o obyvateľstve, jeho priestorovej distribúcii, pohybe a aktivitách.

Lokalizačné údaje mobilných telefónov priniesli možnosť sledovať vzory priestorovej mobility na individuálnej úrovni (González et al. 2008 a Song et al. 2010), ako aj mapovanie pohybu a aktivít obyvateľov využitím agregovaných údajov (Järv et al. 2012, Ratti et al. 2006 a Reades et al. 2009). Značné uplatnenie našlo využitie mobilných dát v dopravných štúdiách (Järv et al 2012 a Calabrese et al. 2011), ale aj v oblastiach ako je krízový manažment (Bengtsson et al. 2011). Napriek rozmanitému využitiu sa lokalizačné údaje mobilných zariadení len v obmedzenej miere využili na plošné mapovanie časovo-priestorovej distribúcie obyvateľstva v rámci širších priestorových kontextov. Dôvodom môže byť nedostupnosť dostatočne presnej lokalizácie, ktorej presnosť závisí najmä od architektúry mobilnej infraštruktúry alebo prístupu k celoštátnym lokalizačným údajom.

Metodika

Základná charakteristika dátovej vzorky

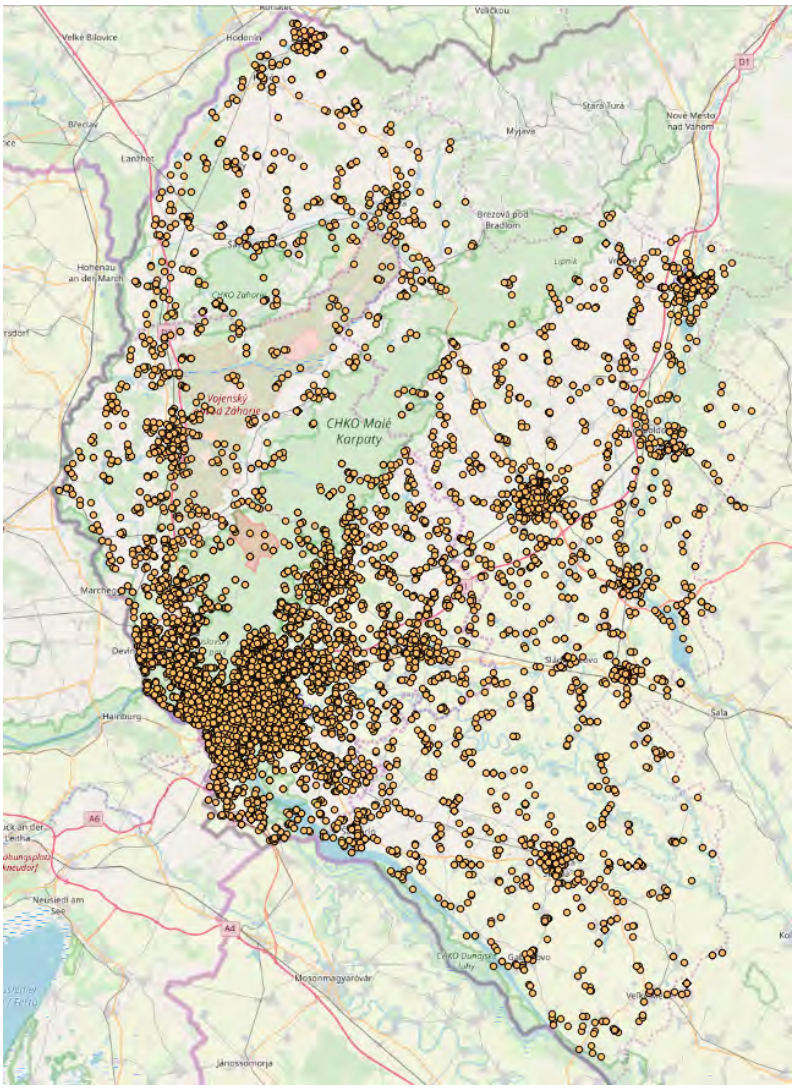
Predkladané výstupy boli spracované na základe dátovej vzorky signalizaćných lokalizačných dát o pohybe SIM kariet od troch najväčších mobilných operátorov na Slovensku (Slovak Telekom, Orange Slovensko a O2 Slovakia). Sledovaným obdobím bolo 14 po sebe nasledujúcich dní v rozmedzí 26.11.2018 (0:00) - 9.12.2018 (23:59), pričom sledovaným územím bol Bratislavský a Trnavský kraj.

Štruktúra dátovej vzorky bola nasledovná,

subscriber_ID	age	contract_type	date	hour	cell_id
a5gfh78b5trf	18-25	P	2018-11-29	11	13590
nj6j7kl9mn23	-	C	2018-12-03	13	34568
b5nh6sf7v35	35-45	P	2018-11-27	5	8742

pričom atribút **subscriber_id** reprezentoval jedinečný identifikátor držiteľa SIM karty registrovanej u jedného z troch analyzovaných operátorov (anonymizovaný hashovacou funkciou, ktorú pozná len zdrojový operátor), **age** reprezentoval vekovú kategóriu držiteľa SIM karty v požadovaných vekových kategóriách (veková kategória je vypočítaná na základe rodného čísla uvedeného v zmluve s operátorom, pričom pri služobných kontraktoch sa tento údaj nevypĺňa), **contract_type** reprezentoval typ kontraktu držiteľa SIM karty (súkromný / služobný), **date** predstavoval dátum, kedy bol záznam o polohe SIM karty zaznamenaný, **hour** predstavuje hodinu, kedy bol záznam o polohe SIM karty zaznamenaný a **cell_id** predstavuje identifikátor BTS stanice, na ktorej bol záznam o polohe SIM karty zaznamenaný.

Nevyhnutnou súčasťou dátovej vzorky boli aj priestorové atribúty vyžarovacích polygónov BTS staníc potrebné na ich namapovanie v priestore. Ako je vysvetlené nižšie, priestorové dáta sme kvôli anonymizácii a potrebnej agregácii vzťahovali na ťažiská vyžarovacích polygónov BTS staníc. Báza priestorových jednotiek (ťažísk), ktoré vstupovali do predkladaných výstupov tvorila vyše 15 tisíc bodov sumárne za všetkých troch analyzovaných operátorov.



Obr 1. Pokrytie sledovaného územia BTS stanicami (body na mape reprezentujú ťažiská vyžarovacích polygónov BTS staníc)

Predkladané výstupy boli spracované na základe pingových (signalizačných) dát. Frekvencia záznamov lokalizačných dát o pohybe SIM kariet bola vopred definovaná objednávateľom a požadovaná u operátorov, pričom základnou jednotkou bola 1 hodina (reprezentovaná cell_id, na ktorej mala SIM karta počas danej hodiny najviac záznamov, resp. reprezentovaná každou cell_id, na ktorej mala SIM karta počas danej hodiny záznam o polohe).

Veľkosť dátovej vzorky možno zhrnúť v nasledujúcej tabuľke 1. Sumárne mala dátová vzorka viac ako 640 miliónov riadkov, nad ktorými prebiehajú rôzne agregačné nápočty.

Počet analyzovaných subscriber_id	2 246 846
Počet analyzovaných subscriber_id, ktorým na základe zvolenej metodológie bolo možné napočítať priemerný pracovný deň (pozri kapitola: Výstupy)	1 402 705
Počet analyzovaných subscriber_id, ktorým na základe zvolenej metodológie bolo možné napočítať priemerný víkendový deň (pozri kapitola: Výstupy)	1 232 761
Počet analyzovaných subscriber_id, ktorým na základe zvolenej metodológie bolo možné napočítať pracovnú lokalitu	1 458 443
Počet analyzovaných subscriber_id, ktorým na základe zvolenej metodológie bolo možné napočítať spaciu lokalitu	1 402 705
Počet analyzovaných subscriber_id, ktorým na základe zvolenej metodológie bola napočítaná rovnaká spaciu a pracovná lokalita	592 730

