

# VYUŽITÍ PERSONALIZOVANÝCH DAT GOOGLE LOCATION PRO ANALÝZU ZMĚN CHOVÁNÍ V PRŮBĚHU PANDEMIE COVID-19

**Jaroslav Burian<sup>1,2</sup>, Vít Pászto<sup>1,2</sup>, Karel Macků<sup>2</sup>, Lukáš Pavlík<sup>1</sup>**

<sup>1</sup> Moravská vysoká škola Olomouc,  
jaroslav.burian@mvso.cz, vit.paszto@mvso.cz, lukas.pavlik@mvso.cz

<sup>2</sup> Univerzita Palackého v Olomouci,  
karel.macku@upol.cz

## **Abstract:**

The paper focuses on a detailed micro-study describing changes in the behaviour of the authors of this article before and during the COVID-19 pandemic based on data from Google Location. Although the study evaluates data on only three individuals and may not be sufficiently representative to draw generalizing conclusions, it is a unique example of possible data processing that is not normally evaluated at such a level of accuracy. The article evaluates changes in the behaviour of authors before and after the introduction of anti-epidemic measures, but the main purpose is to point out the possibilities of analytical processing of such data while being aware of the issue of personal data protection. We recognize that by visualizing our real motion data, we partially disclose our privacy. Nevertheless, we find it very valuable to show how detailed data Google collects about much of the population and how such data can be used effectively.

## **Abstrakt:**

Článek je zaměřen na detailní mikro-studii popisující změny v chování autorů tohoto článku před a během pandemie COVID-19 na základě dat ze služby Google Location. Přestože studie vyhodnocuje data o pouze třech osobách a nemůže tak být dostatečně reprezentativní pro vyvození generalizujících závěrů, jedná se o unikátní ukázkou možného zpracování dat, které běžně nejsou v takové úrovni přesnosti vyhodnocována. V článku jsou hodnoceny změny v chování autorů před a po zavedení protiepidemických opatření, avšak hlavním smyslem je poukázat na možnosti analytického zpracování takovýchto dat současně s vědomím problematiky ochrany osobních údajů. Jsme si vědomi, že vizualizací našich reálných dat o pohybu částečně odhalujeme své soukromí. Považujeme však za velmi cenné ukázat, jak podrobná data o velké části populace společnost Google sbírá a jak lze taková data efektivně využívat.

**Key words:** *Google Location, mikrostudie, geodata, GIS, COVID-19*

**JEL:** C6, C8

## **1 Úvod**

Hodnocení změn chování osob z pohledu pohybu v prostoru bylo po mnoho desetiletí prováděno především pomocí dotazníkových šetření, případně doplněných cestovními deníky. V současnosti jsou tyto metody doplňovány a často nahrazovány moderními zdroji dat označovanými jako big data, jejichž význam roste v posledních několika letech prakticky ve všech oblastech lidské činnosti. Ačkoliv je téma big dat intenzivně diskutováno v posledních několika letech, poprvé se pojem ve spojení s vizualizací

objevil již před více než 20 lety (Cox & Ellsworth, 1997). Batty (Batty, 2016) za big data označuje takové datové sady, které není možné zpracovat běžným hardwarem a softwarem.

Big data jsou sbírána v rámci sledování flotil vozidel, tzv. „Floating Car Data“ (Hao, Zhu, & Zhong, 2015), která umožňují jak sledování pohybu obyvatel, tak především zjišťování cestovní rychlosti, které jsou zásadním zdrojem pro většinu inteligentních dopravních systémů. Dále jde např. o data z GPS přijímačů taxi služeb (Castro & Zhang, 2013), kamerové systémy (Wang, 2013) nebo data logů přístupových bodů WiFi sítí (Sapiezynski, Stopczynski, Gatej, & Lehmann, 2015).

V oblasti sociálních sítí jde nejčastěji o analýzu big dat získaných ze sítě Twitter nebo Foursquare (Hawelka et al., 2014; Kocich, 2018), přičemž časté je využití těchto dat právě pro sledování mobility osob (Noulas, Scellato, Lambiotte, Pontil, & Mascolo, 2012; Long, Jin, & Joshi, 2012; Noulas, Scellato, Mascolo, & Pontil, 2011). Obdobně jsou analyzována data z mnoha dalších sociálních sítí (Frothingham, 2014; Haworth, 2016; Musakwa & Selala, 2016), mezi které patří také sítě zaměřené na záznam sportovních aktivit, ze kterých lze identifikovat nejčastěji využívané cyklostezky, dopravní chování účastníků mobility (Sun, 2017) nebo analyzovat objemy cyklistické přepravy (Hochmair, Bardin, & Ahmouda, 2017; Romanillos, Zaltz Austwick, Etema, & De Kruijff, 2016; Selala & Musakwa, 2016).

