Zber a spracovanie dát o demografickom správaní sa obyvateľstva v hlavnom meste SR v Bratislave –

(Štúdia uskutočniteľnosti)

Martin Bago[1], Peter Barlík[1], Martin Šveda[2], František Križan[2], Norbert Dvorčák[3], Andrej Petrinec[3], Peter Goga[3], Karin Lexmann[3], Michal Babiar[3], Ján Slimák[4], Petra Dzurovčinová[4]

[1] Market Locator SK s.r.o., 29. augusta 36/A, 811 09 Bratislava

[2] Katedra regionálnej geografie, ochrany a plánovania krajiny, Prírodovedecká fakulta Univerzity Komenského v Bratislave, Ilkovičova 6, 841 04 Bratislava

[3] Oddelenie stratégií rozvoja mesta a tvorby územnoplánovacích dokumentov, Hlavné mesto SR Bratislava, Primaciálne nám. 1, 814 99 Bratislava

[4] Oddelenie stratégie a projektov, Hlavné mesto SR Bratislava, Primaciálne nám. 1, 814 99 Bratislava

Zber a spracovanie dát o demografickom správaní sa obyvateľstva v hlavnom meste SR v Bratislave —

Úvod 2

1.Lokalizačné údaje mobilných telefónov ako nástroj územného plánovania 3

1.1.Trvalý pobyt ako trvalý problém 3

1.2. Mobilná lokalizácia v kontexte populačných analýz 4

1.3. Základné princípy lokalizácie prostredníctvom mobilnej siete 5

1.4. Sieťová lokalizácia 5

1.5. Základné charakteristiky lokalizačných údajov mobilných telefónov v rámci sieťovej lokalizácie 8

1.6. Limity využitia údajov sieťovej lokalizácie 9

1.7. Určenie polohy mobilného zariadenia s využitím sieťovej lokalizácie 10

1.8. Extrahovanie „spacej" a „pracovnej" lokality mobilného zariadenia 12

1.9. Využitie lokalizačných údajov mobilnej siete pre analýzu dennej priestorovej mobility v mestskom regióne 14

1.10. Využitie dát od mobilných operátorov v územnom plánovaní mesta 15

1.11. Zoznam pojmov 17

2. Pilotný projekt 18

2.1. Cieľ pilotného projektu 18

2.2. Definovanie formátu dát 18

2.3. Spracovanie dát 19

2.4. Charakteristika užívateľov mobilnej siete - základná klasifikácia 20

2.5. Dátové výstupy 22

2.6. Vizualizácia dát 22

2.7. Replikovateľnosť štúdie 24

2.8. Prípadová štúdia 24

2.9. Manažment pilotného projektu 27

3. Aplikácia a udržateľnosť projektu 29

3.1. Nadväznosť na strategické ciele mesta 29

3.2. Dlhodobá udržateľnosť projektu 29

3.3. Ekonomická analýza (vrátane dopadu riešenia na mesto) 29

3.4. Alternatívne možnosti financovania riešenia v súčasnosti/budúcnosti 30

3.5. Udržateľnosť a efektívnosť projektu 31

3.6. SWOT analýza 31

4. Záverečné zhodnotenie 33

Prílohy 33

Úvod

Projekt *Zber a spracovanie dát o demografickom správaní sa obyvateľstva v hlavnom meste SR v Bratislave* má za cieľ poskytnúť správe mesta nové pohľady na denne prítomné obyvateľstvo optikou telekomunikačných dát. Lokalizačné údaje mobilných telefónov (SIM-kariet) predstavujú unikátny zdroj údajov, ktoré môžu obohatiť tradičné údajové zdroje o obyvateľstve (napr. Sčítanie obyvateľov, domov a bytov, Register obyvateľov), alebo o dopravných tokoch (napr. celoštátne sčítanie dopravy v SR). Vhodným spracovaním údajov z mobilnej siete môžeme získať pohľad na časovo-priestorové správanie užívateľov mesta, ktoré nám tradičné údajové zdroje nedokážu poskytnúť.

Mobilná komunikácia sa stala bezprostrednou súčasťou nášho života a dramatickým spôsobom ovplyvňuje správanie jednotlivcov v čase a priestore. Mobilný telefón nám neslúži výhradne na komunikáciu, ale čoraz častejšie sa stáva dôležitým informačným kanálom a voľnočasovým nástrojom. Najmä služby priestorovej lokalizácie (location-based services) menia spôsob, akým ľudia využívajú mestský priestor (navigácia, vyhľadávanie služieb). Popri zmenách, ktoré prináša do života jednotlivcov nám mobilná komunikácia poskytuje unikátny a bezprecedentný zdroj údajov o dynamike (nielen) mestského prostredia.

V tejto pilotnej štúdii sa pokúšame charakterizovať základné princípy využitia mobilnej komunikácie v územnom plánovaní, definovať základné metodické rámce a navrhnúť vhodné algoritmy pre optimálne spracovanie lokalizačných údajov mobilných telefónov v urbánnom prostredí. Popri tom diskutujeme aj technické a interpretačné limity tohto, na Slovensku doposiaľ neaplikovaného, prístupu. Prvá časť je venovaná metodologickým rámcom využitia údajov mobilnej lokalizácie, zatiaľ čo v druhej časti predkladáme návrh pilotného projektu.

1. Lokalizačné údaje mobilných telefónov ako nástroj územného plánovania

Kľúčovým parametrom pre formulovanie rozmanitých potrieb v rámci územného plánovania je populačná veľkosť. Populačné prognózy a dlhodobé územné plánovanie obvykle vychádzajú z dát o trvalom pobyte. Na základe počtu trvalo bývajúcich obyvateľov dochádza k prerozdeleniu daňových príjmov, plánovaniu verejnej infraštruktúry, či kreovaniu obecných zastupiteľstiev. Na úvod tejto časti je potrebné uviesť niekoľko myšlienok týkajúcich sa tejto fundamentálnej kategórie.

1.1. Trvalý pobyt ako trvalý problém

Kategória „trvalý pobyt“ sa v súčasných podmienkach rôznorodej mobility obyvateľstva javí čoraz viac ako nedostačujúca. Koncepty trvalého bydliska sa stávajú prekonané tak vo výskumných diskurzoch, ako aj v praktickom použití. Motívov pre spochybnenie relevantnosti tejto fundamentálnej kategórie je viacero. Predovšetkým 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 odlišujú. Dôvodom je predovšetkým rastúca fluktuácia a mobilita obyvateľstva, ako aj fenomény postmodernej spoločnosti, predovšetkým informačno-komunikačné technológie. Pohyby obyvateľstva sa stávajú pestrejšími, nadobúdajú nerutinný a nepravidelný charakter. V súčasnosti nie je neobvyklé rozdielne bydlisko členov jednej rodiny, bývanie v podnájme, sezónne bývanie na chate, či dlhodobé pracovné stáže. Stále menšia viazanosť obyvateľstva k jednému miestu sa prejavuje v rôznych spoločenských kontextoch a relativizuje výsledky sociálnych, geografických, či ekonomických analýz založených na tradičných štatistických zdrojoch (cenzy, registre obyvateľstva a pod.).

Tento trend však netýka len Slovenska, ale predstavuje globálny fenomén, ktorému sa vo svetovej literatúre venuje značná pozornosť najmä v súvislosti s priestorovou mobilitou (Järv et al. 2012 a Csáji et al. 2013), manažmentom krízových udalostí (Bengtsson et al. 2011), epidémií (Tatem a Linard 2011 a Weselowski et al. 2013), či spotrebou energií (Roman a Stokes 2015). 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 (Sheller 2004, Shoval 2007 a Reades et al. 2009). Navyše, infraštruktúra umožňujúca tieto toky je pre výskumníkov fyzicky aj právne „neviditeľná“ (Batty 1990). Naša schopnosť zaznamenať procesy prebiehajúce v spoločnosti je tak čoraz zložitejšou úlohou. Ako si všíma Batty (1990), procesy v spoločnosti sa pre nás stávajú „neviditeľnými“ vo viacerých smeroch a zdá sa, že táto „neviditeľnosť“ rastie rýchlejšie, ako naša schopnosť prispôsobiť naše výskumné metódy novým podmienkam postmodernej spoločnosti. Kauzálny reťazec, ktorý nám po desaťročia pomáhal odkrývať vzťahy medzi miestom práce a bývania, lokalizáciou priemyselných podnikov, či priestorovou distribúciou služieb, všetky tieto

tradičné procedúry, ktoré sme používali pre porozumenie priestorovej organizácie spoločnosti, sa oslabujú. Takýto stav však môžeme podľa Battyho (1990) považovať len za dočasnú stratu „hybnosti" v spoločenskom výskume. 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.

1.2. Mobilná lokalizácia v kontexte populačných analýz

Poznanie veľkosti populácie a jej distribúcie v priestore ostáva stále nedostatočné, a to aj napriek jej nespornému významu pre decíznu sféru a výskum naprieč mnohými vednými oblasťami (Tatem a Linard 2011). Od 90. rokov 20. storočia pozorujeme značné úsilie o spresnenie distribúcie svetovej populácie, ktoré vyústilo do vývoja sofistikovaných prístupov redistribujúcich údaje o populácii z cenzov do pravidelnej mriežky so stranou od 100 metrov do 5 km (Tobler et al. 1997, Balk et al. 2006 a Deichmann et al. 2001). Napriek nemalému pokroku v tejto oblasti sú tieto prístupy poznačené viacerými obmedzeniami, v dôsledku čoho sú výsledky do značnej miery ovplyvnené spoľahlivosťou údajov z populačných cenzov, ktoré vytvárajú základnú bázu pre odhady distribúcie obyvateľstva. Napriek pokusom o sofistikované prístupy o zachytenie dynamicky sa meniacej populácie v priestore (napr. Bhaduri et al. 2007 a Leung et al. 2010) doposiaľ absentoval univerzálny nástroj, ktorý by umožňoval použitie naprieč regiónmi sveta (Deville et al. 2014).

Využitie informačno-komunikačných technológií – najmä mobilnej komunikácie – však vytvára nové príležitosti pre štúdium sociálno-priestorového správania (Watts 2007, Blondel et al. 2016 a Ahas a Mark 2005). Vďaka vysokej penetrácii mobilných telefónov do svetovej populácie 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 (Pucci a Colleoni 2016 a Brockmann et al. 2006) a v neposlednom rade i ochoty obyvateľov poskytovať presné údaje v rámci cenzov. V roku 2015 bolo v krajinách EÚ registrovaných 1 213 predplatiteľov hlasovej alebo dátovej služby mobilnej siete na 1 000 obyvateľov (Eurostat 2017). Tento bezprecedentný rozsah pokrytia populácie vytvára unikátny predpoklad pre holistické prístupy sledovania miest a regiónov, čo je principiálne nemožné (a rovnako aj finančné náročné) s využitím tradičných foriem získavania dát. Všadeprítomnosť a štandardizovaný charakter mobilnej infraštruktúry vytvára dlhodobý rámec na sledovanie priestorovej variability populácie vo vysokom časovom a priestorovom rozlíšení (Ahas and Mark 2005, Ratti et al. 2006 a Reades et al. 2009).

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),

sledovanie návštevnosti zahraničných turistov (Ahas et al. 2008), či šírenie epidémií (Weselowski et al. 2012). 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. Využitie lokalizačných údajov mobilnej siete však presahuje teritorialitu štátnych útvarov. Technológia mobilnej komunikácie je principiálne rovnaká naprieč regiónmi sveta a umožňuje bezprecedentné pozorovanie priestorovej dynamiky obyvateľstva. V prípade metropolitného regiónu Bratislavy má táto skutočnosť obzvlášť veľký význam, keďže umožňuje analyzovať funkčné (dochádzkové) väzby v pohraničnom priestore Rakúska a Maďarska. Takéto cezhraničné sledovania sú len ťažko realizovateľné s využitím tradičných údajových zdrojov.