Tab. 1. Počet unikátnych SIM kariet podľa definovaných kategórií

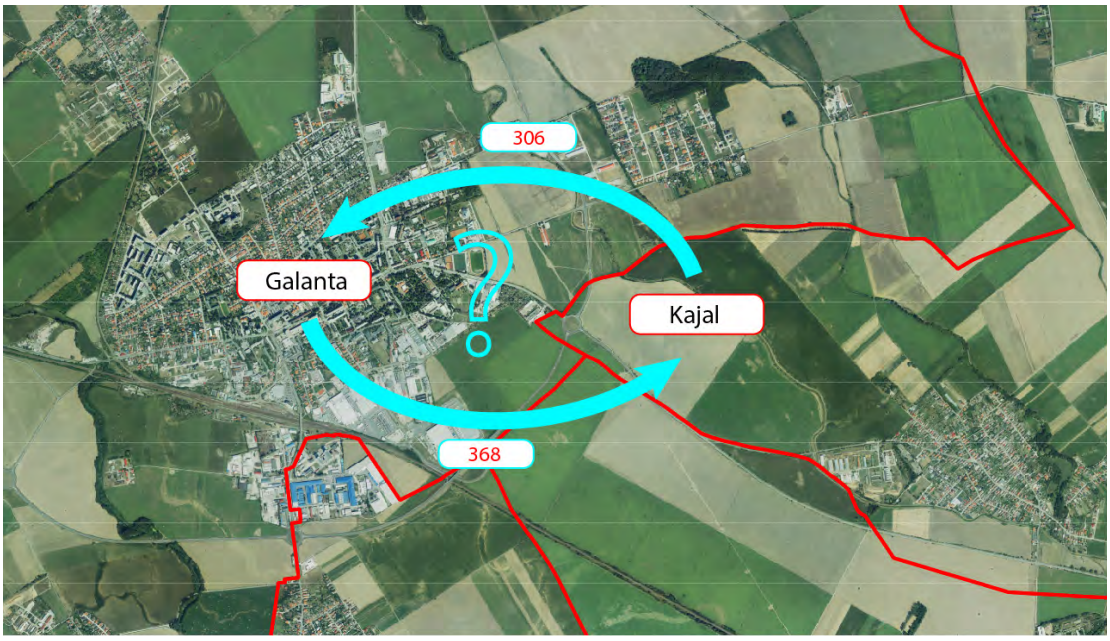
Limity dátovej vzorky

Pri interpretácii lokalizačných údajov mobilných zariadení v rámci sieťovej lokalizácie je dôležité si uvedomiť viaceré limity, ktoré vyplývajú z princípov prevádzky mobilnej siete a z právnych podmienok spracovania údajov jej užívateľov. Pokúsime sa upozorniť na tie najdôležitejšie:

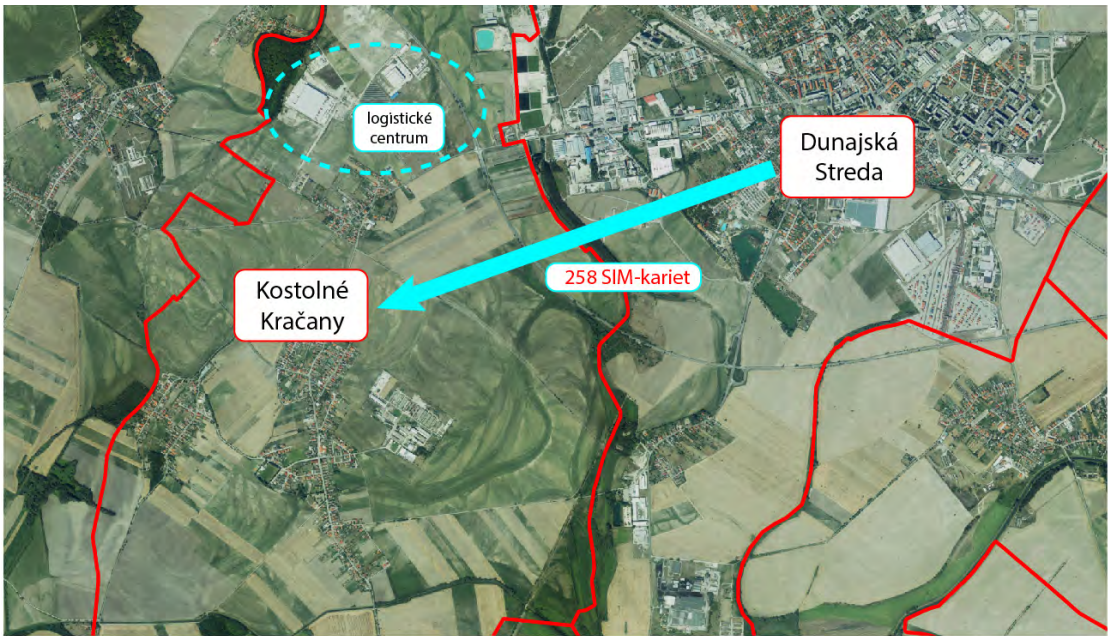
- Počet SIM-kariet nemožno spoľahlivo stotožniť s počtom osôb (individuálnych užívateľov). Dôvodom je skutočnosť, že nemôžeme predpokladať, že každý obyvateľ disponuje mobilným zariadením. Tiež nedokážeme spoľahlivo vylúčiť osoby využívajúce viaceré SIM-karty. Čiastočnú elimináciu duplicitných SIM kariet sme však dosiahli vylúčením nehlasových a neaktívnych SIM kariet.
- Užívateľ SIM-karty nemusí zodpovedať držiteľovi SIM-karty. Napríklad v rámci jednej rodiny môžu deti využívať mobilné telefóny registrované na rodičov. Ak je teda jedna osoba vlastníkom viacerých SIM-kariet, všetky nesú demografické atribúty, ktoré vychádzajú zo zmlúv, resp. kontraktov s operátormi. Poddimenzovanosť detskej zložky je skutočnosť, ktorú je potrebné zohľadniť pri interpretácii výsledkov analýz.
- Pre zachovanie kompletnosti dátovej vzorky bolo u jedného z operátorov potrebné dopočítať chýbajúce záznamy o polohe SIM karty na BTS stanici v

niektorých hodinových agregátoch pri niektorých SIM kartách (napr. chýbajúce údaje medzi 16:00 - 22:00). Po konzultáciách s dátovým špecialistom od zdrojového operátora sme pristúpili k dopočítaniu chýbajúceho atribútu cell_id podľa cell_id na ktorom mala daná SIM karta posledný záznam. Týmto spôsobom došlo k nápočtu približne 15% dátovej vzorky daného operátora.

- Presnosť lokalizácie závisí od architektúry mobilnej infraštruktúry, typu antén (2G, 3G, 4G) a reliéfu. V prípade mestského prostredia sa pohybuje rádovo v stovkách metroch, vo voľnej krajine dosahuje rádovo kilometre. Využitie údajov z mobilnej siete v priestorových analýzach by malo rešpektovať uvedené limity.
- Určenie konkrétnej polohy SIM karty sme vykonali na základe priestorových informácii o BTS staniciach. Vzhľadom na skutočnosť, že samotná poloha BTS staníc ako aj ich vyžarovacie polygóny sú pre jednotlivých operátorov citlivým údajom, kvôli zachovaniu anonymity sme akúkoľvek kvantifikáciu vzťahovali k ťažisku vyžarovacieho polygónu jednotlivých BTS staníc. Keďže požadovanou výstupnou priestorovou jednotkou boli obce, ťažiská vyžarovacích polygónov sme priradili do katastrálneho územia obce, v rámci ktorého sa nachádzali. Je však potrebné poznamenať, že pri využití danej metódy agregácie priestorových dát (vyžarovací polygón → centroid → kataster obce) dochádza k určitému skresleniu výslednej polohy SIM karty, pričom jej reálna poloha môže byť aj v inej (susednej) obci. Pri interpretácii extrahovaných dochádzkových tokov medzi obcami regiónu je tak dôležité prihliadať na presnosť lokalizácie, ktorá je limitované veľkosťou a tvarom katastrálneho územia obce. Osobitne v prípade susediacich obcí, sídelne zrastených obcí, či obcí excentricky umiestnených vzhľadom na tvar katastrálneho územia, je potrebné interpretovať toky osôb s prihliadnutím na priestorové súvislosti. Príkladom je obec Kajal v susedstve mesta Galanta (obr. 2). Extrahované toky s veľkosťou presahujúcou 300 osôb sú pravdepodobne dôsledkom pretiahnutého tvaru katastrálneho územia obce Kajal, ktorý zasahuje do priemyselnej zóny mesta Galanta. Nie každý nadpriemerne veľký tok do malej obce však musí znamenať nepresnosť v lokalizácii. Príkladom je obec Kostolné Kračany (1300 obyv.), kde môžeme zvýšenú dochádzku z Dunajskej Stredy vysvetliť prítomnosťou logistického parku na území tejto malej obce (obr. 3). Uvedené príklady dokumentujú, že pri interpretácii lokalizačných údajov mobilnej siete je dôležité zohľadniť použitú metodiku (priradenie vyžarovacích polygónov do kat. územia obcí) ako aj priestorový kontext a lokálne špecifiká.



Obr. 2. Priradenie ťažiska vyžarovacích polygónov mobilnej siete do hraníc katastrálnych území môže spôsobiť nepresnosti v alokovaní SIM-kariet a z toho vyplývajúcich denných dochádzkových tokov. Je pravdepodobné, že časť extrahovaných záznamov o dennej či nočnej lokalizácii v meste Galanta bola priradená do územia susednej obce Kajal.



Obr. 3. Zvýšená docháchadzka z Dunajskej Stredy do obce Kostolné Kračany môže byť dôsledkom prítomnosti logistického centra na území obce.