Pro sledování pohybu osob je však podle Battyho (Batty, 2013, 2016) klíčová existence mobilních telefonů a s nimi spojené velké množství potenciálních dat. V současnosti je vlastnictví mobilního telefonu považováno takřka za samozřejmé, a proto mohou data mobilních operátorů sloužit jako nejlepší zdroj pro identifikaci skutečné polohy obyvatel v čase. Problémem těchto dat je však jejich nedostupnost, ať už z důvodu ceny, nekompletnosti, nebo jiných objektivních kritérií.

Významnou roli hraje také otázka ochrany osobních údajů, neboť se jedná o osobní údaje osob. Problematika ochrany osobních údajů fyzických osob je řešena v rámci legislativy GDPR (General Data Protection Regulation), která stanovuje pravidla pro jejich zpracování, včetně stanovení možných finančních pokut za únik osobních údajů nebo jejich poškození. V této legislativě jsou také stanoveny požadavky a povinnosti správců osobních údajů na zabezpečení citlivých osobních dat a jejich zdrojů. Osobní údaje o zákaznících mobilních operátorů jsou často využívána např. bankovními institucemi pro TelcoScore, tedy zjišťování solidnosti a důvěryhodnosti posuzovaného klienta. Tento postup není podle dozorových orgánů, které jsou dotčeny řešením problematiky ochrany osobních údajů v rozporu legislativou. Nicméně se jedná o problematiku, o které by měl být konkrétní uživatel mobilního operátora informován a měl by mít právo na poskytnutí informací o tomto hodnocení ze strany mobilního operátora (Kurian et al., 2020; Štarchoň & Pikulík, 2019).

Další významný pohled na problematiku ochrany osobních údajů přináší využití mobilních technologií a dat o poloze jednotlivých uživatelů v rámci pandemie Covid-19. Data, která získávají mobilní operátoři, mohou být využita pro určení trasy nakaženého člověka, včetně identifikace jeho kontaktů v čase. Získaná data mobilních operátorů jsou předávána hygienickým orgánům, které jsou zodpovědné za kontaktování pozitivně nakažených lidí. Uživatelé mobilních telefonů však musí dát souhlas s předáním jejich osobních dat příslušnému orgánu, který je v této věci dále kontaktuje. Mobilní data nejsou jediným zdrojem pro určení polohy člověka. V rámci pandemie mohou být také využity informace z platebních karet. Mobilní operátoři a jiné subjekty však uchovávají a zpracovávají osobní data uživatelů nebo klientů po nezbytně nutnou dobu, která je stanovena zákonem. Po uplynutí této doby jsou většinou data smazána a to především z bezpečnostních důvodů. Požadavky a nároky na ukládání osobních nebo biometrických dat jsou poměrně náročné a žádná z organizací, která má tato citlivá osobní data k dispozici, je ve většině případů nechce uchovávat déle, než stanoveno příslušnou legislativou (Skendzic, Kovacic, & Tijan, 2018).

I přesto začíná narůstat četnost využití těchto dat (Calabrese, Diao, Di Lorenzo, Ferreira, & Ratti, 2013; Kahrik, Novák, Temelová, Kadarik, & Tammaru, 2015; Novák & Temelová, 2012). V nedávné době tato data ve větším rozsahu (avšak agregovaná do základních sídelních jednotek) pořídil v Česku IPR Praha (Čtyrský, Novotný, & Soukup, 2016; Zďerad, 2016; Soukup, 2017) a také Magistrát města Brna a na jejich základě analyzovali pohyb obyvatel a návštěvníků Prahy, Brna a nejbližšího okolí v průběhu dne a noci.