1.3. Základné princípy lokalizácie prostredníctvom mobilnej siete

Na základnej úrovni môžeme rozlíšiť tri spôsoby mobilnej lokalizácie:

1. Sieťová lokalizácia (*network-based*) je výsledkom základných technických daností prevádzky mobilnej siete a je založená na granularite jej infraštruktúry (veľkosť buniek vyžarovacích polygónov) a potreby zbierania údajov o aktivite mobilného zariadenia. Lokalizácia mobilného telefónu vzniká ako „vedľajší produkt" fungovania mobilnej siete, keďže údaje sú generované primárne pre potreby prevádzky mobilnej komunikácie.
2. Lokalizácia prostredníctvom služieb priestorovej lokalizácie je založená na funkcionalite mobilných aplikácií, ktoré používajú polohu mobilného telefónu (tzv. *assisted GPS, A-GPS*). Okrem zabezpečenia funkcionality danej aplikácie (napr. navigácia) sú tieto údaje využívané aj na marketingové účely (ponuka služieb zohľadňujúca polohu užívateľa).
3. Lokalizačné údaje tretích strán vznikajú využitím záznamov polohy mobilných zariadení, ktoré vznikli v rámci rozličných mobilných aplikácií a sú predmetom (väčšinou nelegálneho) obchodovania.

Predkladaná metodika je založená na využití sieťovej lokalizácie, ktorej princípom sa budeme podrobnejšie venovať v nasledujúcej časti.

1.4. Sieťová lokalizácia

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*). Každá anténa je schopná pokryť určité územie a obslúžiť určitý počet zákazníkov. Identifikačné údaje o práve využívanej anténe a ďalšie doplnkové informácie sú uložené v pamäti telefónu a môžu byť využité pre určenie jeho približnej polohy (Novák

2010). Podobne aj mobilní operátori uchovávajú informáciu na anténe BTS o mobilnom zariadení s aktívnou SIM kartou, ktorú využil počas telekomunikačnej aktivity (prichádzajúci a odchádzajúci hovor, posielanie či doručovanie správ, dátový prenos) alebo počas samotnej prítomnosti pripojenia na danú BTS. Tieto lokalizačné údaje sú prirodzenou súčasťou mobilnej komunikácie, keďže identifikácia základňových staníc (BTS), s ktorými komunikuje mobilný telefón, je nevyhnutná pre samotné fungovanie mobilnej siete.

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), Dufková et al. (2008), 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.

1.5. Základné charakteristiky lokalizačných údajov mobilných telefónov v rámci sieťovej lokalizácie

Lokalizačné údaje mobilných telefónov prinášajú nový zdroj údajov o populácii, ktorý môže poskytovať zmysluplné informácii k už existujúcim zdrojom údajov, akými sú celoplošné cenzy (Sčítanie obyvateľstva) a výberové zisťovania. Na rozdiel od týchto „tradičných" zdrojov však prinášajú viaceré výhody, ktoré charakterizuje veľkosť vzorky, rýchlosť a frekvencia zberu údajov, ako aj finančné náklady potrebné na ich obstaranie.

Veľkosť vzorky spravidla dosahuje stovky tisíc (prípadne aj milióny) užívateľov mobilnej siete a vďaka vysokej penetrácii mobilných telefónov v populácii nadobúda charakter celoplošných cenzov. To nám umožňuje eliminovať náhodné a javy zachytiť kľúčové trendy časovo-priestorového správania obyvateľov. Priestorový rámec výskumu je limitovaný len pokrytím mobilnej siete, ktoré s výnimkou horských oblastí zasahuje prakticky celé územie Slovenska.

Štruktúra vzorky zahŕňa aj užívateľov, ktorí pri použití tradičných prieskumov zostávajú spravidla nezastúpení. Ide najmä o príslušníkov vyšších sociálnych vrstiev, ktorí spravidla na výberové zisťovania nereflektujú.

Rýchlosť a frekvencia zisťovania vychádza zo základných princípov sieťovej lokalizácie v mobilnej sieti. Pasívne lokalizačné údaje sú generované v každom okamihu a umožňujú bezprostredné využitie. Získavame tak informácie o obyvateľstve prakticky v reálnom čase. To predstavuje zásadnú výhodu oproti tradičným zdrojom údajov, ktorých aktuálnosť je limitovaná potrebou dlhšieho času na spracovanie (ako napr. pri národných cenzoch). Zisťovanie prostredníctvom údajov z mobilnej siete môže prebiehať akokoľvek dlho a s ľubovoľnou frekvenciou. To je dôležité najmä pri snahe o zachytenie dynamických procesov v spoločnosti, pri ktorých je ročná či desaťročná frekvencia zisťovania nedostatočná.

Porovnateľnosť naprieč rôznymi krajinami je veľmi jednoduchá, keďže mobilná komunikácia využíva podobnú infraštruktúru a údaje, ktoré sa v nej generujú majú rovnakú štruktúru. To nám umožňuje aplikovať rovnaké postupy v rôznych prostrediach (krajinách, mestách) a dospieť tak k relevantným výsledkom, ktoré nie sú limitované rozdielnou metodikou zberu a spracovania.

Spoľahlivosť pasívnych lokalizačných údajov mobilných telefónov je do značnej miery závislá od schopnosti tieto údaje zmysluplne spracovať a interpretovať. Na rozdiel od tradičných foriem zisťovania, ktoré sú závislé od ochoty a schopnosti respondentov poskytnúť presné a pravdivé informácie, údaje z mobilnej siete sú objektívnym záznamom o polohe a aktivite užívateľa mobilného zariadenia. Príkladom môže byť otázka k frekvencii a dĺžke pobytu respondenta v zahraničí, na ktorú nám lokalizačné údaje mobilných telefónov dávajú veľmi presnú odpoveď. Hoci aj tu existujú oprávnené výhrady. Predovšetkým SIM-kartu nemožno výlučne stotožniť s jej držiteľom. Príkladom sú SIM-karty detí registrované na rodičov. Táto skutočnosť však nevytvára zásadnú prekážku zmysluplnému spracovaniu lokalizačných údajov. SIM-karty, ktoré využívajú detskí užívatelia sa dajú identifikovať na základe odlišného časovo-priestorového správania či spôsobu využívania mobilného telefónu. Je však možné pracovať aj s menšími agregátmi, ktoré

môžu reprezentovať jednotlivé domácnosti (rodiny). Pri použití vhodného spracovania (filtrovania) údajov a ich korektnej interpretácii tak dostávame veľmi spoľahlivý zdroj informácií o priestorovej dynamike obyvateľstva.

Priestorová presnosť je závislá od technických parametrov mobilnej infraštruktúry, predovšetkým od granularity buniek mobilnej siete a ich technológie (2G/3G/4G).

Finančné náklady spojené s obstaraním a spracovaním lokalizačných údajov mobilných telefónov sú neporovnateľne nižšie, než sú bežné výdavky spojené s prípravou, priebehom a realizáciou celoplošných zisťovaní (v prípade Sčítania obyvateľov, domov a bytov v roku 2011 to bolo približne 22 mil. Euro, ŠÚ SR 2009, 2010, 2011, 2012).

1.6. Limity využitia údajov sieťovej lokalizácie

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. Duplicitu SIM-kariet u jedného užívateľa je možné odstrániť, ide však o náročné porovnávanie, kde je cieľom identifikovať SIM karty, ktoré podstúpili rovnaké zmeny pripojenia v rovnakom čase. Pri časových agregátoch je však tento postup ešte náročnejší, keďže nepoznáme presný čas zmeny BTS bunky, ale len agregát pre danú hodinu (napr. najčastejšiu bunku v danom hodinovom reze). Ak má užívateľ viaceré SIM-karty u viacerých operátorov, je identifikovanie duplicít znemožnené rozdielnosťou siete - iné polohy BTS buniek, iné vyžarovacie polygóny, anonymizácia MSISDN ID. Častočnú elimináciu duplicitných SIM-kariet však môžeme dosiahnúť vylúčením neaktívnych, nehlasových a neprimárnych 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. Uvedené limity je však možné čiastočne eliminovať použitím vhodnej dátovej analýzy, ktorá dokáže identifikovať jednotlivých členov domácnosti na základe špecifického používania mobilného telefónu (napr. používanie aplikácií a hier určených pre deti).
- 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. Súčasný vývoj v analýze mobilných dát však prináša nástroje umožňujúce zásadné spresnenie polohy v rámci sieťovej lokalizácie. Príkladom je využitie tzv. reťazenia, kedy sa sleduje následnosť

lokalizácie v jednotlivých bunkách mobilnej siete. Pre spresnenie polohy je možné využiť aj sídelnú a cestnú sieť, či iné pokročilé metódy areálovej transformácie údajov (napr. dazymetrická priestorová interpolácia). V praxi najmä ide o transformáciu údajov z buniek mobilnej siete do areálov sídelnej zástavby, čím dostávame reálnejší obraz o priestorovej distribúcii obyvateľstva.

- Jednotliví operátori mobilnej siete môžu byť zacielení na špecifickú skupinu užívateľov (napr. na mladších užívateľov, podnikateľov, seniorov pod.). Aproximácia údajov na celú populáciu je tak značne obmedzená. Pri spracovaní údajov od viacerých operátorov, ktorí spoločne pokrývajú väčšinu populácie, sa však tento problém do značnej miery eliminuje. Napriek tomu však niektoré vekové kategórie ostávajú poddimenzované (najmä malé deti). S využitím štruktúrnych charakteristík populácie (napr. podiel detí vo veku 0-9 rokov) však môžeme dospieť k presnejšej aproximácii populácie (prítomného / bývajúceho obyvateľstva).

Ďalšie limity využitia údajov z mobilnej siete vychádzajú z princípov určenia polohy SIM-karty v mobilnej sieti, v rámci ktorého dochádza k viacerým nevyhnutným skresleniam.

1.7. Určenie polohy mobilného zariadenia s využitím sieťovej lokalizácie

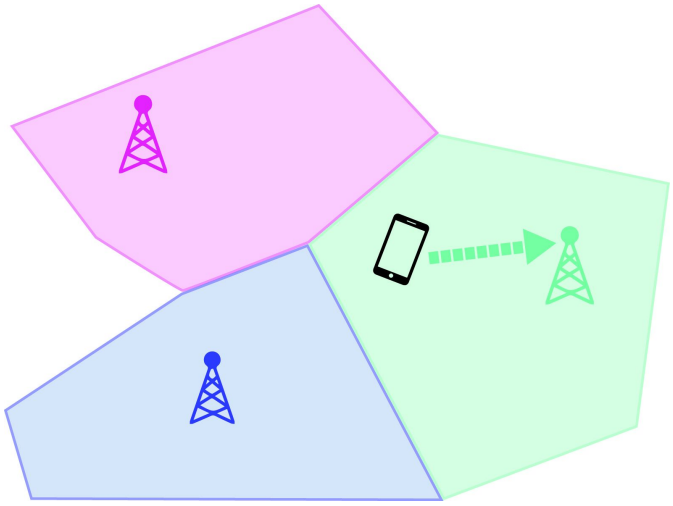
Určenie polohy mobilného zariadenia v sieťovej lokalizácii prebieha nasledovne. Zariadenie s aktívnou SIM kartou sa pripája na BTS stanicu (Obr. 1). Spravidla ide o najbližší vysielač, nemusí to však byť pravidlom. Niektoré BTS stanice totiž slúžia na prenos dát (napr. 4G bunky), iné na prenos hovorov (2G/3G). Zároveň sieť autonómne distribuuje signál pripojeným zariadeniam podľa svojho vyťaženia. Ak je teda niektorá bunka preťažená, pošle časť zariadení na inú bunku. Práve organické zmeny buniek mobilnej siete (tzv. *handover*) môžu vytvárať zdanlivý pohyb SIM-karty, hoci jej poloha je statická. Pri časových agregátoch tento problém zaniká, keďže sa pre danú časovú jednotku (napr. hodinu) vyberá najčastejšia bunka, resp. bunka s najviac záznamami (viď nižšie).

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). Keďže výpočty s nepravidelnými polygónmi sú pomerne náročné, polygóny podliehajú vyhladzovaniu a matematickému zjednodušovaniu, aby s nimi bolo možné jednoduchšie pracovať. Tu prichádza k ďalšiemu skresleniu.

Poloha mobilného zariadenia je limitovaná veľkosťou vyžarovacieho polygónu. Vieme teda, že ak je daná SIM karta pripojená ku konkrétnej BTS bunke, nachádza sa v danom vyžarovacom polygóne. Väčšiu presnosť však nie je možné bez dodatočných informácií (napr. GPS poloha zariadenia) dosiahnuť. Presnosť priestorovej lokalizácie je tak limitovaná veľkosťou vyžarovacích polygónov.