– Anonymizácia dát, ktorú bolo vzhľadom na citlivú povahu zdrojových údajov vykonať. Vzhľadom na citlivosť údajov, sme pre zabezpečenie ochrany jednotlivca pri záverečných dátových výstupoch pristúpili k čiastočnej

anonymizácii dát. Výstupné prechodové matice, ktoré zobrazujú súčet unikátnych SIM kariet v danej filtrácii (napr. v hodine X boli užívatelia v mieste A, a v hodine X+1 v mieste B), nikdy nezobrazia kumulatívne číslo nižšie ako 10. Ak danú podmienku spĺňa menej ako 10 ľudí, softvér zobrazí číslo 10. Ak danú podmienku spĺňa presne 0 ľudí, softvér nepriradí žiadnu hodnotu a bunka ostane prázdna. Ak danú podmienku spĺňa viac ako 10 ľudí, softvér zobrazí konkrétnu hodnotu. Anonymizáciou sa snažíme predísť možnosti identifikácie jednotlivca na základe pohybu, prípadne ďalších atribútov, ale zároveň ju možno pokladať ako jednu z limitov štúdie.

– Pre komplexné pochopenie dochádzkových vzťahov Bratislavskeho a Trnavského kraja by bolo potrebné rozšíriť analyzované územie o priľahlé kraje (hlavne Nitriansky a Trenčiansky kraj). Je teda dôležité poznamenať, že kvatifikované dochádzkové toky do centier Bratislavy a Trnavy nemožno brať ako konečné, a reálne toky sú pravdepodobne vyššie.

Výstupy

Spracovanie lokalizačných údajov mobilných zariadení v prostredí geografických informačných systémov vytvorilo plastický obraz o dennej priestorovej mobilite populácie v Bratislavskom a Trnavskom kraji. Jednotlivé parciálne výstupy poukazujú na značné disproporcie v dennej dynamike oboch sledovaných regiónov a dokumentujú dominantné postavenie Bratislavy, ktorej spádový región výrazne presahuje hranice kraja.

Bratislava ako dominantné sídelné centrum vytvára sieť väzieb rôzneho druhu, intenzita ktorých klesá s rastúcou vzdialenosťou od jadra. Denný pohyb za prácou slúži ako nenahraditeľný syntetický indikátor, ktorý zastupuje celé spektrum vnútroregionálnych väzieb vyplývajúcich z denného cyklu života obyvateľov regiónu. Vďaka poznaniu denných dochádzkových tokov sme schopný lepšie porozumieť nielen dopravným väzbám v regióne, ale aj zložitejším funkčným vzťahom (priestorovej interakcii medzi mestom a zázemím), ktoré formujú tento dynamicky sa rozvíjajúci región Slovenska.

Analýza dennej priestorovej mobility prináša kľúčový pohľad na rozsah a intenzitu priestorových interakcií v Bratislavskom a Trnavskom kraji. Kľúčovým zistením je skutočnosť, že Bratislava formuje metropolitný región, ktorý zahŕňa nielen Bratislavský kraj, ale aj celý Trnavský kraj. Trnava, ako centrum nižšieho sídelného rádu, z hľadiska smerovania a intenzity denných dochádzkových tokov, nevytvára také väzby, ktoré by ju kvalifikovali na prirodzené (spádové) centrum regiónu. Potvrdzuje sa tak už viac krát formulovaná výhrada “umelého” vytvorenia Trnavského kraja, ktorý nerešpektuje prirodzené priestorové väzby. V tomto kontexte je požiadavka na hlbšiu integráciu verejnej dopravy medzi Bratislavským a Trnavským krajom nanajvýš opodstatnená.

Na vybraných tabuľkových a grafických výstupoch je prezentovaný pohľad na dennú dynamiku oboch sledovaných regiónov. Tieto základné analytické výstupy dokumentujú dominantné postavenie Bratislavy, do ktorej smerujú najintenzívnejšie denné toky obyvateľstva. Jej dochádzkový región siaha až na vonkajšie hranice Trnavského kraja. Ide o obce vzdialené viac ako 60 km od Bratislavy. Napriek značnej časovej náročnosti na každodennú dochádzku (viac ako 90 min.), nachádzame v tejto vzdialenosti pomerne silné dochádzkové toky.

Základný pohľad na priestorovú distribúciu obyvateľstva v Bratislavskom a Trnavskom kraji ponúka tab. 1. Hoci nie sú zaznamenaní užívatelia mobilnej siete vyčerpávajúcim

obrazom o populácii, poskytujú zmysluplnú aproximáciu reálnej distribúcie populácie. Ku koncu roka 2017 bolo v Bratislavskom kraji registrovaných k trvalému pobytu vyše 650 000 obyvateľov (ŠÚ SR). Počet SIM-kariet s dlhodobou nočnou lokalizáciou v tomto regióne dosiahol úroveň 862 000. Nočná lokalizácia však nezachytáva všetkých rezidentov v regióne. V tomto počte nie sú zahrnutí obyvatelia, ktorí mobilný telefón nevyužívajú (napr. malé deti), alebo ho využívajú len v obmedzenej miere (viď limity dátovej vzorky). Zohľadnením týchto skupín by sme pravdepodobne dostali ešte vyšší počet rezidentov lokalizovaných v sledovanom regióne. Zaznamenaný počet SIM-kariet s nočnou lokalizáciou tak naznačuje, že v Bratislave (ako aj v jej širokom zázemí) žije väčší počet obyvateľov, než udáva oficiálna štatistika.

Vzhľadom na odhadovaný objem neprihlásených rezidentov môžeme vnímať uvedené zistenia ako značne znepokojujúce. Pri 650 000 obyvateľoch Bratislavského kraja je 210 000 neregistrovaných rezidentov pomerne veľký objem. Navyše skutočnosť, že do Bratislavy a jej bezprostredného zázemia prichádzajú migranti zo všetkých regiónov, rozširuje dôsledky naznačených disproporcií na územie celého Slovenska. Keby sa zaviedla prihlasovacia povinnosť v mieste obvyklého pobytu, ktorá je bežná v mnohých krajinách a dopĺňa evidenciu obyvateľstva, pravdepodobne by sa radikálne zmenil aj náš celkový pohľad na distribúciu obyvateľstva na Slovensku.

Na základe analýzy lokalizačných údajov mobilnej siete môžeme konštatovať, že do Bratislavy denne dochádza vyše 139 000 obyvateľov z obcí Bratislavského a Trnavského kraja (tab. 1). Ak by sme zohľadnili aj dochádzajúcich zo vzdialenejších okresov (najmä Nitrianskeho, Trenčianskeho kraja, resp. blízkeho pohraničia v Maďarsku a Rakúsku), tak môžeme uvažovať o ešte väčšom počte dochádzajúcich, ktorí každodenne využívajú dopravnú infraštruktúru Bratislavy.

Pri bližšom pohľade na cieľové lokality dochádzky v rámci administratívneho členenia Bratislavy (graf. 1) pozorujeme, že najväčšie toky dochádzajúcich sústreďuje MČ Bratislava - Ružinov (26 % dochádzajúcich) a Bratislava - Staré Mesto (21 %). Do MČ Nové Mesto a MČ Petržalka dochádza len niečo cez 13 %, resp. 10 % osôb z celkového objemu tokov smerujúcich do Bratislavy. Ostatné mestské časti koncentrujú počas pracovného dňa úhrnom 28 % dochádzajúcich.

Do Trnavy dochádza na dennej báze približne 27 000 obyvateľov Bratislavského a Trnavského kraja, z toho dochádzajúci z Bratislavy tvoria viac ako jednu tretinu (9943). Veľkosť najvýznamnejších dochádzkových tokov pre ostatné mestá regiónu znázorňuje obr. 4. V rámci tejto úrovne pozorovania môžeme identifikovať väčšie toky osôb (SIM-

kariet) najmä medzi “tradičnými” dvojicami miest, ktorých intenzívne väzby majú hlboké historicko-sídelné väzby. Príkladom sú dvojice miest Skalica – Holíč a Pezinok – Modra (obr. 5).

Abstrahovaním od základnej kostry dochádzkových tokov, ktoré reprezentuje štruktúra tokov do Bratislavy a Trnavy, môžeme odhaliť viaceré tangenciálne toky, ktorých veľkosť (počet SIM-kariet) môže predstavovať dôležitý vstup do dopravných analýz. Príkladom je dochádzkový tok medzi Piešťanmi a Dunajskou Stredou, ktorý prekvapuje nielen svojou dĺžkou (90 km) a smerovaním, ale aj počtom identifikovaných osôb (294 SIM-kariet). Väčšina extrahovaných tangenciálnych tokov má však racionálne zdôvodnenie. Vo väčšine prípadov ide o toky z mesta do blízkej obce s väčším priemyselným či logistickým areálom. Príkladom je docházkový tok mezi Malackami a Lozornom (204 SIM-kariet) alebo medzi Dunajskou Stredou a Kostolnými Kračanmi (258 SIM-kariet). V oboch prípadoch sú obce sídlom logistického areálu, ktorý poskytuje pracovné príležitosti pre širšie okolie.

Použitie lokalizačných údajov mobilnej siete prináša potrebné analytické východiská, vďaka ktorými vieme lepšie porozumieť dynamickému prostrediu metropolitného regiónu Bratislavy. Na výsledkoch čiastkových analýz môžeme demonštrovať, že použité údaje poskytujú zmysluplné výsledky, v rámci ktorých sa neobjavili toky, ktoré by sa nedali racionálne interpretovať. Výsledky tak podporujú vhodnosť metodického postupu a jeho korektnú implementáciu. Údaje z mobilnej siete nám poskytujú aktuálne a presné záznamy o dennej priestorovej dynamike, ktoré nedokážeme získať z tradičných údajových zdrojov. Vhodnou interpretáciou a pri rešpektovaní limitov mobilnej lokalizácie dostávame cenný analytický nástroj pre priestorový a dopravný management.