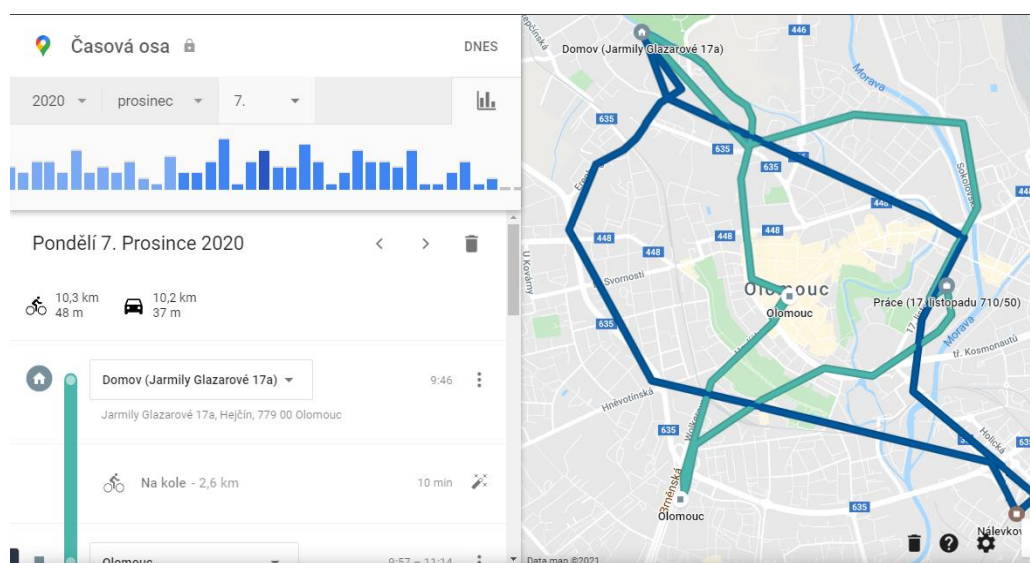
S daty mobilních operátorů je pracováno v současné době také při řešení problémů spojených s nemocí COVID-19. V Česku jsou data od mobilních operátorů využívána v rámci tzv. „chytré karantény“, která je založena na vizualizaci individuálních dat mobilních operátorů pro snadnější zjištění pohybu nakažených osob. Data jsou na základě souhlasu nakaženého dodávána soukromé společnosti CleverMaps, která vytváří tzv. vzpomínkové mapy, tedy grafickou vizualizaci pohybu za posledních 5 dní. Takováto vizualizace může následně pomoci nakažené osobě vzpomenout si na pravděpodobná místa pohybu a následně osoby, se kterými se setkala. Data však podléhají vysokému stupni ochrany a nejsou dostupná komukoliv dalšímu.

Pro analýzu pohybu obyvatel se však kromě dat mobilních operátorů nabízí využití dat ze služby Google Location (sledování polohy). Tato data jsou sbírána společností Google, pokud uživatel aplikace Google Maps (v rámci které funguje služba Google Location) souhlasí se sdílením své polohy. Poloha zařízení (nejčastěji mobilního telefonu) je zjišťována nejen na základě polohy okolních BTS stanic, ale také podle připojení k WiFi sítím a zejména také pomocí GPS (pokud je v zařízení tato možnost povolena). Díky tomu je možné z dat lépe vyhodnocovat, v jakém dopravním prostředku se daná osoba pohybovala, zda byla statická v jedné lokalitě a jaká byla přesnost zaměřené polohy. Data využívá především společnost Google, nicméně každý uživatel má možnost si svá data nejen zobrazit (Obr. 1), ale také stáhnout ve formátu JSON a dále zpracovávat v prostředí GIS a vizualizovat (Obr. 2). Vhodné jsou např. volně dostupné aplikace <https://locationhistoryvisualizer.com> (Obr. 3) nebo <https://theyhaveyour.info>.

Na základě těchto dat je tak možné mnohem přesněji identifikovat polohu uživatele a následné analýzy jsou mnohem přesnější, a to jak polohově, tak i konextuálně. Data je možné využít např. pro identifikace nejčastějších tras (Löchtfeld, 2019) pro analýzy pohybu obyvatel (Ruktanonchai, Ruktanonchai, Floyd, & Tatem, 2018), vyhodnocování prostorových vzorů pohybu (Rybníkář, 2019) nebo analýzy návštěvnosti (Romero, 2019).