Ako bolo spomínané vyššie, jednotlivé bunky majú rôzny dosah v závislosti od ponúkanej technológie. Polygóny 2G buniek majú dosah v kilometroch, polygóny 4G buniek už len rádovo stovky metrov. Pri snahe o dosiahnutie vyššej presnosti priestorovej lokalizácie je možné z datasetu vylúčiť práve 2G bunky, ktoré zvyčajne slúžia iba na prenos hovorov. Takýto prístup je vhodný napríklad v urbanizovaných územiach s vysokým počtom BTS buniek. Pri analýze vidieckych oblastí by však vyradenie 2G buniek dramaticky znížilo celkovú vzorku, keďže vidiecka a prírodná krajina je pokrytá oveľa menším počtom buniek. Riešením je rozdelenie územia do pravidelnej siete, kde je pre každú jednotku definovaný podiel pokrytia jednotlivých typov BTS staníc. Na jeho základe sa následne alokujú údaje o užívateľoch mobilnej siete do uvedenej pravidelnej siete. Takýto postup prináša jednoduchý priestorový systém záznamov, ktoré môžeme agregovať do ľubovoľných väčších priestorových jednotiek. Popri skladobnosti je výhodou aj prispôsobenie týchto záznamov iným zdrojom priestorových údajov (napr. údaje zo Sčítania obyvateľstva agregované do priestorovej mriežky s veľkosťou 1 x 1 km).

S príchodom technológie 5G siete bude možné dosiahnuť ešte väčšiu presnosť, keďže dosah týchto buniek bude iba v desiatkách, maximálne stovkách metrov. 5G sieť teda prinesie oveľa hustejšie pokrytie zastavaných území BTS bunkami, pričom vo vysoko frekventovaných oblastiach (napr. v nákupných centrách) bude musieť byť ich hustota dostatočne vysoká na to, aby 5G bunky dokázali doručovať kvalitný dátový prenos množstvu pripojených zariadení. V Európe 5G sieť aktuálne používa napríklad San Marino.



Obr. 1. Princíp lokalizácie telefónu v mobilnej sieti s využitím sieťovej lokalizácie.

Lokalizácia mobilného zariadenia na úrovni buniek mobilnej siete nám umožňuje analyzovať priestorovú distribúcia obyvateľstva v relatívne detailnej mierke a v každom okamihu. To prináša bezprecedentný zdroj údajov pre analýzu dennej dynamiky mesta. Môžeme tak jednoducho sledovať nielen početnosť, ale aj štruktúru prítomného obyvateľstva na úrovni jednotlivých lokalít (napr. urbanistických obvodov) v ľubovoľných časových rezoch (napr. každú hodinu). To má zásadný význam pre lepšie porozumenie denných rytmov jednotlivých lokalít a ich funkcii v rámci mnohovrstevného mestského organizmu. Bezprostredné využitie nachádzajú takéto

údaje aj v krízovom manažmente, v ktorom pomáhajú distribuovať záchranné zložky a kapacity pre evakuáciu obyvateľstva.

Poznanie distribúcie obyvateľstva v priestore a čase je základnou schopnosťou lokalizačných údajov sieťovej lokalizácie. Ďalším krokom spracovania týchto údajov je extrahovanie „spacej" a „pracovnej" lokality užívateľov mobilnej siete a následná konštrukcia dochádzkových tokov.

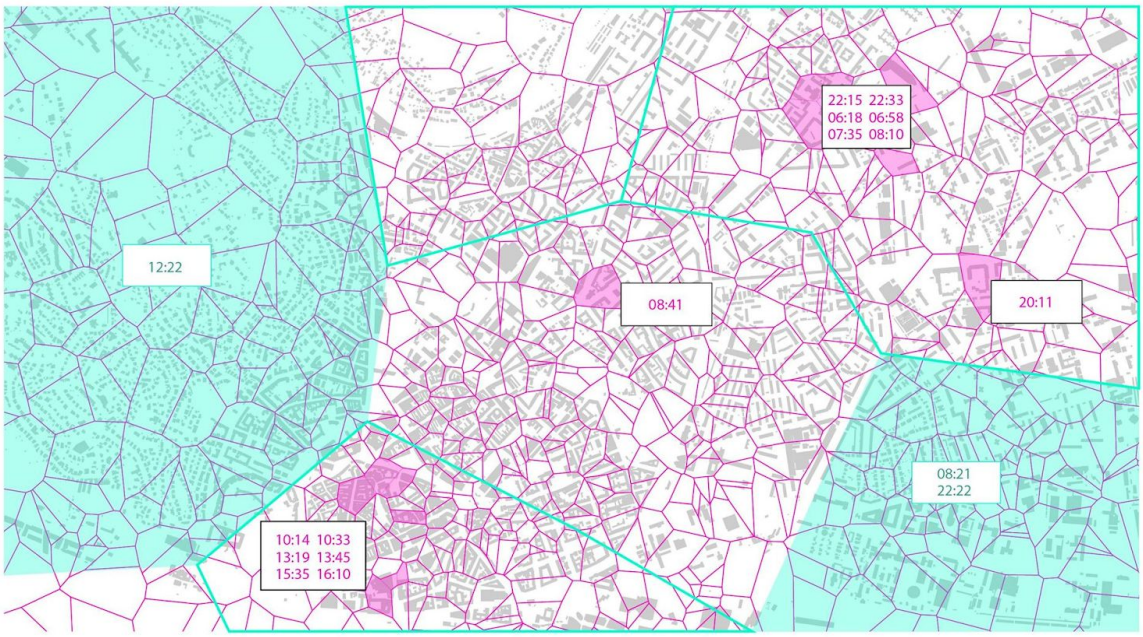
1.8. Extrahovanie „spacej" a „pracovnej" lokality mobilného zariadenia

Základným krokom pre zmysluplné využitie mobilnej lokalizácie v urbánnom plánovaní je určenie „pracovnej" a „spacej" lokality užívateľov mobilnej siete. Metodický a konceptuálny rámec vychádza z konceptu tzv. kotevných bodov – anchor-points (Nurmi a Koolwaaij 2006 a Ahas et al. 2009). Kotevné body sú chápané ako hlavné uzly aktivít človeka, ktoré vytvárajú kostru jeho každodenných pohybov. Použitím vhodného algoritmu môžeme z lokalizačných dát mobilnej siete vyťažiť informácie o základných kotevných bodoch „domov" a „práca". V praxi ide predovšetkým o identifikovanie miest dennej a nočnej lokalizácie. Údaje o koncentrácii denných a nočných kotevných bodov nám umožňujú nielen spresniť priestorovú distribúciu obyvateľstva (napr. koľko ľudí býva v časti mesta), ale aj extrahovať údaje o predpokladaných denných dopravných prúdoch (napr. koľko ľudí dochádza denne do danej časti mesta).
V praxi existuje celá rada techník (pozri napr. Yooet al. 2005, Nurmi and Koolwaaij 2006, Ahas et al. 2009, Huang et al. 2010) pre extrahovanie zmysluplných denných (pracovných) a nočných (domácich) lokalít z množstva polohových záznamov, ktoré vznikajú v mobilnej sieti. V slovenskom prostredí bol úspešne otestovaný algoritmus pre určenie pravdepodobnej spacej a pracovnej lokality užívateľa mobilného zariadenia (Šveda a Barlík 2018), ktorého základom je sledovanie počtu CDR záznamov (aktivít telefónu) v bunkách mobilnej siete v špecifických častiach dňa. Pre pracovnú polohu je to interval PO-PI od 9:00 do 15:00. Pre spaciu polohu SIM-karty je to interval PO-PI od 20:00 do 05:00. Presnejšie je vzťah daný nasledovne:

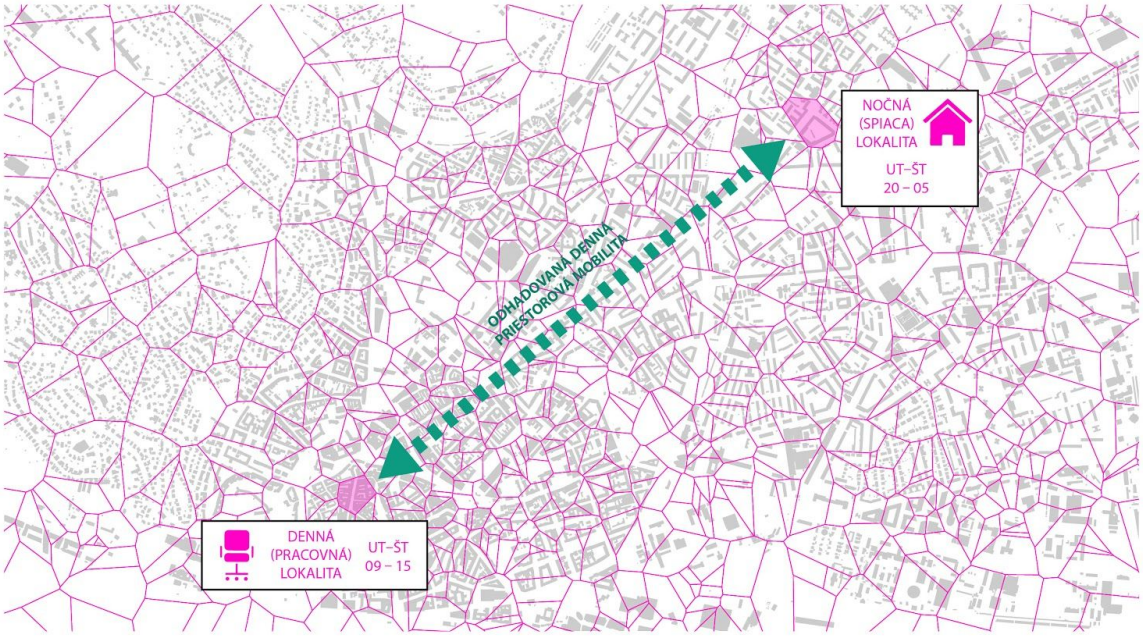
$$W_{is}^{d} = \frac{A_{is}^{d}}{\sum_{j}^{t} A_{is}^{d}}$$

kde W je podiel aktivít mobilného telefónu Ais na celkovom počte všetkých aktivít (správ, hovorov), ktoré sa vyskytli počas dňa d na konkrétnej BTS stanici s pre individuálneho užívateľa (SIM-karty) i. Celkový počet BTS staníc (buniek) v sledovanom území je t.
Pre eliminovanie náhodných a nerutinných lokalizácií je potrebné analyzovať dlhší časový priebeh. Pre vyprofilovanie rutinných denných a nočných lokalizácií sa ako optimálny javí interval 14 dní. Každý monitorovací deň sú tri najaktívnejšie BTS bunky porovnávané s troma najaktívnejšími bunkami z predchádzajúceho dňa, ktorých váha je redukovaná koeficientom $k = 0.9^{(D-d)}$, kde D je pozorovaní (dní) a d je poradie pozorovania (dňa). Výsledná poloha pre spaciu a nočnú lokalizáciu je určená ako BTS bunka s najväčším podielom aktivít:

$$\max W_{is}^{d} = \left\{W_i^1 k^{(D-1)}, W_i^2 k^{(D-2)}, \dots, W_i^d k^{(D-d)}\right\}$$



Obr. 2. Bunky mobilnej siete so zaznamenanou aktivitou jedného užívateľa počas 24 hodín. Príklad z vysoko urbanizovaného prostredia s vysokou granularitou buniek mobilnej siete (polygóny buniek mobilnej siete sú ilustračné).

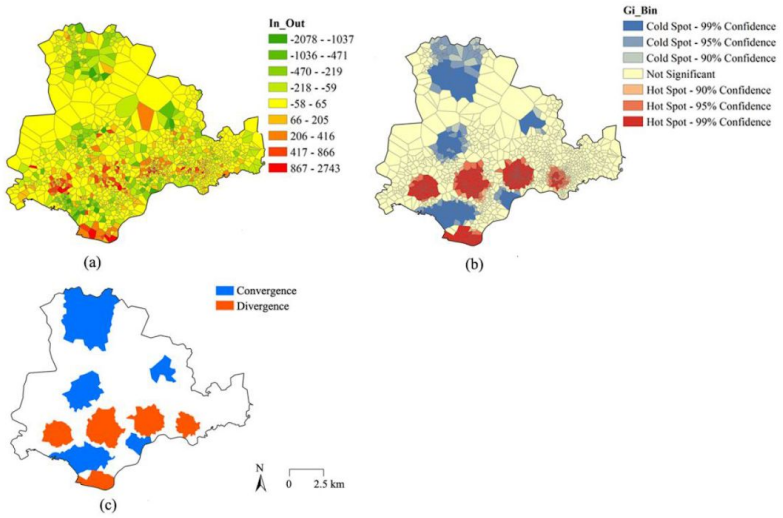


Obr. 3. Identifikácia dennej (pracovnej) a nočnej (spacej) lokality užívateľa mobilného telefónu (polygóny buniek mobilnej siete sú ilustračné).