Tabuľkové a grafické výstupy

Extrahovanie spacej a pracovnej lokality

- Spacia lokalita

Spacia lokalita predstavuje oblasť, v ktorej bola SIM karta najčastejšie lokalizovaná počas nočných hodín, pričom za nočné hodiny bol považovaný interval 23:00 - 5:59 v pracovných dňoch PO - PIA počas dvoch skúmaných týždňov.

Vstupom pre výpočet spacej lokality bolo 7 hodinových rezov so záznamami o polohe SIM karty na staniciach BTS stanici v danej hodine, počas 10 pracovných dní

dvojtýždňovej vzorky. Za každý hodinový agregát počas uvedených pracovných dní sme extrahovali BTS, kde SIM karta mala najčastejší výskyt a z týchto 56 hodnôt bola následne vybraná BTS stanica s najvyšším počtom záznamov o polohe SIM karty. Následne bola daná BTS stanica priradená do katastra obce, do ktorého zasahovalo ťažisko "spacej" BTS stanice. Výsledný kataster obce následne považujeme za spaciu lokalitu užívateľa, resp. SIM karty.

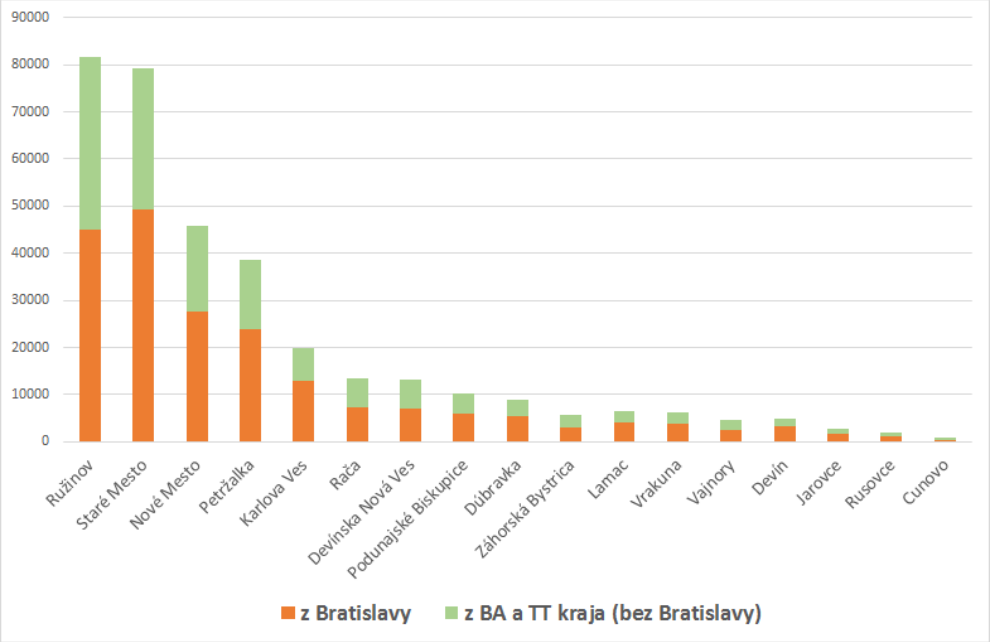
- Pracovná lokalita

Za pracovnú lokalitu považujeme tú oblasť, v ktorej bol najčastejšie telefón lokalizovaný v pracovných hodinách medzi 9:00 a 14:59 počas pracovných dní sledovanej dvojtýždňovej vzorky. Za pracovné dni sme považovali utorok, stredu a štvrtok, nakoľko sme vychádzali z predpokladu, že počas týchto pracovných dní sú dochádzkové toky najpravidelnejšie a nemali by byť ovplyvnené neštandardnou dochádzkou počas pondelka a piatka (predĺžený víkend, home office, skorší odchod a podobne).

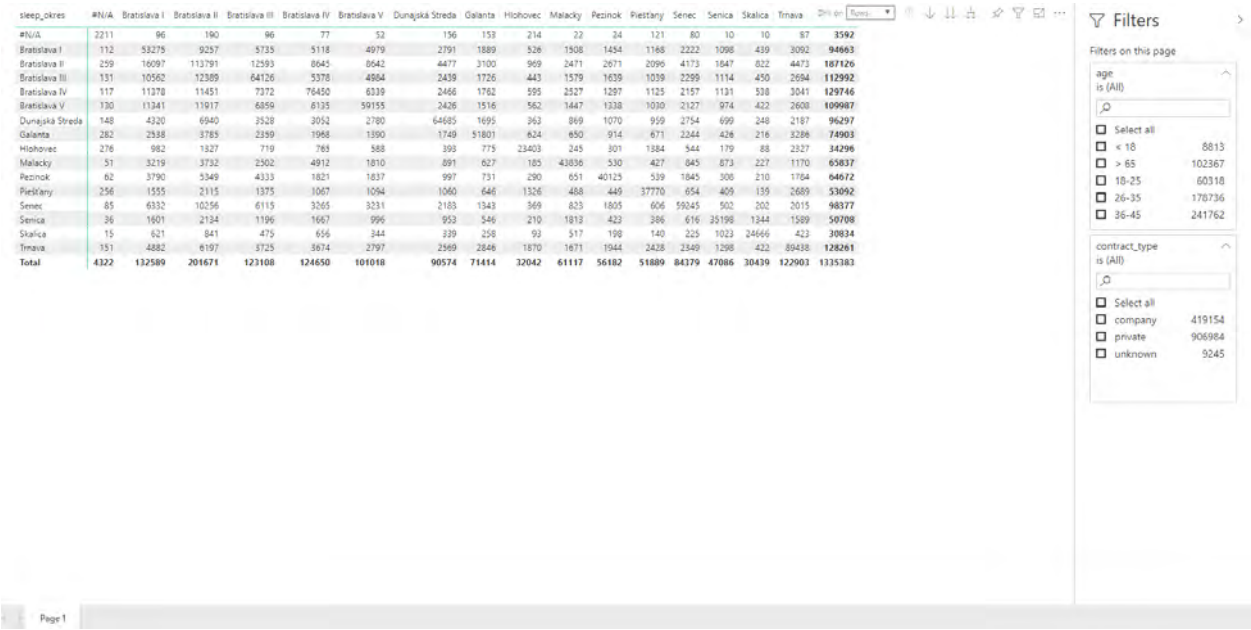
Vstupom pre výpočet pracovnej lokality bolo 6 hodinových rezov so záznamami o polohe SIM karty na staniciach BTS stanici v danej hodine, pre 6 pracovných dní dvojtýždňovej vzorky. Za každý hodinový agregát počas uvedených pracovných dní sme extrahovali BTS, kde SIM karta mala najčastejší výskyt a z týchto 36 hodnôt bola následne vybraná BTS stanica s najvyšším počtom záznamov o polohe SIM karty. Následne bola daná BTS stanica priradená do katastra obce, do ktorého zasahovalo ťažisko "pracovnej" BTS stanice. Výsledný kataster obce následne považujeme za pracovnú lokalitu užívateľa, resp. SIM karty. Pri kvantifikácii SIM kariet podľa pracovnej a spacej lokality, je potrebné uviesť, že hovoríme iba o tých SIM kartách, ktoré mali na základe zvolenej metodológie napočítanú aj pracovnú aj spaciu lokalitu na území Bratislavského a Trnavského kraja.

		Cieľový región (pracovná lokalita)				
		Bratislava	Bratislavský kraj (bez Bratislavy)	Trnava	Trnavský kraj (bez Trnavy)	SPOLU
Východiskový región (spacia lokalita)	Bratislava	543641	31015	9943	49207	633806
	Bratislavský kraj (bez Bratislavy)	62507	149467	3010	13636	228620
	Trnava	11604	3030	41031	10081	65746
	Trnavský kraj (bez Trnavy)	64810	17975	14065	304632	401482
	SPOLU	682562	201487	68049	377556	1329654

Tab. 2. Sumárne počty užívateľov mobilnej siete na základe ich východiskovej (spacej) a cieľovej (pracovnej) lokality.