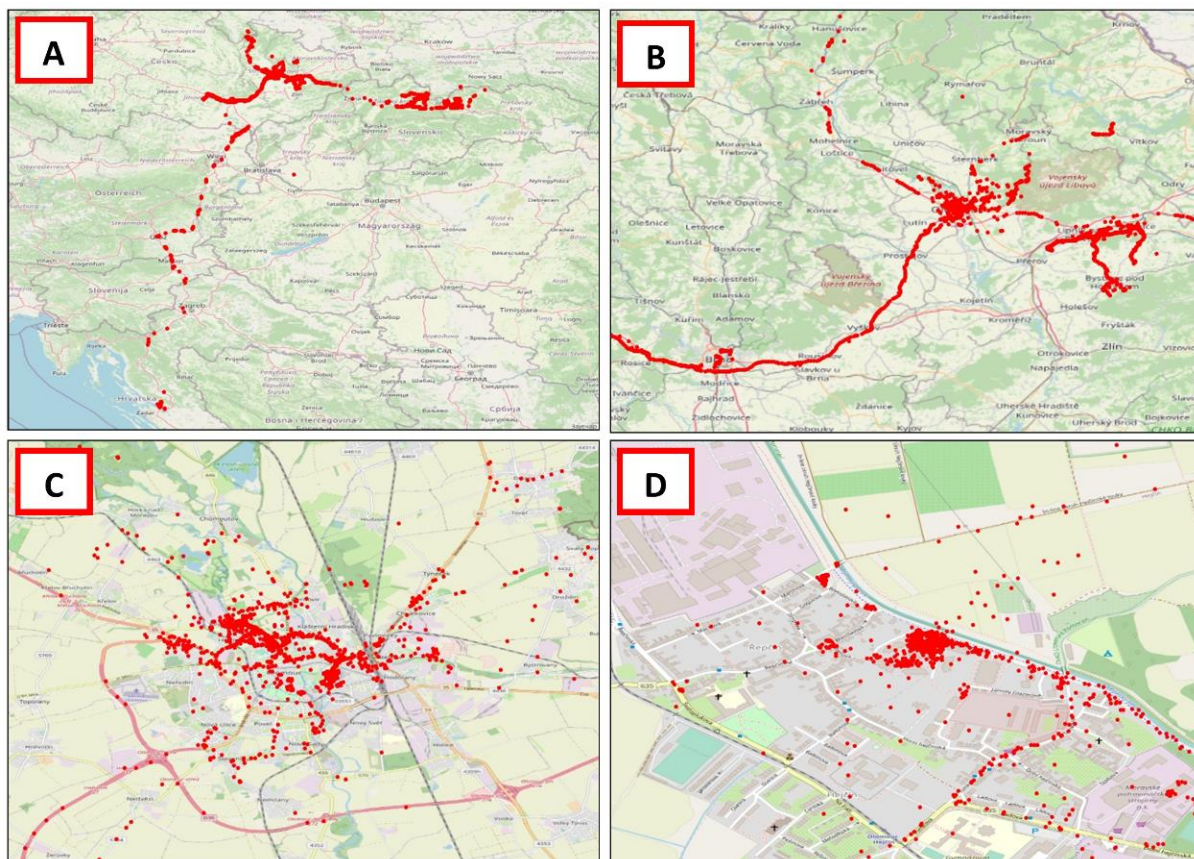
Pro tuto studii byla využita data ze služby Google Location od tří autorů tohoto článku. Vyhodnocován byl celý rok 2020, a to ve fázích, které odpovídají nejzásadnějším změnám v chování dotčených osob v souvislosti s nejrůznějšími restriktivními opatřeními při řešení problémů s nemocí COVID-19. Článek navazuje na předchozí studii (Pászto, Burian, & Macků, 2021), která však byla zaměřena pouze na první vlnu COVID-19 na začátku roku 2020. Cílem článku je na základě dat tří osob ze služby Google Location detailně analyzovat jejich prostorové chování na území města Olomouce v souvislosti s pandemií COVID-19 a představit tak možnosti využití těchto dat.