1.9. Využitie lokalizačných údajov mobilnej siete pre analýzu dennej priestorovej mobility v mestskom regióne

Po extrahovaní pracovnej a spacej lokality mobilného telefónu je konštrukcia dochádzkových tokov jednoduchým krokom. Treba však upozorniť, že v tomto prípade ide len o geometrickú konštrukciu očakávaných tokov osôb z ich bydliska do ich pracoviska. Hoci presnú trajektóriu pohybu osôb nepoznáme, údaje o odhadovanej veľkosti dochádzkových tokov sú kľúčovým ukazovateľom vstupujúcim do územného plánovania a sú tiež dôležité pre zvládanie čoraz závažnejších dopravných problémov a ich vplyvu na kvalitu života v mestách. Na rozdiel od tradičných prístupov, ktoré analyzujú dennú priestorovú mobilitu na relatívnej malej vzorke respondentov, údaje z mobilnej siete umožňujú zachytiť veľkú časť mestskej populácie (najmä tú ekonomicky aktívnu) a sledovať aj jej štruktúrne charakteristiky.

Poznanie denných a nočných koncentrácií obyvateľstva je dôležitou súčasťou výskumu dennej dynamiky mesta. V tomto kontexte môžeme uvažovať o konvergentných a divergentných lokalitách. Konvergentné lokality predstavujú miesta, kde je počet dochádzajúcich väčší ako odchádzajúcich. Naopak, divergencia z lokality naznačuje, že počet ľudí opúšťajúcich lokalitu je väčší ako počet prichádzajúcich ľudí. V porovnaní s jednoduchými údajmi o počte užívateľov v danej bunke, bilancia odchodu a príchodu SIM-kariet z lokality (bunky) indikuje zmenu v koncentrácii obyvateľstva v danom časovom intervale (napr. v jednotlivých hodinách počas dňa). Na základe rôznej intenzity príchodov a odchodov v rôznych častiach dňa môžeme vytvoriť klasifikáciu, ktorá zachytáva základné vzory časového priebehu bilancie konvergencie a diveregencie. Napríklad rezidenčné a administratívne lokality budú mať odlišný časový priebeh zmien v koncentrácii obyvateľstva, rovnako sa môžu odlišovať aj „robotnícke" sídliská od rezidenčných areálov s vyšším sociálnym statusom.



Obr. 4. Identifikácia konvergentných a divergentných areálov v mestskom regióne (Shenzhen, Čína) s využitím „hot spot" analýzy (v tomto prípade Getis–Ord Gi index). Zdroj: Yang et al. 2018.*

Uvedené využitie lokalizačných údajov mobilných telefónov je len elementárnym spracovaním vstupných údajov. Ďalšie spracovanie (filtrovanie) údajov z mobilnej siete a jej zmysluplné prepojenie s rôznymi dátovými vrstvami (cestná infraštruktúra, demografické štruktúry, využitie zeme a pod.) prinášajú unkátny nástroj pre rozmanité potreby v rámci územného plánovania, či verejnej správy. Zahraničné štúdie nám ponúkajú nespočetne veľa príkladov využitia údajov z lokalizácie mobilných telefónov:

- dopravné plánovanie (Lai et al. 2018)
- návštevnosť zahraničných turistov a ich časovo-priestorové správanie (Ahas et al. 2008)
- návštevnosť rekreačných areálov (Zhai et al. 2018)
- krízový manažment (Kubíček et al. 2018)
- časovo-priestorový pohyb rôznych vekových skupín obyvateľstva s dôrazom na špecifické potreby starších obyvateľov (Masso et al. 2010)
- optimalizácia verejnej dopravy (lokalizácia zastávok, trasovanie a kapacita liniek a pod.)

1.10. Využitie dát od mobilných operátorov v územnom plánovaní mesta

Pre strategický územný rozvoj každého mesta je dôležité územné plánovanie, ktorého základnou súčasťou je poznanie demografického správania obyvateľstva a následná prognóza demografického vývoja. Každý územný alebo priestorový plán potrebuje kvalitné socioekonomické a demografické informácie, ak chce vyvážene zabezpečovať služby obyvateľstvu mesta a zachovať jeho proporcionálny rozvoj. Takýmito informáciami môžu byť rôzne dátové zdroje o obyvateľoch mesta vrátane dát od mobilných operátorov.

Socioekonómia a využívanie jej dát v územnom plánovaní sa prejavuje najmä vo vzájomných väzbách medzi

- demografickou štruktúrou
- bilanciou pracovných síl
- jednotlivými zložkami osídlenia
- potrebou základných druhov občianskej vybavenosti
- požiadavkami na technickú vybavenosť
- jednotlivými systémami dopravy
- prírodnými zložkami

Socioekonómia z hľadiska využívania dát ovplyvňuje celú štruktúru jednotlivých zložiek osídlenia, ktoré sú základnými činnosťami v rámci územného rozvoja. Dominantnú pozíciu v rámci socioekonómie zohráva človek a jeho základné potreby pre udržanie života. Bez poznania základných potrieb človeka nie je možné urobiť dobrý územný plán. Územné plánovanie vo vymedzenom priestore mesta a jeho regiónu je teda dôležitým nástrojom pre pochopenie organizácie územného priestoru pre všetky činnosti človeka, ktorého rozhodujúcim nástrojom sú socioekonomické údaje a zber rôznych dát o obyvateľstve.

Základom pre správne analyzovanie a pochopenie priestorového správania sa obyvateľstva v meste sú údaje od rôznych inštitúcií zaoberajúcich sa sledovaním obyvateľstva. Za základné sa považujú výsledky z národných cenzov za príslušné dekády, informácie z registra obyvateľov (REGOB MV SR), informácie z rôznych iných databáz príslušných správcov a prevádzkovateľov zariadení, vrátane dát od mobilných operátorov.

Lokalizačné údaje mobilných telefónov môžu byť v rámci územného plánovania využité vo viacerých oblastiach:

1. Pri sledovaní priestorovej distribúcie obyvateľstva a jeho dennej priestorovej mobility, predovšetkým dochádzky do práce. Zmysluplné využitie však prináša aj analýza širších vzorcov časovo-priestorového správania (nákupné správanie, víkendová mobilita a pod.). Popri sledovaní smeru a veľkosti jednotlivých tokov, nám údaje z mobilnej siete umožňujú analyzovať aj štruktúrne charakteristiky, pri ktorých má osobitné postavenie vek, ekonomická aktivita a nočná lokalizácia (bydlisko) užívateľov mobilnej siete.
2. V oblasti pracovných príležitostí a trhu práce najmä vo využívaní pohybu obyvateľstva za prácou do príslušných centier zamestnanosti.
3. V oblasti bývania najmä vo využívaní lokalít bývania s trvalým a prechodným pobytom počas dňa a noci.
4. V oblasti obchodu z hľadiska návštevnosti obchodných centier, ich spádovej oblasti a profilu návštevníkov.
5. V oblasti kultúry z hľadiska návštevnosti kultúrnych zariadení najmä počas kultúrnych podujatí i mimo nich.
6. V oblasti športových zariadení a využívania areálov športu organizovanej a neorganizovanej telovýchovy najmä z hľadiska využívania týchto priestorov a zariadení, návštevnosti organizovaných športových podujatí a pod.
7. V oblasti cestovného ruchu sledovaním počtu návštevníkov a ich časovej a priestorovej distribúcie na území mesta v pracovných a víkendových dňoch.
8. V oblasti zdravotníctva najmä v sledovaní využívania zdravotníckych zariadení (predovšetkým nemocníc) a sledovanie ich spádovej oblasti.
9. V oblasti školstva sledovaním časovo-priestorových vzorcov správania študentov v rámci školskej a mestskej infraštruktúry.
10. V oblasti dopravy v zdrojoch a cieľoch dopravy, v zaťaženiach dopravných komunikácií, križovatiek a výpadoviek z mesta a do mesta, v zariadeniach dopravy najmä jednotlivých staniciach osobnej dopravy či už cestnej, železničnej a vodnej dopravy, pohybu obyvateľov v MHD, pohyb obyvateľov v priebehu dňa, pracovných a víkendový dní, v jednotlivých dopravných trasách a komunikáciách a pod.
11. V oblasti poskytovania služieb a využívania zariadení poskytujúcich služby na území mesta pre obyvateľov.

Priestorový systém agregácie lokalizačných údajov mobilných telefónov by mal rešpektovať prirodzenú priestorovú štruktúru mesta, resp. príslušné základné sídelné jednotky - urbanistické obvody, ako aj ich skladobnosť v rámci administratívneho členenia SR.

Sledovanie dát by malo vychádzať zo sledovania počas pracovných dní v rozsahu dňa a noci, v príslušných hodinových intervaloch a počas víkendových dní. Analýza dát o pohybujúcom sa obyvateľstve by mala vychádzať z týchto hlavných výstupov:

- z analýzy prítomného obyvateľstva (užívateľov mobilnej siete) v územných jednotkách mesta a v časových rezoch (hod.) v členení rezidenti, pracujúci, tranzitujúci a ostatní návštevníci,
- z analýzy denných rytmov jednotlivých územných jednotiek mesta z hľadiska ich dominantnej funkcie (rezidenčná, výrobná, administratívna, rekreačná a pod.),
- z analýzy dochádzky za prácou medzi jednotlivými územnými jednotkami mesta,
- z analýzy dochádzkových tokov medzi mestom a jednotlivými obcami v zázemí Bratislavy, najmä regiónu denných dochádzkových tokov (tzv. funkčný mestský región),
- dochádzky do centier príslušnej mestskej vybavenosti podľa jednotlivých funkcií školstvo, zdravotníctvo, sociálna vybavenosť, kultúra, šport a rekreácia, služby a centrá cestovného ruchu,
- zaťaženia dopravnej siete mesta a jej jednotlivých nosných smerov aj v rámci MHD.

1.11. Zoznam pojmov

BTS – Base Transceiver Station - je jedna bunka siete, jeden “vysielač” signálu.
2G (GSM) – táto technológia priniesla mobilným telefónom po prvýkrát možnosť odosielať SMSky, dodnes slúži hlavne na prenos hlasových hovorov a SMS. Vo vidieckych oblastiach môžu 2G bunky vysielať signál až do vzdialenosti väčšej ako 6 kilometrov. V mestách vysielajú signál spravidla do 1 kilometra.
3G (UMTS) – prvé 3G siete začali vo svete fungovať už v roku 1998, k nám sa dostali o niečo neskôr. 3G technológia prenosu informácií po prvýkrát umožnila bezdrôtovo prenášať aj dáta, a to až do rýchlosti 384 kbit/s. Neskôr dorazili protokoly HSPA a HSPA+, pomocou ktorých bolo možné dáta prenášať až rýchlosťou 56 Mbit/s. 3G bunky na vidieku vysielajú do vzdialenosti niekoľko kilometrov, v mestách iba niekoľko sto metrov.
4G (LTE) – štvrtú generáciu na Slovensku poznáme aj pod označením LTE a jej teoretické prenosové maximum dosahuje až 1 Gbit/s. 4G bunky v mestách majú zásah niekoľko desiatok metrov.
CDR - Call detail record - pod skratkou CDR rozumieme záznamy, ktoré vznikajú pri interakcii užívateľa so zariadením, ktoré používa SIM kartu. Za interakciu v tomto zmysle slova rozumieme akýkoľvek hovor, SMS správu či dátový prenos. CDR majú jasne definovanú štruktúru, ktorá obsahuje typ interakcie, zainteresované čísla, či dĺžku hovoru.
MSISDN - skratka pre telefónne číslo asociované so SIM kartou. V našej práci neprichádzame do styku s MSISDN, ale len s kryptovanými unikátnymi identifikátormi SIM kariet.
hash - zhluk znakov, ktorý spravidla maskuje pôvodný údaj neznámym kľúčom, čím plne anonymizuje pôvodnú vzorku.
Vyžarovací polygón - mnohostranný plošný útvar, ktorý definuje zónu pokrytú signálom práve jednej BTS bunky. Vyžarovací polygón buď poskytuje operátor z dominancie siete a následne takto komplikovaný útvar zjednodušujeme alebo je dopočítaný na základe technológie danej bunky, uhla vyžarovania, vzdialenosti a polohy bunky.