Graf. 1. Podiel dochádzajúcich z Bratislavy a jej zázemia (Bratislavský a Trnavský kraj) do jednotlivých mestských častí Bratislavy.



sleep_okres	#N/A	Bratislava I	Bratislava II	Bratislava III	Bratislava IV	Bratislava V	Dunajská Streda	Galanta	Hlohovec	Malacky	Pezinok	Piešťany	Senec	Senica	Skalica	Trnava	
#N/A	2211	96	190	96	77	52	156	153	214	22	24	121	80	10	10	87	**3592**
Bratislava I	112	53275	9257	5735	5118	4979	2791	1889	526	1508	1454	1168	2222	1098	439	3092	**94663**
Bratislava II	259	16097	113791	12593	8645	8642	4477	3100	969	2471	2671	2096	4173	1847	822	4473	**187126**
Bratislava III	131	10562	12389	64126	5378	4984	2439	1726	443	1579	1639	1039	2299	1114	450	2694	**112992**
Bratislava IV	117	11378	11451	7372	76450	6339	2466	1762	595	2527	1297	1125	2157	1131	538	3041	**129746**
Bratislava V	130	11341	11917	6859	6135	59155	2426	1516	562	1447	1338	1030	2127	974	422	2608	**109987**
Dunajská Streda	148	4320	6940	3528	3052	2780	64685	1695	363	869	1070	959	2754	699	248	2187	**96297**
Galanta	282	2938	3785	2359	1968	1390	1749	51801	624	650	914	671	2244	426	216	3286	**74903**
Hlohovec	276	982	1327	719	765	588	393	775	23403	245	301	1384	544	179	88	2327	**34296**
Malacky	51	3219	3732	2302	4912	1810	891	627	185	43836	530	427	845	873	227	1170	**65837**
Pezinok	62	3790	5349	4333	1821	1837	997	731	290	651	40125	539	1845	306	210	1764	**64672**
Piešťany	256	1555	2115	1375	1067	1094	1060	646	1326	488	449	37770	654	409	139	2689	**53092**
Senec	85	6332	10256	6115	3265	3231	2183	1343	369	823	1805	606	59245	502	202	2015	**98377**
Senica	36	1601	2134	1196	1667	996	953	546	210	1813	423	386	616	35198	1344	1589	**50708**
Skalica	15	621	841	475	656	344	339	258	93	517	198	140	225	1023	24666	423	**30834**
Trnava	151	4882	6197	3725	3674	2797	2569	2846	1870	1671	1944	2428	2349	1298	422	89438	**128261**
Total	**4322**	**132589**	**201671**	**123108**	**124650**	**101018**	**90574**	**71414**	**32042**	**61117**	**56182**	**51889**	**84379**	**47086**	**30439**	**122903**	**1335383**

Obr. 4. Ukážka výstupov v softvéri Power BI. Výstup zobrazuje O-D maticu tokov medzi pracovnou a spacou lokalitou na úrovni okresov. V rovnakom výstupe je možnosť zobrazenia tokov medzi jednotlivými obcami alebo kombináciou obec/okres.V pravej časti okna je možná filtrácia podľa veku/typu kontraktu.

Extrahovanie polohy SIM karty v časových rezoch

- Nápočet hodinových rezov

Jedným zo základných pracovných výstupov tejto štúdie je matica presunov medzi (vopred zadefinovaný) polygónmi obcí. Prechodovú maticu si objednávateľ v prílohe zmluvy zadefinoval ako hodinové rezy najsilnejšieho dňa. Najsilnejší deň pre lepšiu výpovednú hodnotu nahradili dňom priemerným.

Následne sme pristúpili k napočítaniu hodinových agregátov zo všetkých eventov/udalostí. Jednotlivé datasety v niektorých prípadoch obsahovali viacero prislúchajúcich záznamov pripojení na BTS stanicu v danej hodine, a tak algoritmus vybral za reprezentatívnu BTS stanicu tú s najvyšším počtom výskytov v danej hodine počas vybraných dní. Priemerný pracovný deň sme pritom počítali zo 6 dní v rámci dvoch po sebe nasledujúcich týždňov, utorok až štvrtok. Pre konkrétny hodinový rez sme sa takto dopracovali k 6 záznamom BTS, a následne bola vybraná BTS s najvyšším počtom výskytov. Ak bola najčastejšia hodnota neznáma, daný hodinový rez ostal prázdny.

V jednotlivých datasetoch sa pritom vyskytujú všetci užívatelia, ktorí sa v danom čase nachádzali na území BA alebo TT kraja a ktorým bolo možné aspoň jeden slot napočítať.

- Nápočet 3-hodinových rezov

Pre praktickejšiu využiteľnosť zdrojových dát na identifikovanie vzorcov presúvania obyvateľstva, bola zo strany BID vznesená požiadavka na kalkuláciu trojhodinových slotov s najčastejšou polohou SIM kariet. Pristúpili sme teda ku kalkulácii časových agregátov nad pôvodnými zdrojovými dátami, kde:

timeslot_id	Časový rozsah (index hodín)
1	0-2
2	3-5
3	6-8
4	9-11
5	12-14
6	15-17
7	18-20
8	21-23

Do timeslotu 1 teda spadali všetky eventy SIM kariet vytvorené v čase od 0:00 do 2:59. Algoritmus spočítal unikátne udalosti na jednotlivých BTS staniciach a zoradil ich podľa poradia. Za polohu SIM karty v timeslote 1 teda považujeme tú BTS, ktorá mala najviac výskytov v danom hodinovom rozptyle, a teda strávila najviac času na tejto konkrétnej BTS. Nápočet trojhodinových slotov prebehol pre priemerný pracovný deň (celková vzorka 6 dní), aj pre priemerný víkendový deň (celková vzorka 2 dni).

- Priemerný pracovný deň v hodinových rezoch

Pre pochopenie cestovných návykov počas bežných pracovných dní sme sa rozhodli vypočítať polohy SIM kariet v hodinových rezoch. Konkrétny agregovaný hodinový záznam je pritom priemerom až 6 pracovných dní. Za vzorku sme totiž vybrali pracovné dni utorok až štvrtok počas dvoch po sebe nasledujúcich týždňov.

Použité BTS stanice v danej hodine každého jedného dňa zoradili podľa počtu výskytov, pričom sme pre danú hodinu konkrétneho dňa vybrali tú najčastejšiu BTS stanicu. Takto algoritmus postupoval pri každom unikátnom dni, až nakoniec podľa počtu výskytov zoradil dovedna šesť najčastejších BTS staníc v 6 dňoch. Finálnym vstupom teda bolo 6 najčastejších BTS staníc pre 6 dní - ako hľadaný priemer následne algoritmus vybral tú BTS stanicu, ktorá sa vo finálnom zozname ocitla najčastejšie.

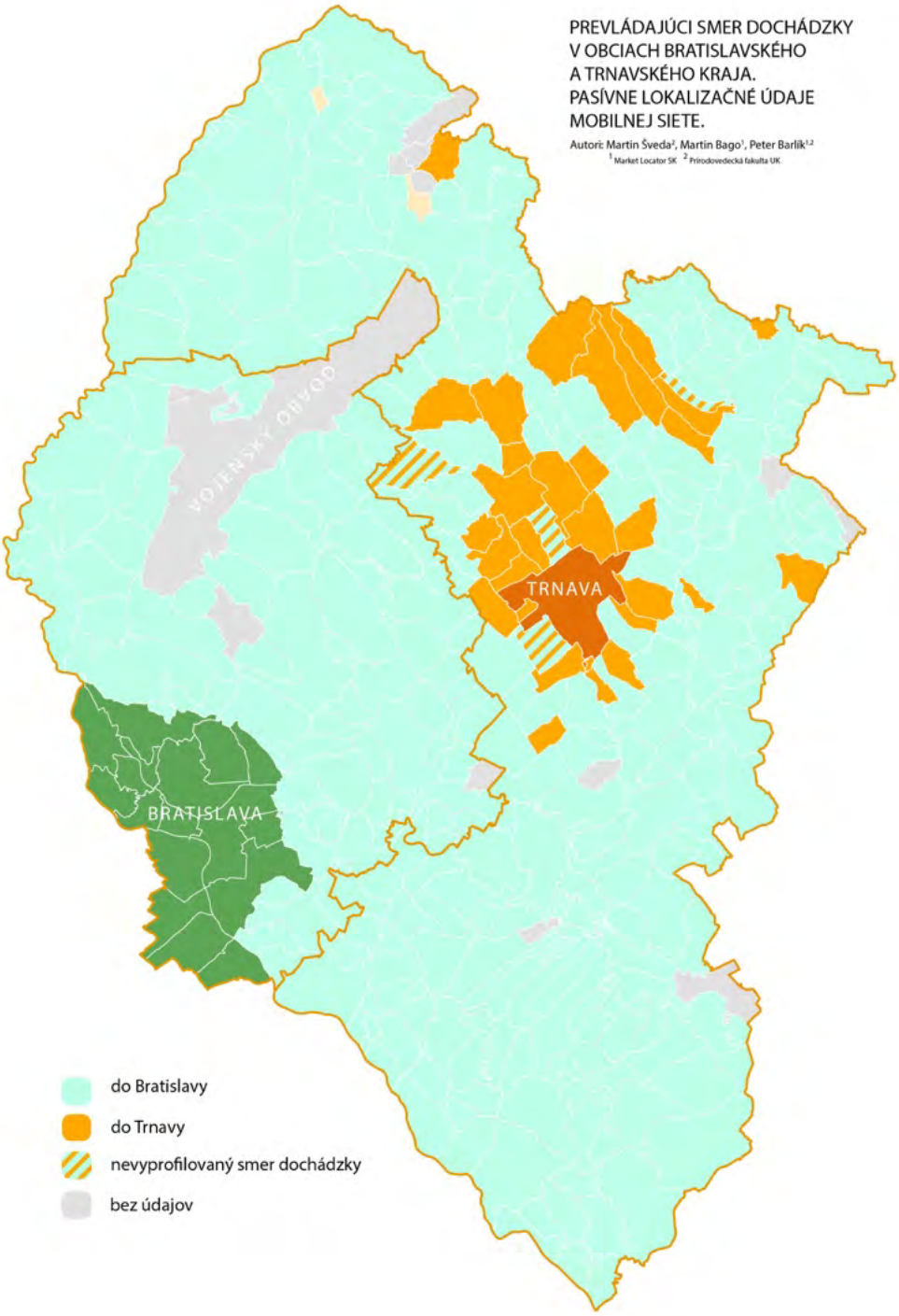
Pre konkrétnu BTS bol vypočítaný centroid jej vyžarovacieho polygónu, ktorý sme následne priradili k príslušnej obci, resp. Okresu.