**Obrázek 1: Základní vizualizace data Google Location**



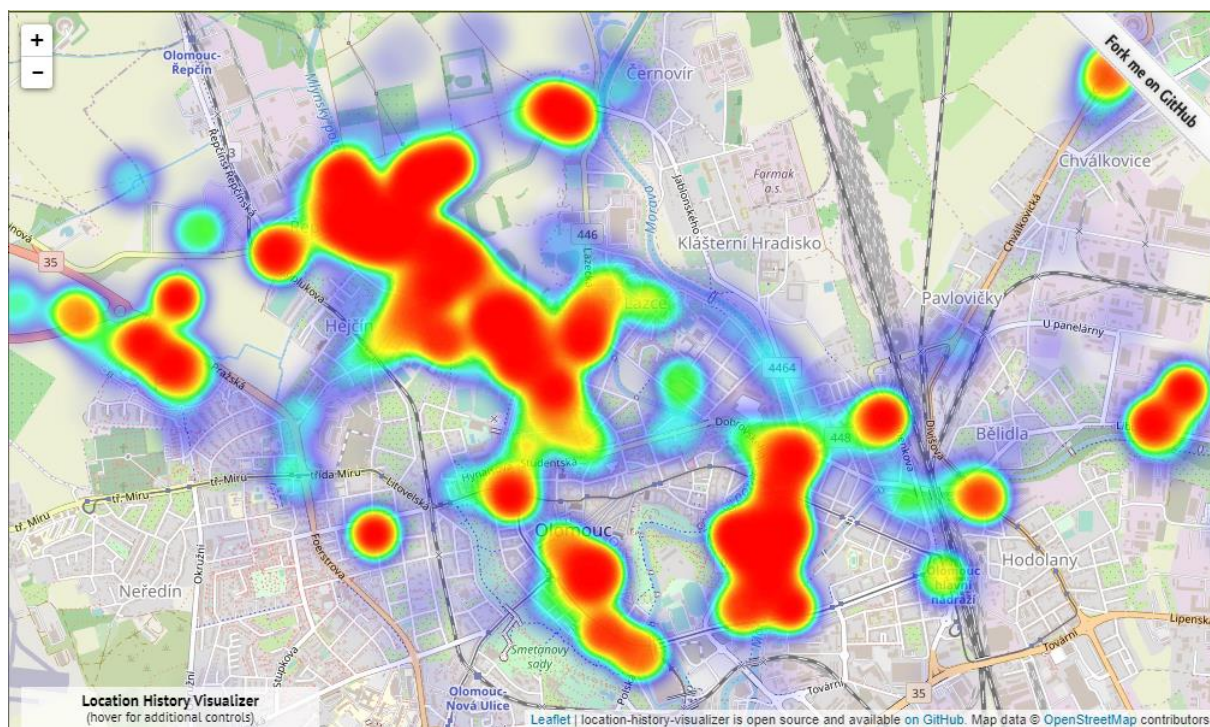
Zdroj: vlastní zpracování

Obrázek 2: Vizualizace data Google Location v prostředí GIS



Zdroj: vlastní zpracování

Obrázek 3: Vizualizace dat Google Location formou intenzity výskytu (metoda heat mapy)



Zdroj: vlastní zpracování

## 2 Data a jejich zpracování

Pro detailní analýzu jsou jednotliví autoři dále označováni jako osoba A, B a C za účelem částečné anonymizace. Po stažení byla data z formátu JSON převedena pomocí volně dostupného skriptu (Boates, 2020) do formátu GeoJSON a následně importována do prostředí Esri geodatabáze. S ohledem na velký objem a časový rozsah dat bylo nutné data dále upravit pomocí základní filtrace dat, eliminací polohových chyb, a především výběrem časového období pro následnou analýzu. Dále byla z dat vybrána pouze část pokrývající území města Olomouce. Analyzovány nebyly jiné cesty nebo pobyty spojené s prací nebo volným časem v jiných zemích, městech nebo přírodě. Detailně byl studován celý rok 2020, který byl rozdělen do sedmi fází (Tab. 1) představujících období s výrazně odlišnými podmínkami ve společnosti (především vlivem vládních nařízení) a lze tedy předpokládat výrazné změny v chování sledovaných osob. První období (Standard 1) představuje stav, kdy nebyly na území ČR stanoveny žádné restriktce a sledované osoby se tedy chovaly obvyklým způsobem. Následuje období postižené velmi tvrdými opatřeními (Lockdown 1), která vedla do té doby k bezprecedentním změnám v životech lidí – například uzavření škol, provozoven maloobchodu a výraznému nárůstu práce z domova. Další fáze (Rozvolňování 1) postihuje období s postupným návratem k běžnému životu, která přechází v období letních dovolených (Léto) a v období začátku nového školního roku s opětovným běžným životním stylem (Standard 2). Závěr roku je rozdělen do fáze opětovného zpřísnění pravidel (Lockdown 2) a do závěrečné fáze, kdy docházelo k rozvolňování a opětovnému zpřísnění.

Data byla za jednotlivá studovaná období analyticky zpracována v prostředí GIS pomocí agregace do pravidelné hexagonální sítě. V dřívější studii (Pászto et al., 2021) bylo testováno několik metod (např. hot spot analýza nebo analýza jádrových odhadů - kernel density), avšak pro věrohodné vyhodnocení změn v chování a zachování maximální jednoduchosti pro interpretaci dat je plně dostačující agregace bodů do hexagonální sítě. Tento tvar byl v minulosti celou řadou autorů potvrzen jako nejvhodnější pro agregaci bodových dat (blíže např. Birch, Oom, & Beecham, 2007; Burian, Stachova, & Vondrakova, 2018). Hexagonální síť byla generována v prostředí ArcGIS Pro pomocí nástroje Bins, který generuje hexagony z bodové vrstvy umístěné v databázi na serveru takzvaně on-the-fly dle velikosti zobrazeného území a úrovně přiblížení (měřítko mapy). Výsledné hexagony byly následně vizualizovány do map uvedených v kapitole výsledky.

Pro přesnější kvantifikaci změn byly každou osobou definovány vybrané body zájmu (bydliště, zaměstnání, nákupy a volný čas – pouze indoor aktivity). Kolem těchto bodů byly vygenerovány kružnice (buffery) o poloměru 300 m a následně byl vyhodnocen rozdíl v počtu zaznamenaných bodů polohy uvnitř kružnic mezi sledovanými obdobími. Poloměr byl zvolen s ohledem na polohovou přesnost vstupních dat, která v případě nepřesného měření (určení polohy pomocí BTS stanice) dosahuje výraznějších odchylek. Z důvodu rozdílného počtu dnů v jednotlivých obdobích a rozdílného počtu bodů zaznamenaných u jednotlivých osob (vlivem různé frekvence záznamu polohy) byly počty zaznamenaných bodů nejprve normovány na srovnatelné období (1 den) a teprve potom porovnávány.