2. Pilotný projekt

2.1. Cieľ pilotného projektu

Cieľom pilotného projektu je zhodnotiť možnosti lokalizačných údajov mobilných telefónov pre potreby územného plánovania a poskytnúť zmysluplnú aproximáciu priestorovej distribúcie obyvateľstva a jeho časovo-priestorového správania v mestskom regióne Bratislavy. Zámerom je nielen spresniť údaje o priestorovej distribúcii obyvateľstva, ale aj otestovať zmysluplnosť mobilnej lokalizácie ako doplnkového zdroja údajov pre priestorové a populačné analýzy. Zároveň tak chceme diskutovať možné interpretačné mantinely tohto netradičného zdroja údajov, ktorého význam v priestorových analýzach bude pravdepodobne narastať.

2.2. Definovanie formátu dát

Telekomunikačných operátorov budeme na základe definovaných požiadaviek dopytovať o datasety signalizačných dát slovenských hlasových SIM kariet v požadovanej časovej a geografickej vzorke. Pre potreby tejto štúdie si vystačíme s hodinovými agregátmi – a teda nepotrebujeme všetky signalizačné záznamy danej vzorky SIM kariet užívateľov, ale iba ich agregáty v 60 minútových intervaloch. Ako dátový zdroj preferujeme záznamy o aktivite domácich SIM kariet, zahraničných roamingujúcich SIM kariet ako aj národného roamingu. Podľa potrieb si následne dáta vyčistíme.

Dátovú vzorku si na základe našich predchádzajúcich skúseností definujeme takto:

SUBSCRIBER_ID	AGE_GROUP	CONTRACT_TYPE	DAY_ID	HOUR_ID	CELL_ID

Kde:

SUBSCRIBER_ID je unikátny anonymizovaný identifikátor konkrétneho používateľa. Spravidla ide o zhluk rôznych znakov, ktoré maskujú pôvodné telefónne číslo. Pri spracovaní dát viacerých operátorov je hashovacia funkcia odlišná, a tak dáta nemožno spätne dešifrovať.

AGE_GROUP je definíciou vekovej skupiny, do ktorej vlastník spadá na základe údajov uvedených na faktúre.

CONTRACT_TYPE je typ kontraktu, pod akým je daná SIM karta u operátora zazmluvnená – súkromný alebo služobný.

DAY_ID je časový údaj (dátum), kedy bol signálny záznam vytvorený.

HOUR_ID je časový údaj (hodina), kedy bol signálny záznam vytvorený.

CELL_ID je unikátny identifikátor bunky (BTS stanice), na ktorej bola daná SIM karta v danom dni a danej hodine pripojená. V prípade, že pre danú hodinu vzniklo u konkrétneho užívateľa viacero

CELL_ID záznamov, tieto záznamy budú agregované do jedného – toho, s najčastejším výskytom v tej hodine.

Aby sme mohli určiť polohu konkrétnej SIM karty v danom čase, okrem CELL_ID atribútu budeme potrebovať aj polohu vysielača a jeho vyžarovací polygón. Takáto požiadavka bude zadaná aj operátorom, ktorí budú mať povinnosť dodať nám mapu všetkých CELL_ID (*cell map*), ktoré sa v tejto dátovej vzorke budú vyskytovať. Mapa bude obsahovať presnú lokalitu umiestnenia BTS stanice, jej vyžarovací polygón a unikátny identifikátor kompatibilný s dátovou vzorkou.

Požadované dáta nám budú sprístupnené na back-endových serveroch infraštruktúry operátora, (najpravdepodobnejšie vo formáte .sql dátového extraktu, prípadne vo formáte .csv), ktorý bude následne importovaný na databázové servery priamo u operátora. Samotné dáta pritom neopustia infraštruktúru operátora, budeme k nim pristupovať na diaľku cez autorizačné protokoly a virtuálne privátne siete.

2.3. Spracovanie dát

Dodané datasety spravidla obsahujú podstatne viac záznamov, než je pre potreby konkrétnej štúdie treba. Preto hneď pri úvodnom spracovaní pristúpime k filtrácii záznamov, aby sme dátovú vzorku očistili od nepotrebných údajov a skreslení.

Filtrácia bude prebiehať po vzájomnej dohode so zastupiteľmi Magistrátu Hlavného mesta SR na definíciách jednotlivých typov denne prítomného obyvateľstva / majiteľov SIM kariet – tranzitujúci/trvalo bývajúci a pod.

V prvom kroku však odstránime všetky nevyhovujúce záznamy, ako napríklad SIM karty s nedostatočným množstvom záznamov, SIM karty ktoré nie sú predmetom analýzy a podobne. Odrežeme extrémy, teda SIM karty s podozrivo vysokým, resp. podozrivo nízkym počtom záznamov. Pre potreby tejto štúdie vo vzorke zároveň ponecháme aj SIM-karty, ktoré mali práve jednu udalosť v danej geografickej lokácii - budeme ich považovať za tranzitujúce karty (viď nižšie).

V zozbieraných dátach, sa môžu vyskytnúť aj tieto prípady:
- jedna osoba má dve a viac SIM kariet u jedného operátora - dostupné riešenie pre elimináciu týchto záznamov
- jedna osoba má dve a viac SIM kariet u viacerých operátorov - nedostupné riešenie
- držiteľom SIM karty je iná osoba než tá uvedená v kontrakte - dostupné čiastočné riešenie
- v dátach budú technologické (dátové) aj hlasové SIM karty - dostupné riešenie pre elimináciu týchto záznamov
- nepravidelne aktívne SIM karty, resp. karty, ktoré sú väčšinu času vypnuté a zapnuté iba náhodne - dostupné riešenie pre elimináciu týchto záznamov
- biznis aj súkromné karty vlastnené fyzickými osobami - nevytvára potrebu špecifickej filtrácie datasetu

Cieľom spracovania je vyťažiť z existujúcej vzorky len tú časť záznamov, ktorú možno považovať za relevantnú pre analýzu.

Extrapolácia dát jedného operátora na celkovú populáciu je možná len s výraznými obmedzeniami, keďže portfólia klientov jednotlivých operátorov môžu byť dramaticky vychýlené (napr. vekovou skladbou, historickou dominanciou v regióne a pod.). Líšiť sa takisto môže aj relatívny počet SIM kariet so súhlasmi u jednotlivých operátorov. Ani jeden z týchto prípadov však pri tejto analýze nie je relevantný, keďže zámerom je získať údaje od všetkých troch dominantných operátorov na Slovensku.

Pre prípadné spresnenie počtu reálnych užívateľov SIM kariet je možné vzorku upraviť vzhľadom na podiel užívateľov s aplikovanými geolokačnými súhlasmi (tento podiel sa líši v závislosti od operátora), prípadne vzorku ešte pred spracovaním očistiť o "dátové" a "nehlasové" SIM karty. Uvedené je možné dosiahnuť detailným zadaním pri dopyte dát a získaním dodatočných informácií priamo od operátorov.

Takto očistený dátový súbor je pripravený pre analýzu, ktorej východiskom je identifikácia základných kategórií užívateľov mobilnej siete.

2.4. Charakteristika užívateľov mobilnej siete - základná klasifikácia

Pre potreby tejto štúdie sme spoločne so zástupcami Oddelenia stratégií rozvoja mesta a tvorby územnoplánovacích dokumentov Magistrátu hl. mesta SR Bratislavy vypracovali definície jednotlivých typov užívateľov (SIM kariet). Z pohľadu územného plánovania je kľúčové identifikovať rezidentov, návštevníkov či tranzitujúcich užívateľov, ktorí využívajú zdroje mesta, avšak každý iným spôsobom a iným účelom.

Tieto definície platia v rámci tejto štúdie, a pri jej replikovaní je, samozrejme, možné modifikovať jednotlivé parametre takým spôsobom, aby čo najlepšie zohľadňovali stanovené ciele.

Základom klasifikácie je nočná (spacia) a denná (pracovná) lokalita, ktoré môžeme pri 14-dňovej vzorke definovať nasledovne:

Nočná lokalizácia: Bunka mobilnej siete s najčastejším výskytom užívateľa v dňoch utorok, streda a štvrtok v čase od 20:00 do 05:00 v období dvoch po sebe nasledujúcich týždňov (6-dňový kĺzavý priemer).

Denná lokalizácia 1: Bunka mobilnej siete s najčastejším výskytom užívateľa v dňoch utorok, streda a štvrtok v čase od 09:00 do 15:00 v období dvoch po sebe nasledujúcich týždňov (6-dňový kĺzavý priemer).

Pre zachytenie diferencovanejšieho režimu dňa, než je len tradičná trajektória domov-práca-domov, zavádzame ďalšiu kategóriu dennej lokalizácie, ktorej aplikovaním môžeme zachytiť potenciálny tretí vrchol trojuholníka domov – práca – mimopracovné aktivity.

Denná lokalizácia 2: Bunka mobilnej siete s najčastejším výskytom užívateľa (okrem bunky dennej lokalizácie 1) v dňoch utorok, streda a štvrtok v čase od 06:00 do 22:00 v období dvoch po sebe nasledujúcich týždňov, (6-dňový kĺzavý priemer).

Víkendová nočná lokalizácia: Bunka mobilnej siete s najčastejším výskytom užívateľa v dňoch piatok, sobota, nedeľa v čase od 20:00 do 05:00 v období dvoch po sebe nasledujúcich týždňov (4-dňový priemer).

Víkendová denná lokalizácia: Bunka mobilnej siete s najčastejším výskytom užívateľa v sobotu v čase od 09:00 do15:00 v období dvoch po sebe nasledujúcich týždňov (2-dňový priemer).

Na základe uvedených parametrov môžeme užívateľov v mestskom regióne Bratislavy rozdeliť na nasledujúce elementárne kategórie:

Kategória užívateľa (SIM-karty)	Nočná lokalizácia UT-ŠT, 20:00–05:00	Denná lokalizácia 1 UT-ŠT, 09:00–15:00	Denná lokalizácia 2 UT-ŠT, 06:00–22:00	Víkendová nočná lokalizácia PI-NE 20:00–05:00	Víkendová denná lokalizácia SO 09:00 – 15:00
Rezident A	Bratislava	Bratislava	1. totožná s nočnou lokalizáciou 2. totožná s dennou lokalizáciou 1 3. iná lokalita	Bratislava / mimo Bratislavy	Bratislava / mimo Bratislavy
Rezident B	Bratislava	BA kraj (okrem Bratislavy)	1. totožná s nočnou lokalizáciou 2. totožná s dennou lokalizáciou 1 3. iná lokalita	Bratislava / mimo Bratislavy	Bratislava / mimo Bratislavy
Rezident C	Bratislava	mimo BA kraj	1. totožná s nočnou lokalizáciou 2. totožná s dennou lokalizáciou 1 3. iná lokalita	Bratislava / mimo Bratislavy	Bratislava / mimo Bratislavy
Dochádzajúci A	BA kraj (okrem Bratislavy)	Bratislava	1. totožná s nočnou lokalizáciou 2. totožná s dennou lokalizáciou 1 3. iná lokalita	x	x
Dochádzajúci B	mimo BA kraj	Bratislava	1. totožná s nočnou lokalizáciou 2. totožná s dennou lokalizáciou 1 3. iná lokalita	x	x
Tranzitujúci	x	x	Eventy počas max. 2 hodín na BTS na území mesta	x	x
Nezaradený užívateľ	Eventy počas 3 a viac hodín v mobilnej sieti na území Bratislavy, avšak bez vyprofilovanej dennej a nočnej lokalizácie			x	x

Tab.1. Elementárne kategórie užívateľov mesta Bratislava.

Vyššie uvedené slúži ako teoretické východisko pre aplikáciu nad reálnymi dátami. Po vzájomnej dohode však môže prísť k zmene parametrov v prípade, ak nad finálnym datasetom identifikujeme iné (efektívnejšie, zmysluplnejšie) možnosti spracovania.

2.5. Dátové výstupy

Dátové výstupy budú recipientovi doručené v dohodnutom formáte. Výstupná prechodová matica (origin-destination matrix) bude vo formáte .xls/.xlsx, súbor teda bude čitateľný nástrojom Microsoft Excel. Predmetom dodania sú len analytické výstupy a nie pôvodné zdrojové dáta.

Príklad výstupnej prechodovej matice (údaje sú ilustračné), jednotlivé definície nájdete v kapitole Charakteristika užívateľov - základná klasifikácia:

Pracovná/Spacia location	BA I	BA II	BA III	BA IV	..	..
BA I	16 883	12 444	7 596	1 245		
BA II	7 998	5 689	11 247	9 889		
BA III	4 468	1 234	2 346	6 554		
BA IV	7 895	5 868	4 468	3 321		
...						
...						