- Priemerný víkendový deň v hodinových rezoch

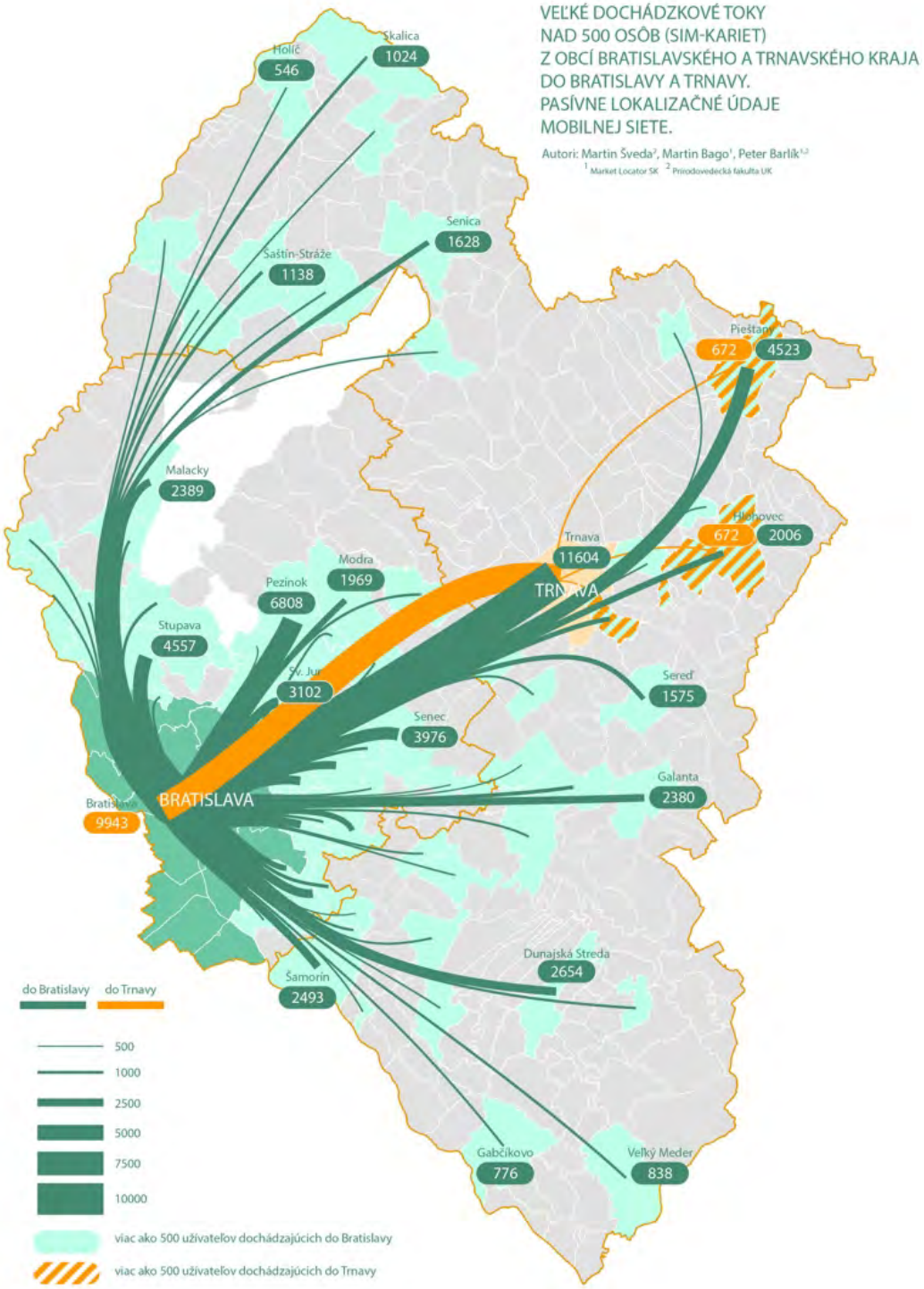
Pri výpočte priemerného pracovného dňa sme postupovali obdobne ako pri výpočte priemerného dňa, len vstupná časová vzorka bola odlišná - išlo o dve soboty v dvoch po sebe nasledujúcich dňoch.

Pre výpočet hodinového agregátu sme v tomto prípade použili všetky zaznamenané zmeny BTS staníc v daných hodinových rezoch, ktoré algoritmus zoradil podľa počtu výskytov a vybral najčastejšiu/najpoužívanejšiu BTS v danej hodine, počas dvoch sledovaných dní. Pre konkrétnu BTS bol vypočítaný centroid jej vyžarovacieho polygónu, ktorý sme následne priradili k príslušnej obci, resp. okresu.