**Tabulka 1: Fáze roku 2020 v souvislosti s COVID-19**

Fáze	Označení	Popis
1. 1. - 10. 3.	Standard 1	Běžný stav
11. 3. - 24. 5.	Lockdown 1	Nouzový stav, zavření škol, home office
25. 5. - 30. 6.	Rozvolňování 1	Rozvolňování opatření, postupný návrat do práce
1. 7. - 31. 8.	Léto	Léto, dovolená
1. 9. - 11. 10.	Standard 2	Běžný stav
12. 10. - 18. 11.	Lockdown 2	Nouzový stav, zavření škol, home office
19. 11. - 31. 12.	Rozvolňování / Lockdown 3	Rozvolňování opatření – otevření škol, restaurací a obchodů, následně opět zpřísnění a Vánoce

Zdroj: vlastní zpracování

### 3 Výsledky

Následující popis se nejprve zaměřuje na vybrané prostorové aspekty změn v chování sledovaných osob za účelem poukázat na velmi vysokou podrobnost možného vyhodnocení. Na rozdíl od studií, které pracují s anonymizovanými a často také agregovanými daty je možné vést interpretaci velmi konkrétním a detailním způsobem. Ačkoliv dále uvedená zjištění velmi dobře korespondují s obecnými a předpokládanými trendy, jedná se o unikátní způsob hodnocení dat, který názorně ukazuje nejen možnosti a také rizika využití (a případného zneužití) dat tohoto typu.

Všechny tři sledované osoby jsou akademickými pracovníky a pracují v širším centru města. Pracoviště je na obrázcích 4, 5 a 6 reprezentováno hexagonem s nejvyššími intenzitami výskytu, vždy uprostřed mapy. Osoba A má bydliště lokalizované v jihozápadní části města, v blízkosti se vyskytuje také nejčastěji navštěvovaný obchod, tři nejčastější místa zábavy jsou rozložena po celém městě. Osoba B má bydliště umístěné v těsné blízkosti nejčastější místa pro nakupování, které je v pěší dostupnosti. Jako bod zábavy navštěvuje akvapark, který byl však navštěvován méně často ve všech sledovaných obdobích. Osoba C má bydliště lokalizované v severozápadní části města, centra zábavy ve východní části a navštěvuje 3 různé obchody pro nákup potravin.

Z vizualizace všech zaznamenaných bodů všech sledovaných osob na obr. 4, 5 a 6 je patrná výrazná změna prostorové distribuce pohybu osob mezi sledovanými obdobími. Nejvyšší hodnoty koncentrace počtu bodů (výskytů osobu) jsou dle očekávání patrné zejména v lokalitách pracoviště a bydliště, přičemž lze pozorovat významný pokles aktivity v místě pracoviště v obdobích označených jako *Lockdown 1* a *Lockdown 2*. Nejvýznamnější je tento pokles u osoby C, která se v období *Lockdown 1* vyskytovala v místě pracoviště pouze jedenkrát a naopak rozšířila časový i prostorový rozsah svého výskytu v oblasti svého bydliště (včetně rodinných vycházek nebo outdoorového sportu v okolí bydliště). Oproti tomu osoba B vykazuje pokles v období *Lockdown 1* v menší intenzitě, neboť i v tomto sledovaném období docházela do místa pracoviště alespoň dvakrát týdně, což je dáno odlišnou rodinnou situací od osoby C. Ve všech případech jde však o velmi signifikantní pokles výskytu v místě pracoviště. Oproti tomu v průběhu *Lockdownu 2* a *Lockdownu 3* takto patrný pokles není zřejmý, neboť všechny osoby i v tomto období docházely prakticky denně do práce (s výjimkou osoby C v období *Lockdown 2*). Opět, osoby A a B mají obdobnou rodinnou situaci (dítě v mladém věku), takže vzhledem k stále selektivně otevřeným školským zařízením (mateřská škola a první stupeň základní školy) je oproti období *Lockdown 1* nelimitováno uzavření školských zařízení. Navíc, jakožto akademičtí pracovníci, spíše využívali technické vybavení pracoviště pro online výuku.

U všech osob je v obdobích *Rozvolňování 1*, *Léto* a *Standard 2* velmi dobře patrné navýšení intenzity výskytu i v dalších lokalitách ve městě, tedy zejména v nákupních centrech, případně dalších lokalitách spojených s volným časem. Nejvíce patrné je to zejména v letním období, kdy lze chování všech osob považovat za obvyklé, prakticky bez jakýchkoliv omezení. V datech zobrazených na obr. 4, 5 a 6 jsou pochopitelně zobrazeny výskyty pouze na území města Olomouce a nejsou tedy plně podchyceny všechny aspekty chování spojené např. s dovolenými a volným časem. Obzvláště v období *Léto* je vidět jednak vyšší prostorový rozptyl výskytu sledovaných osob, ale hlavně je zřejmá vyšší intenzita navštěvování více lokalit ve městě.