Tab. 2. Príklad výstupnej prechodovej O-D matice s počtom užívateľov mobilnej siete agregovaných do okresov Bratislavy.

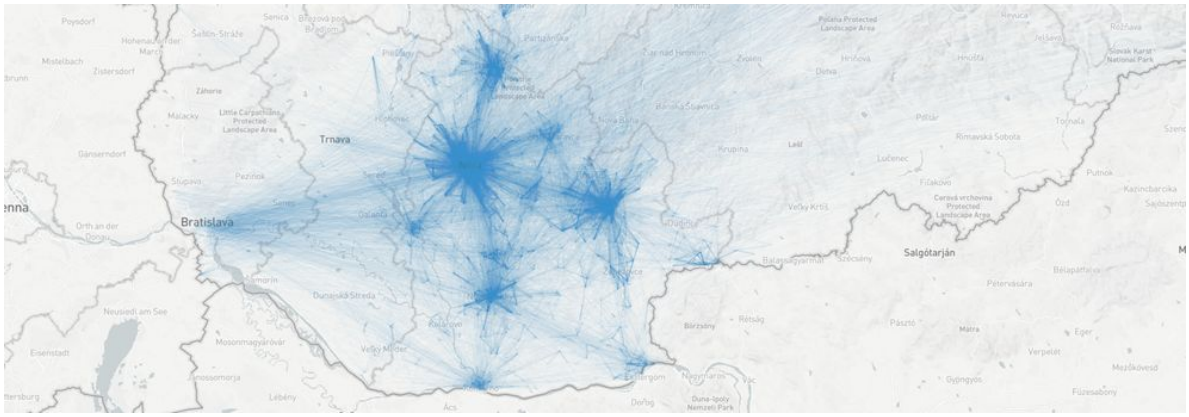
Predmetná prechodová matica s vymyslenými údajmi obsahuje pre každú dvojicu počet SIM kariet, ktoré spĺňajú danú podmienku. Príklad: v okrese BA I má predpočítanú pracovnú lokalitu 16 833 SIM kariet, ktoré majú spaciu lokalitu taktiež v okrese BA I. V okrese BA II má predpočítanú pracovnú lokalitu 7 998 SIM kariet, ktoré majú spaciu lokalitu v okrese BA I.

Rovnaké matice, agregované na dohodnuté polygóny/zóny/okresy mesta bude aj predmetom výstupu finálnej štúdie. Vďaka takýmto maticiam budú môcť zástupcovia mesta nahliadnuť do správania jeho obyvateľov. Matice ponúknu odpovede na otázky, ako napríklad kam sa najčastejšie presúvajú SIM karty s predpočítanou nočnou lokalitou v Petržalke, či z ktorých satelitných oblastí prichádza do Starého mesta najviac SIM kariet. Podľa dostupnosti ďalších atribútov (vek, typ kontraktu, …) bude možné vyrobiť aj prechodové matice z filtrovaných dimenzií.

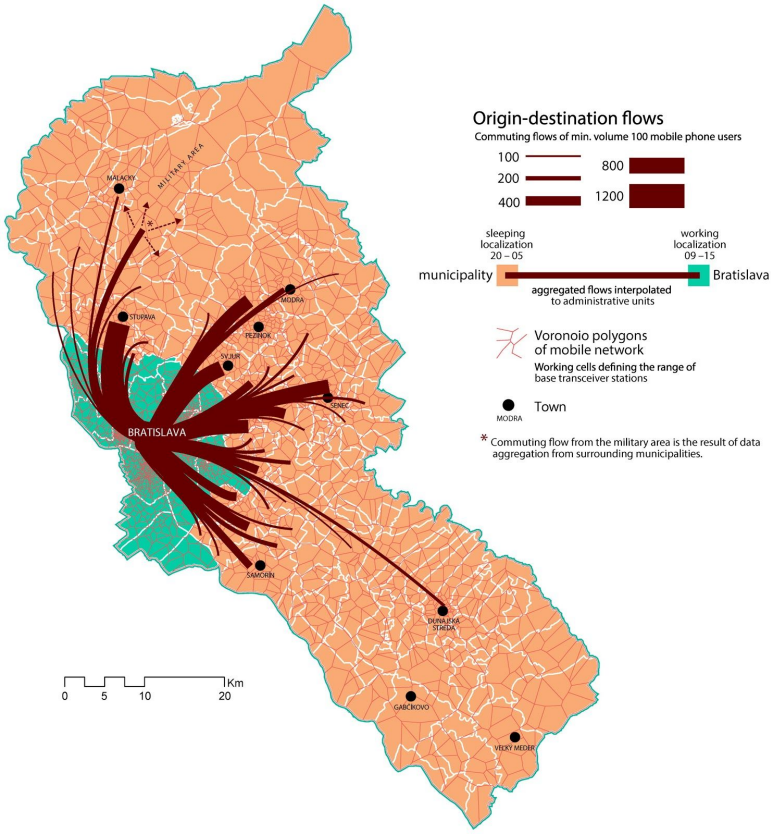
2.6. Vizualizácia dát

Pre finálnu (popularizačnú) vizualizáciu dát budú použité viaceré nástroje. Základné prehľady a grafy vytvoríme v Microsoft Office Excel, pre pokročilejšie vizualizácie použijeme prostredie ArcGIS a programovacie jazyky R a Javascript, resp. ich knižnice a frameworky ako Shiny, Mapdeck a Leaflet.
Výstupom vizualizácií nebude zdrojový kód, ale vizualizácia samotná, dostupná v samostatne spustiteľnej verzii, bez ohľadu na operačný systém či programovú výbavu recipienta.

Predpokladáme obrázkové formáty ako .jpg/.png, a webové formáty ako .html – obe vhodné pre statické a dynamické mapové zobrazenie.
Pre ilustráciu uvádzame príklady vizualizácií z našich predošlých skúseností, v podobnom trende chceme pokračovať aj pri tejto štúdii:



Obr. 5. Príklad vizualizácie prepojenia spacej a pracovnej lokality pre Nitriansky kraj nad umelo vygenerovanými dátami



Obr. 6. Príklad vizualizácie odhadovaných dochádzkových tokov v zázemí Bratislavy s využitím pasívnych lokalizačných údajov sieťovej lokalizácie. Zdroj: Šveda a Barlík 2018

2.7. Replikovateľnosť štúdie

Štúdiu uskutočniteľnosti financovanú Schémou na podporu malých a stredných podnikov pri implementovaní inovatívnych riešení v mestách sme sa snažili od začiatku poňať tak, aby bola v podobných podmienkach bez väčších problémov replikovateľná. Detailne popisujeme zber i formát zdrojových dát, rovnako ich spracovanie a interpretáciu.

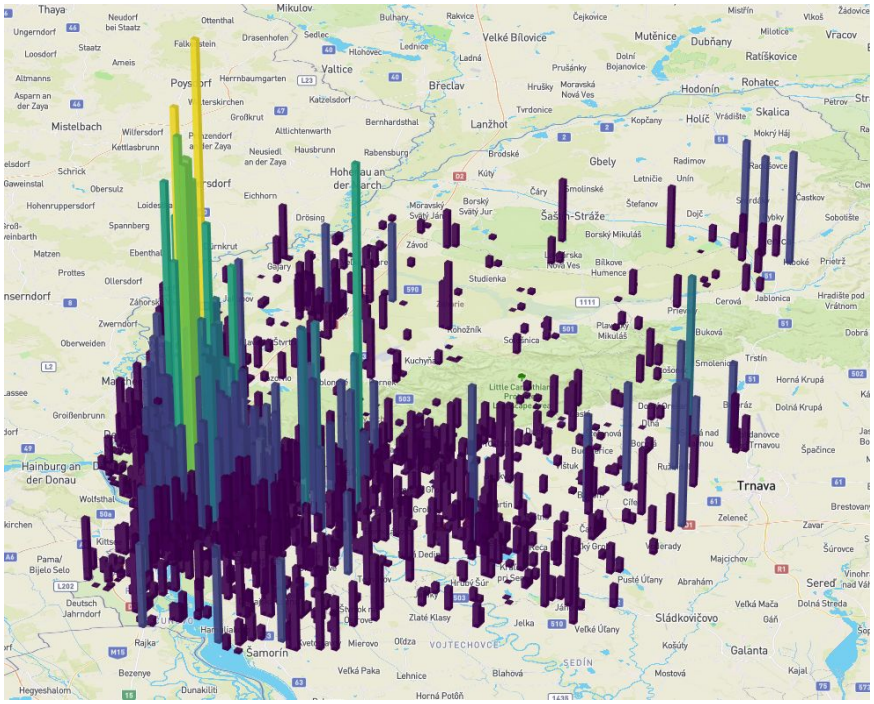
V prípade dostatočných prostriedkov pre nákup anonymizovaných telekomunikačných dát je preto možné obdobnú štúdiu spravovať na teoreticky akomkoľvek regióne s dostatočne detailným pokrytím signálu. Vzhľadom na charakter mobilnej infraštruktúry v neurbanizovanom prostredí (voľnej krajine) neodporúčame predkladanú metodiku aplikovať v periférnych vidieckych či horských oblastiach, kde je pokrytie signálom spravidla zabezpečované niekoľkými BTS bunkami s rozsiahlymi vyžarovacími polygónmi.

2.8. Prípadová štúdia

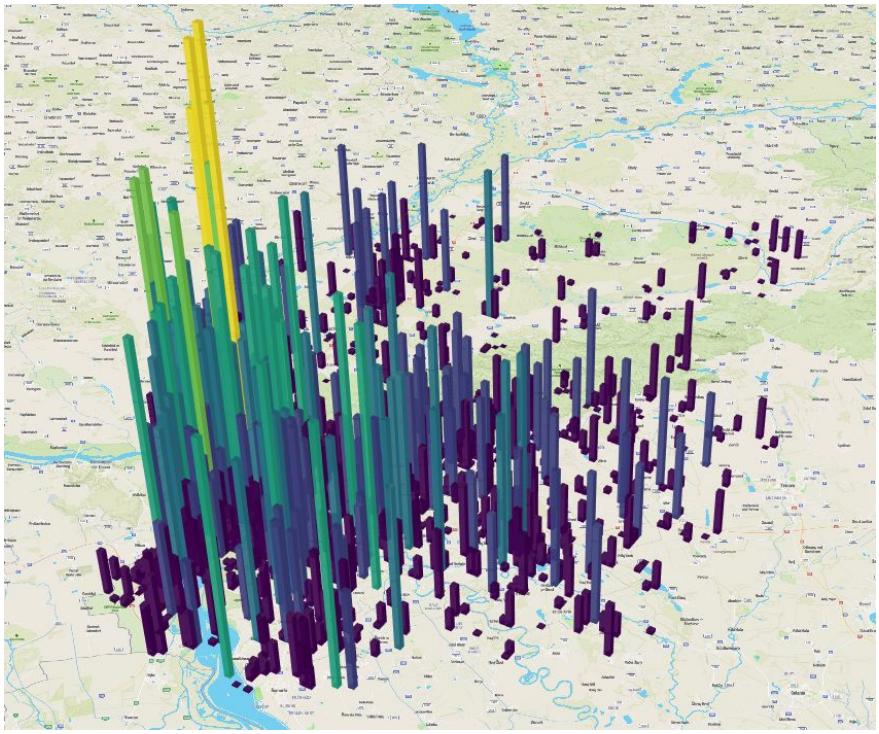
Pre overenie relevantnosti navrhovaného navrhovaného prístupu k spracovaniu telekomunikačných údajov sme spracovali dátovú vzorku za jeden bežný pracovný deň v decembri 2018. Testovacie dáta pochádzajú len od jedného mobilného operátora a nereprezentujú tak potenciál celkovej celoplošnej vzorky zloženej z troch dominantných operátorov.
Obstaranie uceleného dátového súboru (traja dominantí mobilní operátori, dlhšie časové obdobie) je možné iba v prípade pokračovania projektu do II. fázy. Vzorka použitá v predkladanom návrhu projektu ani výstupy nad ňou nie sú určené na publikovanie, iba na overenie relevantnosti teoretickej časti štúdie uskutočniteľnosti.