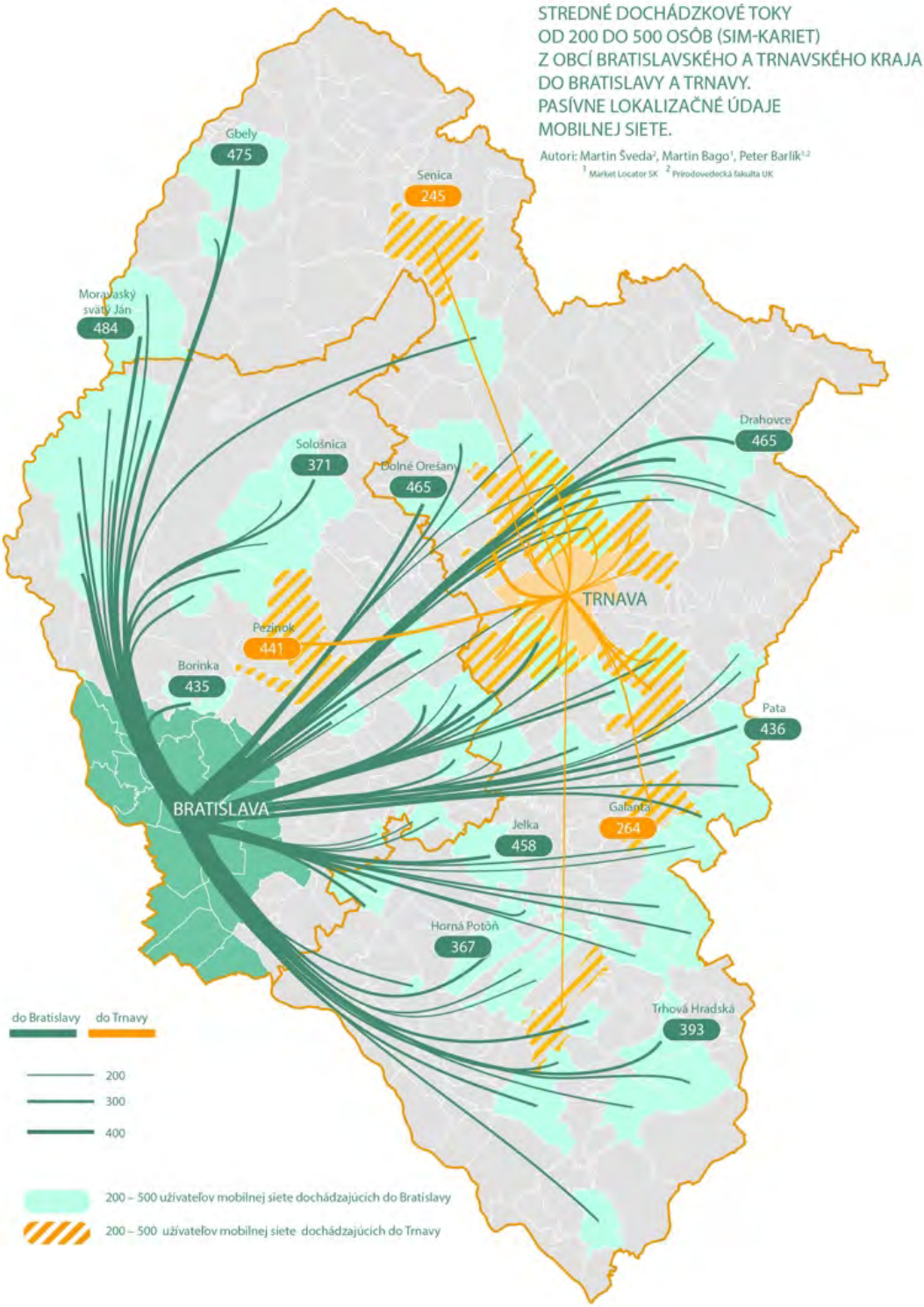
Mapové výstupy



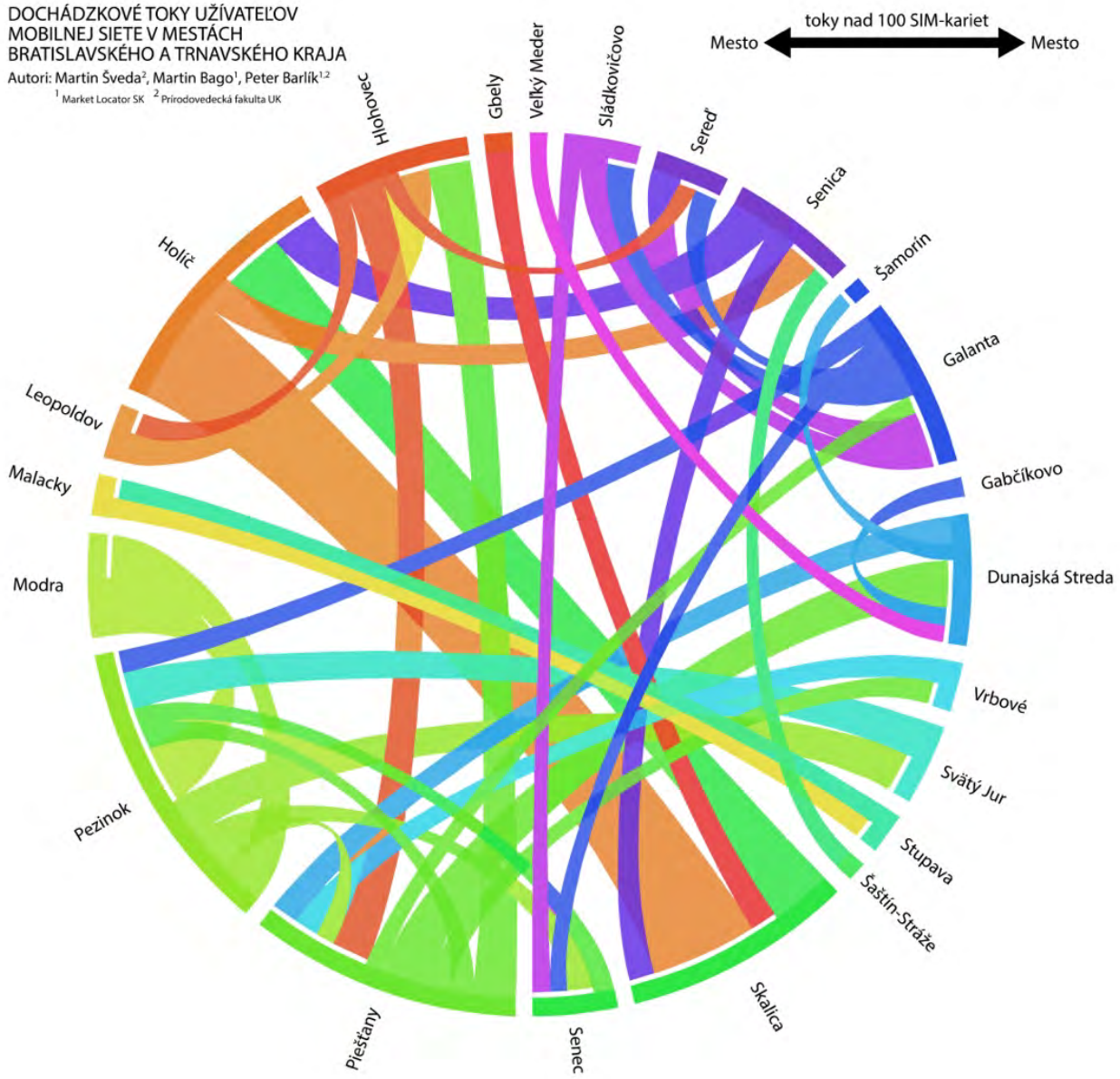
Obr. 5. Prevládajúci smer dochádzky z obcí Bratislavského a Trnavského kraja do Bratislavy a Trnavy. Pasívne lokalizačné údaje mobilnej siete.



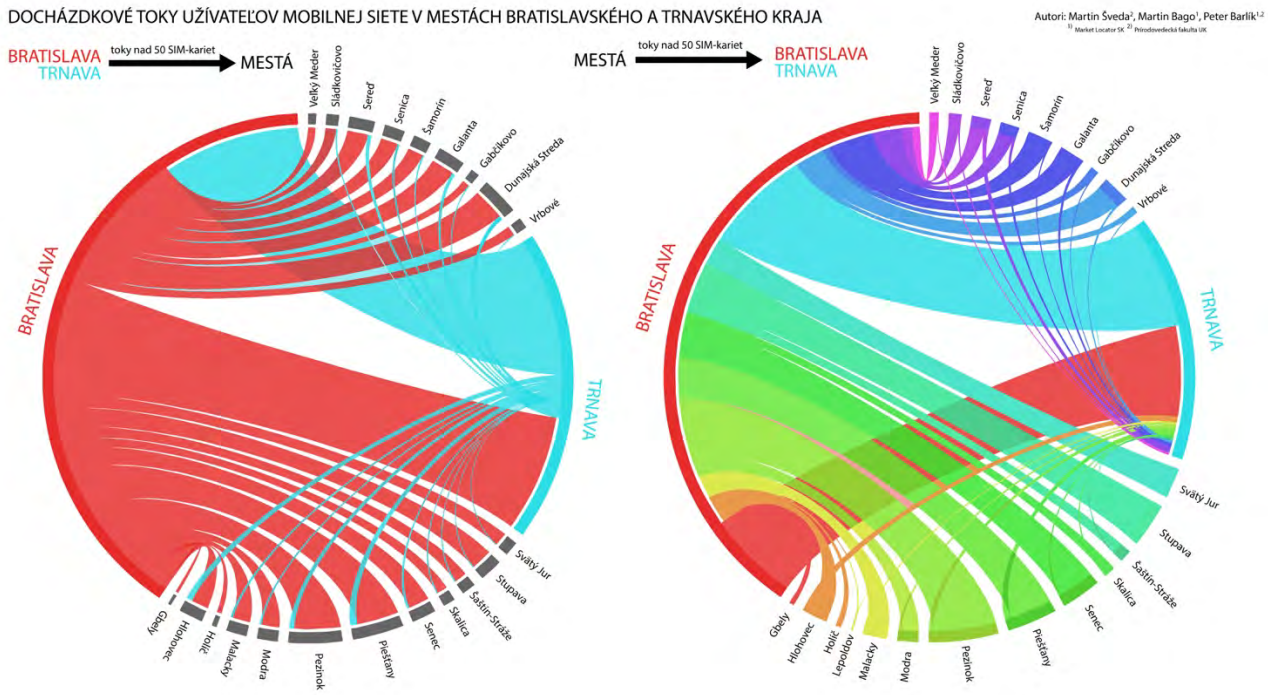
Obr. 6. Veľké dochádzkové toky nad 500 osôb (SIM-kariet) z obcí Bratislavského a Trnavského kraja do Bratislavy a Trnavy. Pasívne lokalizačné údaje mobilnej siete.