Osoba A vykazuje v obdobích *Lockdown 1–3* výrazný úbytek pohybu po městě, zúžený především na lokalitu bydliště a pracoviště, přičemž nejznatelnější je tento úbytek právě v nejpřísnějším období *Lockdown 1*. Kromě bydliště a pracoviště zvýrazňuje prostorová vizualizace v tomto období také další hexagony, které pokrývají místa nákupu potravin a nebo místa trasy přepravy mezi zaměstnáními a bydlištěm. Dalším dobře pozorovatelným jevem je vypuštění pohybu v místě pravidelného trávení volného času (v tomto případě osoby A spojené s pravidelným setkáváním větší skupiny lidí), které je zřetelně patrné např. mezi fázemi *Standard 1 – Lockdown 1* a *Standard 2 – Lockdown 2* (jedná se o lokalitu poblíž vlakového nádraží). Zvýšená prostorová aktivita je pak dle očekávání zachycena v obdobích *Léto* a *Standard 2*, kdy se osoba pochybovala bez omezení po místech trávení volného

času, hlavních lokalitách nákupních středisek anebo plošně více rozprostřených lokalitách trávení času v parcích a přírodě.

Osoba B v období Standard 1 (před pandemií) měla svůj běžný prostorový vzor rozmístěný víceméně rovnoměrně mezi svým pracovištěm a bydlištěm, které zároveň bylo místem nákupů. Z obrázku 5 je patrné, že se jednalo o lokality v centru města nebo blízko něj. Navíc je zde zastoupeno také místo aktivního trávení volného času a bydliště rodinných příslušníků. Při pohledu na stejně klidná období pandemie (zejména Léto), je vidět další výskyt osoby B na jiných lokalitách Olomouce, což je spojeno s vyšší volnočasovou aktivitou i vzhledem k příznivému počasí a také vzhledem k frekventovanějšímu setkávání v rámci rodiny. Paradoxně v období Lockdown 2 bylo u osoby B daleko více koncentrované ve srovnání s Lockdownem 1 a její výskyt se omezoval pouze v oblasti bydliště, pracoviště a místa pobytu nejbližší rodiny. A to vše i přesto, že zejména opatření v oblasti školství nebyla tak limitující jako v období Lockdown 1. Prostorový vzor intenzity výskytu osoby C je velmi výrazný zejména ve všech třech obdobích *Lockdownu*. Výskyt je omezen téměř výhradně na bydliště a pracoviště s naprostým minimem návštěv obchodu či jiných lokalit. První *Lockdown* je navíc charakteristický výrazně nižším výskytem na pracovišti a vyšším podílem práce z domova. Ve druhém a třetím *Lockdownu* je kromě domova a pracoviště patrný ještě vyšší výskyt v jižní části města, které reprezentuje místo volného času, které bylo možné navštěvovat i v období restrikcí. Druhý typ prostorového vzoru je patrný v rámci období *Rozvolňování 1* a *Standard 2*, kdy jsou místa bydliště a pracoviště doplněna dalšími lokalitami, které reprezentují především nákupní centra a místa trávení volného času v rámci města. Tento vzor je z prostorového pohledu rozšířen o další častěji navštívené lokality v průběhu léta (zejména restaurace, parky a obchody).