Počet riadkov (celkový počet udalostí): 13 654 011
Počet unikátnych používateľov: 552 367
Geografický rámec: Bratislavský kraj
Časový rámec: jeden pracovný deň, streda

Pri pilotnej analýze sme pristúpili k orezaniu menej než 1% užívateľov, ktorí mali najviac záznamov. Užívateľ s nízkym počtom záznamov sme ponechali, poslúžia nám na lepšie pochopenie tranzitujúcich.
Na tejto vzorke sme vykonali testovanie, spracovanie a čistenie dát podľa predošlých kapitol, a následne sme podľa definícií z I. časti tejto štúdie uskutočniteľnosti napočítali dodatočné atribúty užívateľov.
Na testovacom datasete sme pristúpili k napočítavaniu zadefinovaných lokalizácií podľa metodiky vyššie. Vypočítať jednotlivé atribúty sa nám podarilo pre:
Nočná lokalizácia: 54,5%
Denná (pracovná) lokalizácia 1: 87,8%
Denná lokalizácia 2: 75,6%
Tranzitujúci: 5%
Nezaradený užívateľ: 4,6%



Obr. 7. Príklad vizualizácie spacej lokality nad testovacím datasetom. Výška stĺpca predstavuje počet unikátnych SIM kariet, ktoré majú v danej lokalite predpočítanú spaciu lokalitu. Šírka stĺpca je 500x500 metrov a spadajú do nej všetky BTS bunky s centroidom vyžarovacieho polygónu v danom štvorci



Obr. 8. Príklad vizualizácie pracovnej lokality nad testovacím datasetom.

Pre výpočet víkendovej lokácie bola testovacia dátová vzorka nedostatočná. Ostatné lokality sa nám však v rámci možnosti testovacej vzorky podarilo vypočítať na základe stanovených definícií. Uvedený podiel jednotlivých vypočítaných lokácií bude pri viacdňovej vzorke už iba narastať, výrazným limitom pilotu bola práve obmedzená dátová vzorka.

SUBSCRIBER_ID	AGE_GROUP	CONTRACT_TYPE	DAY_ID	HOUR_ID	CELL_ID
abde45Gaymm	56-65	B	20183012	17	3485
cca49tRuwY3a	18-25	C	20183012	17	1123
aa97dcgfeXa1a	.	.	20183012	19	4409

Tab. 3. Príklad dátového extraktu, ktorý vstupuje do analýzy (dáta sú umelo vygenerované)

CELL_ID	lat	long	Polygon
3485	48.1532858411	17.2255080862	POLYGON((17.16335325910063 48.19623225179323, ...))
1123	48.5748134791	17.3385843789	POLYGON((17.22652463741963 48.19073949092182, ...))
4409	48.1853605964	17.2165228896	POLYGON((17.15561113137835 48.13840454630733

Tab. 4. Príklad dátového extraktu polygonálnej mapy (dáta sú umelo vygenerované)

Uvedená tabuľka reprezentuje príklad polygonálnej mapy, kde pre konkrétnu bunku poznáme jej polohu a vyžarovací polygón. Práve vyžarovací polygón sa následne spája s vypočítanými lokalitami jednotlivých užívateľov a agreguje sa na príslušnú územnú jednotku.

SUBSCRIBER_ID	počet zázna mov	Nočná lokalizácia	Denná lokalizácia I.	Denná lokalizácia II.	Víkendová nočná lokalizácia	Víkendová denná lokalizácia
abde45Gaymm	315	4456	4409	1378	null	1228
cca49tRuwY3a	124	1123	null	3248	4476	1978
aa97dcgfeXa1	1	null	null	null	null	null

Tab. 5.: Príklad nápočtu lokalít pre konkrétnych užívateľov

Vo finálnej analýze získame pre anonymizovaný unikátny identifikátor SIM karty (SUBSCRIBER_ID) nápočty lokácií reprezentované príslušnou BTS bunkou. Jej vyžarovací polygón bude reprezentovaný jediným bodom - centroidom. Následne získame pre jednotlivé GPS body počet pripojených SIM kariet. Ak takýto pohľad preložíme štvorcovou maticou

získame ucelený pohľad pripojených zariadení v danom čase na danom mieste, resp. v hraniciach definícií.

Samozrejme, nie všetky unikátne SIM karty budú mať dostatok záznamov pre nápočet všetkých lokácií. Pre takéto SIM karty sme pripravili označenie “Tranzitujúci” a “Nezaradený užívateľ”.

2.9. Manažment pilotného projektu

Do pilotného projektu a tvorby metodických usmernení pre II. fázu projektu bola zapojená mnohočlenná pracovná skupina, tvorená zástupcami spoločnosti Instarea/Market Locator, Oddelenia Územného plánovania magistrátu Bratislavy a odbornými garantmi z Prírodovedeckej fakulty Univerzity Komenského v Bratislave.

Komunikácia prebieha na báze pravidelných pracovných stretnutí a podľa potreby aj na báze e-mailovej a telefonickej komunikácie. Mesto, ako objednávateľa projektu, na stretnutiach zastupuje Oddelenie stratégií rozvoja mesta a tvorby územnoplánovacích dokumentov, ktoré zhotoviteľovi zadáva požiadavky na spracovanie získaných dát a podklady do metodológie.

Za Market Locator/Instarea: Mgr. Peter Barlík PhD., Ing. Martin Bago,

Za Územné plánovanie: Ing. arch. Dvorčák, Ing. Petrinec, RNDr. Goga, Ing.arch. Lexmann, RNDr. Babiar Michal PhD.

Za Prírodovedeckú fakultu Univerzity Komenského v Bratislave: Mgr. Martin Šveda, PhD., doc. RNDr. František Križan, PhD.

Výsledkami stretnutí pracovnej skupiny sú okrem definícií záujmových skupín aj tento dokument ako celok.

Pevne veríme, že v prípade ak náš projekt získa podporu pre svoju II. fázu, bude táto pracovná skupina naďalej poskytovať širokú podporu a tím doručí riešenie prospešné všetkým zúčastneným.

Harmonogram I. fázy projektu:

aktivita/mesiac	I.	II.
I.	Príprava metodologických podkladov, stretnutia so zástupcami Magistrátu, tvorba Metodiky	
II.		Tvorba Metodiky a pilotu

Plánovaný harmonogram II. fázy projektu:

aktivita/mesiac	I.	II.	III.	IV.	V.	VI.
I.	Oslovenie 1-3 operátorov pre nákup dát					
II.		Spracovanie, čistenie a kompletizácia dát				
III.				Napočítavanie lokalít, tvorba O-D matice, vizualizácie, štúdia		

Na realizáciu II. fázy sú odhadované finančné náklady vo výške 100 000€, ktoré budú primárne použité na obstaranie lokalizačných dát od troch najväčších operátorov na Slovensku. Nakoľko finančné náklady potrebné na obstaranie dát presahujú 50 000€, ktoré sú alokované vo fáze II., na zvyšné prostriedky (max do 50%) sa spoločnosť Market Locator SK, s.r.o. v spolupráci s operátormi pokúsi vyjednať zľavu, ako spôsob spolufinancovania projektu. Vzhľadom na uvedené okolnosti a výšku zľavy, ktorú sa podarí vyjednať s operátormi, bude pre naplnenie cieľov II. fázy pravdepodobne nutné pristúpiť k okresaniu dátovej vzorky (dáta od 1-3 operátorov, menší priestorový či časový rozsah dát a pod.). Tým zabezpečíme nielen dodržanie finančného plánu, ale po úspešnom obstaraní dát aj časového rozvrhu s úspešným dodaním štúdie na konci sledovaného obdobia.

Aj v II. fáze plánujeme pokračovať v rovnakom nastavení spolupráce so zástupcami mesta a akademickou pôdou.

3. Aplikácia a udržateľnosť projektu

3.1. Nadväznosť na strategické ciele mesta

Projekt je previazaný s prioritami Programu hospodárskeho a sociálneho rozvoja hl. mesta SR Bratislavy na roky 2010-2020, menovite:

A.B.I) Posilnenie vitálneho podnikateľského prostredia, vďaka ktorému bude v záujme investora v Bratislave zostať, čím si mesto udrží konkurencieschopnú pozíciu aj voči okolitým metropolám ako Viedeň a Budapešť, ale aj posilní atraktivitu pre nových investorov.

a

B. III) Program spolupráce akademických a výskumných inštitúcií s podnikmi.

3.2. Dlhodobá udržateľnosť projektu

Projekt predstavuje pilotnú aplikáciu poznatkov v oblasti využívania lokalizačných údajov mobilných operátorov pre potreby rozvoja hlavného mesta SR Bratislavy. Svojim úzkym časovým záberom (dva po sebe idúce týždne) a špecifickým zameraním na sledovanie odhadu pobytu a pohybu denne prítomného obyvateľstva po Bratislave, ako aj denne dochádzajúceho obyvateľstva pre potreby územného plánovania by mal pre hl. mesto SR Bratislavu slúžiť ako úvod do problematiky využívania mobilných dát pre potreby správy mesta.
Skúsenosti a poznatky z tohto projektu, vrátane metodiky spracovania dát a prípadných nedostatkov a výziev zistených pri realizácií projektu, budú využité v budúcnosti pri realizácii podobných projektov vo väčšom rozsahu. Schopnosť zachytiť priestorovú dynamiku populácie považujeme za nevyhnutný predpoklad pre moderný manažment mesta a dlhodobé územné plánovanie.

3.3. Ekonomická analýza (vrátane dopadu riešenia na mesto)

Hlavné mesto v súčasnosti nerealizuje na pravidelnej báze žiadne alternatívne riešenie, ktorého cieľom by bolo systematicky získavať obdobné informácie, ako poskytuje riešenie na báze lokalizačných údajov mobilných telefónov. Na svoje potreby využíva štatistické údaje a údaje generované v rámci činnosti mesta a mestských organizácií. Avšak hlavné mesto tento zber a vyhodnocovanie údajov v súčasnosti nerobí cielene a strategicky vzhľadom na absentujúcu komplexnú dátovú politiku. Chýbajúca dátová politika a tlak na plánovanie založenom na dátach, nevytváral potrebu vytvorenia rozpočtovej položky na obstarávanie externých dát alebo ich cielené získavanie od externých partnerov.

Navrhované riešenie je preto z hľadiska pomeru nákladov a výnosov vysoko efektívnym spôsobom zberu rozsiahlych údajov, ktorých správne využitie v praxi môže mať pozitívny dopad

aj na samotné mesto a jeho obyvateľov. Skúsenosti zo zahraničia ukazujú, že analýza dennej priestorovej mobility na základe lokalizačných údajov z mobilných telefónov má vysoký potenciál v oblasti riadenia dopravy, územného plánovania, cestovného ruchu či krízového riadenia s dlhodobým dopadom na ekonomiku, zdravie či spokojnosť obyvateľov a návštevníkov mesta. Najvýraznejší prínos tohto riešenia sa javí práve v oblasti mobility a územného plánovania. Mesto, ktoré pozná dennú priestorovú dynamiku svojho obyvateľstva, má kapacitu vyvíjať adresnejšie úsilie pri skracovaní vzdialeností dochádzania za prácou alebo zábavou, navrhovaní vhodných trás alternatívnej prepravy alebo posilňovaní spojov hromadnej dopravy a občianskeho vybavenia tam, kde je to najviac potrebné, a tým skracovať čas potrebný na presun a podnecovať využívanie alternatívnych spôsobom dopravy (chôdza, cyklodoprava, väčšie využívanie mestskej hromadnej dopravy). Ďalším pozitívom je znalosť pohybu obyvateľstva v mestskom priestore a príprava politík na plánovanie zón, ktoré by vyhovovali potrebám obyvateľov a návštevníkov, a tiež oživili "zanedbané" verejné priestory.

Získané dáta môžu využívať okrem mesta aj mestské časti, ktoré môžu dáta využiť pri príprave strategických dokumentov, ako je územný plán či Program hospodárskeho a sociálneho rozvoja.
Pilotná štúdia je podkladom na rozhodovanie pre využitie dát mobilných operátorov, ich nedostatky a potenciálne prínosy. Presnosť údajov, ich možný prekryv s existujúcimi urbanistickými jednotkami a schopnosť kombinovať ich s inými dátami generovanými mestom a o meste je kľúčová pri budúcom použití lokalizačných dát mobilných operátorov.

3.4. Alternatívne možnosti financovania riešenia v súčasnosti/budúcnosti

Hlavné mesto Bratislava má viaceré možnosti financovania experimentálnych projektov z oblasti plánovania mesta založeného na dátach prostredníctvom fondov EÚ či iných finančných zdrojov zo zahraničia alebo na národnej úrovni.