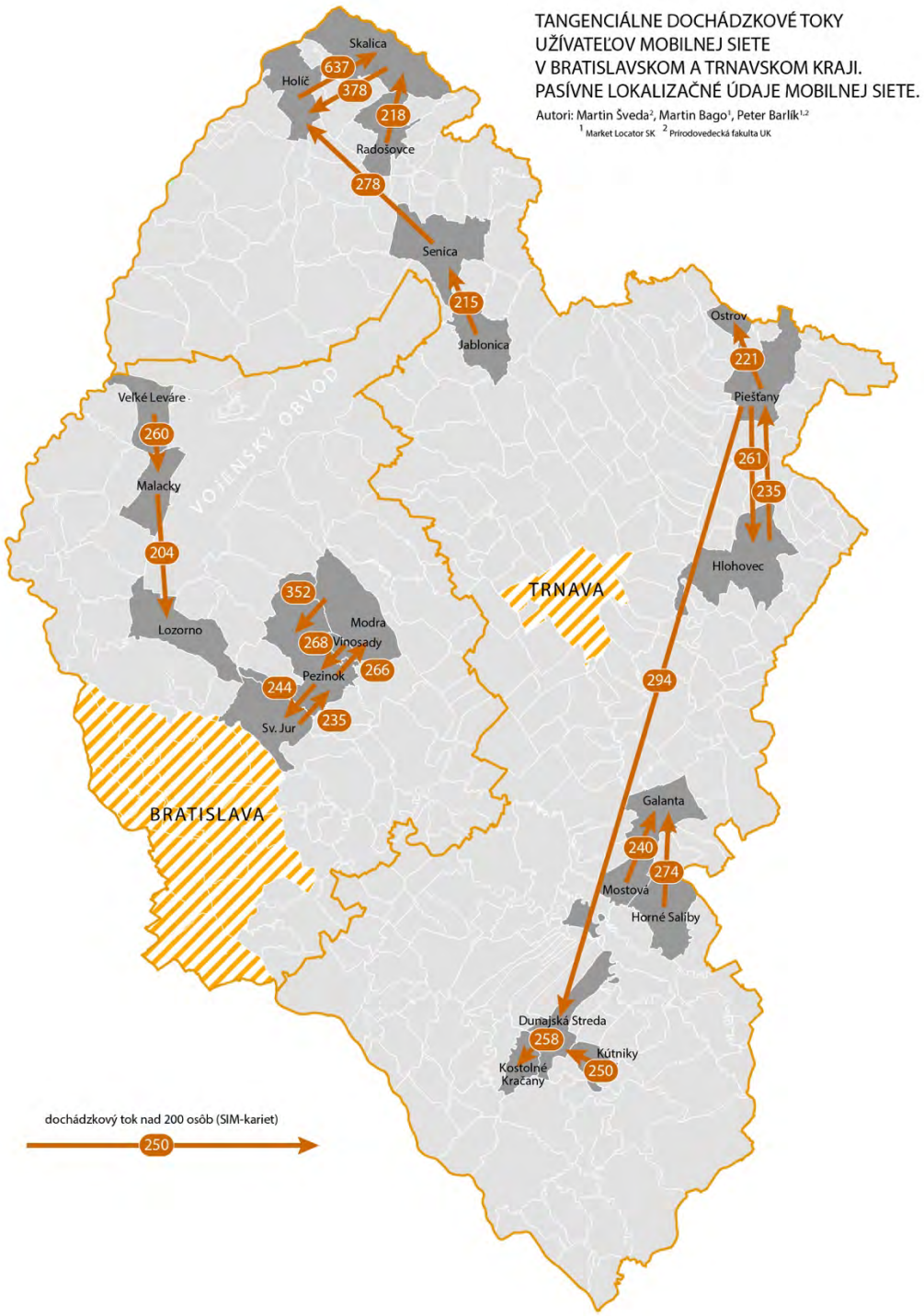
Obr. 7. Stredné dochádzkové toky od 200 do 500 osôb (SIM-kariet) z obcí Bratislavského a Trnavského kraja do Bratislavy a Trnavy. Pasívne lokalizačné údaje mobilnej siete.



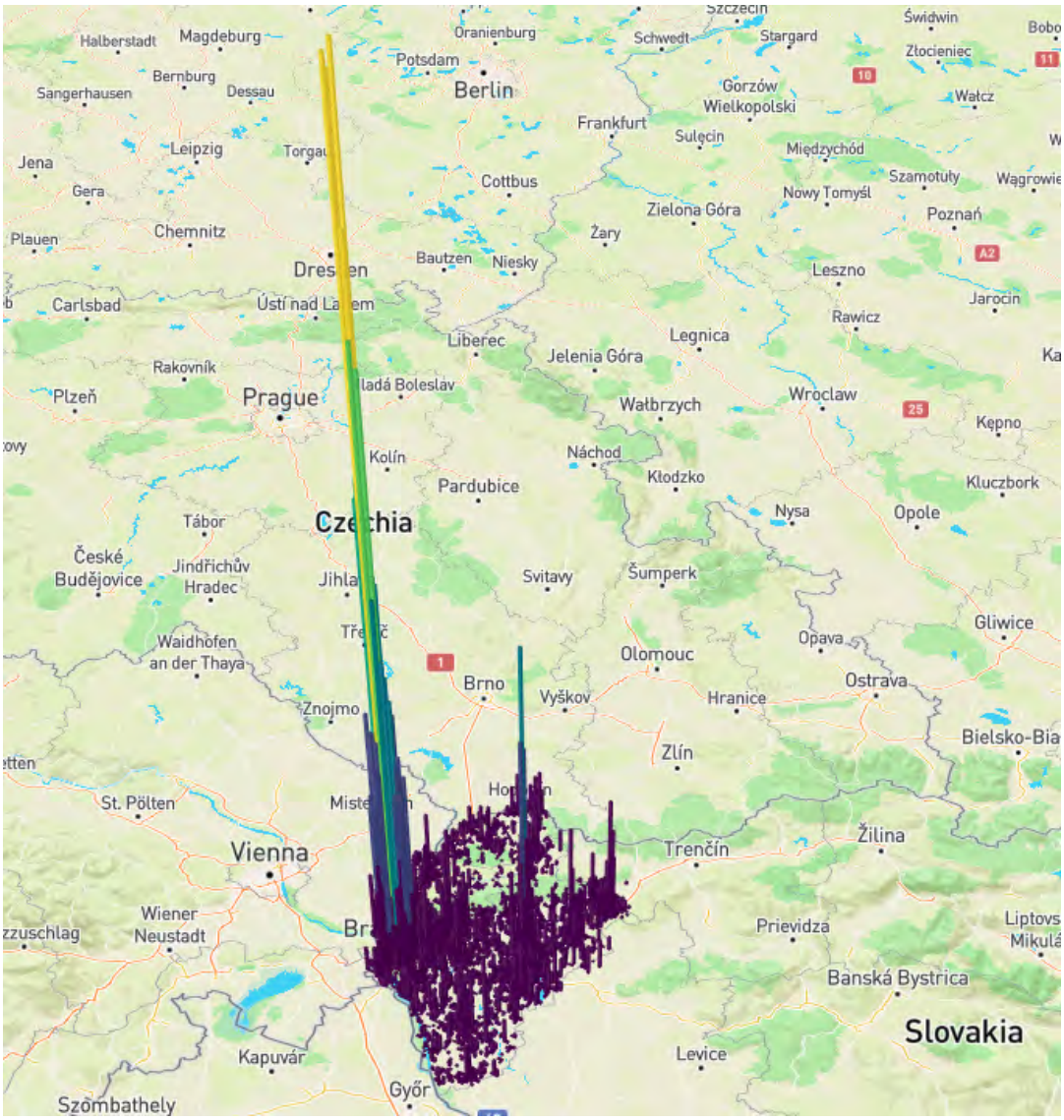
Obr. 8. Dochádzkové toky užívateľov mobilnej siete medzi mestami Bratislavského a Trnavského kraja (okrem tokov do Bratislavy a Trnavy). Znázornené sú len toky väčšie ako 100 osôb (SIM-kariet). Pasívne lokalizačné údaje mobilnej siete.



Obr. 9. Dochádzkové toky užívateľov mobilnej siete medzi Bratislavou (resp. Trnavou) a mestami Bratislavského a Trnavského kraja. Znázornené sú len toky väčšie ako 50 osôb (SIM-kariet). Pasívne lokalizačné údaje mobilnej siete.



Obr. 10. Tangenciálne dochádzkové toky užívateľov mobilnej siete v bratislavskom a Trnavskom kraji. Znázornené sú toky väčšie ako 200 osôb (SIM-kariet) s vylúčením tokov začínajúcich alebo končiacich v Bratislave a Trnave. Pasívne lokalizačné údaje mobilnej siete.



Obr. 11. Vizualizácia počtu SIM kariet pre jednotlivé BTS stanice v Bratislavskom a Trnavskom kraji (Print screen HTML mapy).

Zoznam použitej literatúry

AHAS, R., AASA, A., SILM, S., AUNAP, R., KALLE, H., MARK, Ü. (2007a). Mobile positioning in spacetime behavior studies: social positioning method experiments in Estonia. Cartography and Geographic Information Science, 34, 4, 259-273.

NOVÁK, J. (2010). Lokalizační data mobilních telefonů: možnosti využití v geografickém výzkumu. Dizertačná práca, Praha (Univerzita Karlova v Praze).

AHAS, R., AASA, A., MARK, Ü., PAE, T., KULL, A. (2007b). Seasonal tourism spaces in Estonia: case study with mobile positioning data. Tourism Management, 28, 898-910.

SONG, C., QU, Z., BLUMM, N., BARABÁSI, A. L. (2010). Limits of predictability in human mobility. Science, 327, 5968, 1018-1021.

JÄRV, O., AHAS, R., SALUVEER, E., DERUDDER, B., WITLOX, F. (2012). Mobile phones in a traffic flow: A geographical perspective to evening rush hour traffic analysis using call detail records. PLoS One, 7, 11

RATTI, C., FRENCHMAN, D., PULSELLI, R., WILLIAMS, S. (2006). Mobile landscapes: Using location data from cell phones for urban analysis. Environment and Planning B: Planning and Design, 33, 727-748.

READES, J., CALABRESE, F., RATTI, C. (2009). Eigenplaces: Analysing cities using the space and time structure of the mobile phone network. Environment and Planning B: Planning and Design, 36, 824-836.

BENGTSSON L., XIN LU, THORSON. A., GARFIELD, R., VON SCHREEB, J. (2011). Improved response to disasters and outbreaks by tracking population movements with mobile phone network data: a post-earthquake geospatial study in Haiti, PLoS Medicine, 8, 8.