Ďalšie pilotné štúdie môžu overiť relevantnosť lokalizačných dát pre ďalšie využitie pre strategické plánovanie.

Na základe skúseností zo zahraničia sa predpokladá dlhodobá implementácia riešenia v budúcnosti formou vytvorenia informačného systému, ktorý na základe špecifikácií jednotlivých užívateľov (mesta, mestských častí, mestských organizácií) bude schopný v takmer reálnom čase spracovať vstupné dáta od mobilných operátorov do podoby výstupných údajov, ktoré budú využiteľné pre rôzne cieľové skupiny. Možná návratnosť vynaložených finančných prostriedkov je aj v možnosti poskytovania výstupných údajov tretím stranám aj na komerčné účely, čím by sa čiastočne pokryli náklady na prevádzku systému a nákup dát od mobilných operátorov.

3.5. Udržateľnosť a efektívnosť projektu

Udržateľnosť a efektívnosť projektu závisí od prípravy a implementácie dátovej politiky, ktorá zadefinuje formálne, technologické aj obsahové ciele a následne bude reagovať na strategické a operatívne potreby mesta.

Využiteľnosť údajov mobilných operátorov pre potreby územného plánovania, bude overená pilotnou štúdiou.

3.6. SWOT analýza

Celoplošné riešenie

Silné stránky:

- Demografické pokrytie: využitie mobilných telefónov je dnes rovnomerne distribuované naprieč všetkými významnými socioekonomickými skupinami bez ohľadu na pohlavie, vek, sociálny status, zamestnanie, atď. a dáta z mobilných telefónov teda poskytujú vysoko reprezentatívnu vzorku.
- Penetrácia mobilných zariadení je v rámci Bratislavy vysoká.
- Rýchlosť a frekvencia: v porovnaní s tradičnými zdrojmi údajov, ktorých aktuálnosť je limitovaná potrebou dlhšieho času na spracovanie (ako napr. pri národných cenzoch), zisťovanie prostredníctvom údajov z mobilnej siete môže prebiehať akokoľvek dlho a s ľubovoľnou frekvenciou.
- Finančné náklady spojené s obstaraním a spracovaním lokalizačných údajov mobilných telefónov sú neporovnateľne nižšie, než sú bežné výdavky spojené s prípravou, priebehom a realizáciou celoplošných cenzov.
- Veľké množstvo dát: využívanie mobilných telefónov počas každodenných činností obyvateľov je veľmi vysoké a tak umožňuje relevantný pohľad do využívaniu mesta

Slabé stránky:

- Priestorová presnosť: územné pokrytie BTS vysielačov nekorešponduje s hranicami urbanistických jednotiek v meste (mestské časti/sídliská/zóny/...), ktoré majú pre účely územného a dopravného plánovania kľúčový význam.
- Finančná náročnosť obstarania veľkého množstva údajov a ich pravidelné získavanie.

Príležitosti:

- Spolupráca s mobilnými operátormi a ich následné vyhodnocovanie vrátane nákupu dát za obojstranne výhodných podmienok
- Spolupráca s akademickým sektorom pri tvorbe metodológie a kalibrovaní vstupných údajov pre zabezpečenie čo najpresnejších výstupov
- Spolupráca s podnikateľským a tretím sektorom pri zlepšení správy mesta

Riziká:

- Negatívny postoj verejnosti a obavy z nedostatočnej ochrany osobných údajov
- Možnosť kombinovania dát s inými dátovými zdrojmi mesta

Pilotný projekt

Silné stránky:

- Získanie historických dát o odhade pobytu a pohybu denne prítomného obyvateľstva a priestorové využitie mesta
- Historické porovnanie pracovného dňa a víkendu pobytu a pohybu obyvateľov
- Prvá štúdia svojho druhu pre hlavné mesto, kde mesto získa reálny pohľad na potreby obyvateľov

Slabé stránky:

- Časový rozsah vzorky
- Veľkosť monitorovanej oblasti
- Definovanie vzťahu počtu SIM kariet na 1 obyvateľa Bratislavy
- Vylúčenie SIM kariet registrovaných na podnikateľské a firemné účely
- Priestorová presnosť: územné pokrytie BTS vysielačov nekorešponduje s hranicami urbanistických jednotiek v meste (mestské časti/sídliská/zóny/...), ktoré majú pre účely územného a dopravného plánovania kľúčový význam.

Príležitosti:

- Overenie potreby lokalizačných dát pri strategickom územnom plánovaní
- Validácia potreby dát pri iných funkciách mesta
- Porovanie reálneho stavu oproti dostupným štatiským údajom a definovanie miest na zlepšenie v zbere a vyhodnocovaní dát

Riziká:

- Využiteľnosť dát pre iné oddelenia
- Nedostatočné finančné zdroje na rozvoj ďalších pilotných projektov
- Možnosť prepojenia dát s inými dátovými zdrojmi a ich efektívne vyhodnotenie
- Neupresnené percentuálne pokrytie trhu (nezahrnutí zákazníci virtuálnych mobilných operátorov)

4. Záverečné zhodnotenie

Aplikácia úloh z časti I. tejto štúdie uskutočniteľnosti na malej testovacej vzorke potvrdila, že pre účely zisťovania denne prítomného obyvateľstva, sú lokalizačné údaje mobilných telefónov zmysluplným dátovým zdrojom. Testovaciu vzorku sme čistením a filtráciou pripravili tak, aby bolo možné vypočítať hlavné kotevné body priestorovej mobility obyvateľstva (spaciu a pracovnú lokalitu užívateľov SIM kariet) a identifikovať tak vzorce ich pohybu v hlavnom meste Slovenskej republiky, Bratislave.

Z pilotnej realizácie vyplýva, že vhodne spracované telekomunikačné dáta je možné používať ako dodatočný informačný zdroj pri analýzach populácie či dopravy. Vhodným kombinovaním telekomunikačných dát s inými dátovými zdrojmi - napríklad od Štatistického úradu SR, jednotlivých Ministerstiev SR, či dátami hlavného mesta - je možné lepšie/hlbšie nahliadnuť do priestorového správania rezidentov i užívateľov Bratislavy, a poskytnúť im tak lepšie služby, či napríklad vyššiu bezpečnosť.

Zároveň zo štúdie vyplývajú viaceré limitácie telekomunikačných dát. Či už z hľadiska presnosti, socio-demografických údajov alebo z hľadiska absolútneho počtu prítomného obyvateľstva môžu telekomunikačné dáta obsahovať skreslenia definované v kapitole Limity dát. Vzhľadom na uvedené, tento dátový zdroj má charakter doplnkového dátového zdroja pri územnom plánovaní či analýze priestorového správania obyvateľstva.

Vhodnou kombináciou dostupných dátových zdrojov sa však limitácie lokalizačných údajov mobilných telefónov znižujú a užívateľ týchto dát, v tomto prípade Magistrát hlavného mesta Bratislava a jeho jednotlivé sekcie, tak získavajú vhodný dátový zdroj pre riešenie širokého spektra otázok.

Prílohy

Príloha 1: Vyúčtovanie výdavkov za I. fázu

Príloha 2: Kópie zmluvy s mestom Bratislava o prebratí diela a Protokolu o odovzdaní a prevzatí

Príloha 3: Zoznam použitej literatúry

Príloha 3: Zoznam použitej literatúry

AHAS, R., AASA, A., ROOSEA, A., MARK, Ü., AND SILMA, S. (2008). Evaluating passive mobile positioning data for tourism surveys: An Estonian case study. Tourism Management, 29, 3, 469-486.

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.

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.

AHAS, R., MARK, Ü. (2005). Location based services – new challenges for planning and public administration? Futures, 37, 6, 547-561.

BALK, D.L., DEICHMANN, U., YETMAN, G., POZZI, F., HAY, S.I., NELSON, A. (2006). Determining global population distribution: methods, applications and data. Advances in Parasitology, 62, 119-156.

BATTY, M. (1990). Invisible cities. Environment and Planning B: Planning and Design, 17, 127-130.

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.

BHADURI, B., BRIGHT, E., COLEMAN, P., URBAN, M. (2007). LandScan USA: A high-resolution geospatial and temporal modeling approach for population distribution and dynamics. GeoJournal, 69, 1, 103-117.

BLONDEL, V.D., ESCH, M., CHAN, C., CLÉROT, F., DEVILLE, P., HUENS, E., MORLOT, F.,

BROCKMANN, D., HUFNAGEL, L., GEISEL, T. (2006). The scaling laws of human travel. Nature 439, 462-465.

CSÁJI, B., BROWET, A., TRAAG, V.A., DELVENNE, J.C., HUENS, E., VAN DOOREN, P.,

DEICHMANN, U., BALK. D., YETMAN, G. (2001). Transforming Population Data for Interdisciplinary Usages: From Census to Grid. Washington, DC (Center for International Earth Science Information Network).

DEVILLE, P., LINARD, C., MARTIN, S., GILBERT, M., STEVENS, F.R., GAUGHAN, A.E., BLONDEL, V.D., TATEM, A.J. (2014). Dynamic population mapping using mobile phone data. Proceedings of the National Academy of Sciences, 111, 45, 15888-15893.

EUROSTAT (2017). Africa – EU: Key statistical indicators.
GONZÁLEZ, M. C., HIDALGO, C. A., BARABÁSIM A. L. (2008). Understanding individual human mobility patterns. Nature, 453, 7196, 779-782.

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

Kubíček, P., Konečný, M., Stachoň, Z., Shen, J., Herman, L., Řezník, T., ... & Leitgeb, Š. (2018). Population distribution modelling at fine spatio-temporal scale based on mobile phone data. International Journal of Digital Earth, 1-22.

Lai, J., Chen, Y., Wu, Z., Yuan, G., & Li, M. (2018). Application of Cellular Data in Traffic Planning. In Big Data Support of Urban Planning and Management (pp. 273-296). Springer, Cham.

LEUNG, S., MARTIN, D., COCKINGS, S. (2010). Linking UK public geospatial data to build 24/7 space-time specific population surface models. GIScience 2010: Sixth International Conference on Geographic Information Science. Zurich (Springer LNCS).

Masso, A., Silm, S., & Ahas, R. (2018). Generational differences in spatial mobility: A study with mobile phone data. Population, Space and Place, e2210.

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

PUCCI, P., COLLEONI, M. (eds.) (2016). Understanding Mobilities for Designing Contemporary Cities. Basel (Springer).

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.

ROMÁN, M. O., STOKES, E. C. (2015). Holidays in lights: Tracking cultural patterns in demand for energy services. Earth's Future, 3, 182-205.

SHOVAL N. (2007). Sensing human society. Environment and Planning B: Planning and Design, 34, 191-195.

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

ŠVEDA, M., BARLÍK, P. (2018): Daily commuting in the Bratislava metropolitan area: case study with mobile positioning data. Papers in Applied Geography, DOI: 10.1080/23754931.2018.1540357

ŠÚ SR (2009). Výročná správa Štatistického úradu SR za rok 2009. Štatistický úrad Slovenskej republiky.

ŠÚ SR (2010). Výročná správa Štatistického úradu SR za rok 2010. Štatistický úrad Slovenskej republiky.

ŠÚ SR (2011). Výročná správa Štatistického úradu SR za rok 2011. Štatistický úrad Slovenskej republiky.

ŠÚ SR (2012). Výročná správa Štatistického úradu SR za rok 2012. Štatistický úrad Slovenskej republiky.

TATEM, A., LINARD, C. (2011). Population mapping of poor countries. Nature, 474, 36, 400-401.

TOBLER, W., DEICHMANN, U., GOTTSEGEN, J., MALOY, K. (1997). World population in a grid of spherical quadrilaterals. International Journal of Population Geography, 3, 3, 203-225.

WATTS, D.J. (2007). A twenty-first century science. Nature, 445, 489.

WESOLOWSKI, A., EAGLE, N., NOOR, A.M., SNOW, R.W., BUCKEE, C.O. (2013). The impact of biases in mobile phone ownership on estimates of human mobility. Journal of the Royal Society Interface, 10, 81.

Yang, X., Fang, Z., Yin, L., Li, J., Zhou, Y., & Lu, S. (2018). Understanding the spatial structure of urban commuting using mobile phone location data: a case study of Shenzhen, China. Sustainability, 10(5), 1435.

Zhai, Y., Wu, H., Fan, H., & Wang, D. (2018). Using mobile signaling data to exam urban park service radius in Shanghai: methods and limitations. Computers, Environment and Urban Systems, 71, 27-40.