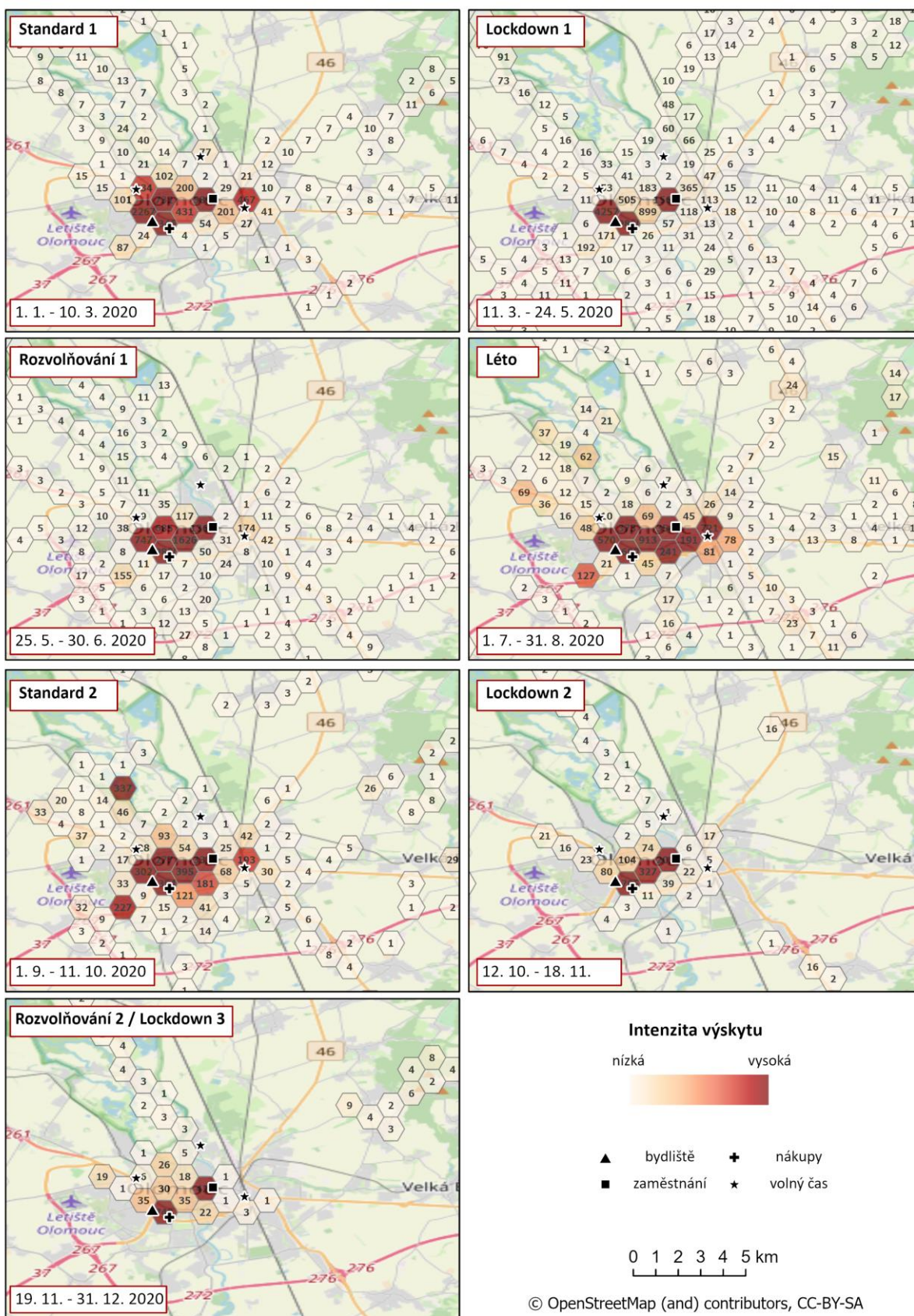
Výsledky je možné interpretovat také na základě kvantitativní analýzy, která byla zpracována pomocí počtu bodů v obalových zónách (buffer zónách) kolem zvolených bodů zájmu (bydliště, nákupy, volný čas a zaměstnání). Výsledky jsou zobrazeny v grafu na obr. 7 a dále ve dvou tabulkách. Tabulka 2 zobrazuje počty zaznamenaných bodů ve sledovaných lokalitách normované dle počtu dnů ve sledovaném období. V tabulce 3 je potom zobrazena procentuální změna ve výskytu osob ve sledovaných lokalitách oproti výchozímu období (Standard 1).

Z dat je velmi dobře patrný výrazný nárůst pobytu v místě bydliště v obou obdobích lockdownu, přičemž v prvním z nich je tento nárůst výraznější. Výrazně vyšší nárůst výskytu v místě bydliště u osoby A je způsobem nejen skutečným častějším výskytem osoby, ale také použitým zařízením, které v místě bydliště generuje větší počet zaznamenaných bodů než v případě jiných lokalit. V období rozvolňování a také v letním období je naopak velmi dobře patrný pokles výskytu v místě bydliště.

Opačná je situace v případě pracoviště, kde je v období obou lockdownů patrný pokles výskytu, s výjimkou osoby A, která v prvním období musela častěji pracovat. Období rozvolňování je naopak charakteristické zvýšeným výskytem v místě pracoviště s výjimkou letního období dovolených. V posledním sledovaném období je patrné odlišné chování sledovaných osob, kdy každá reagovala na rozvolňování a třetí lockdown jiným způsobem.

Hůře interpretovatelné jsou změny v nakupování, což je způsobeno odlišnou frekvencí v nakupování jednotlivých osob. Obecně lze říci, že sledované osoby nenakupují tolik často, aby bylo možné souhrnně hodnotit jejich chování. V případě volného času je velmi dobře čitelný výrazný pokles aktivit jak v prvním, tak ve druhém sledovaném období lockdownu. Ve fázích rozvolňování je zřejmý postupný návrat k původním aktivitám. Analýza však u žádné z osob nepostihuje venkovní volnočasové aktivity, které je však obtížné pomocí hodnocených dat analyzovat.

Obrázek 4: Intenzita výskytu osoby A



Zdroj: vlastní zpracování

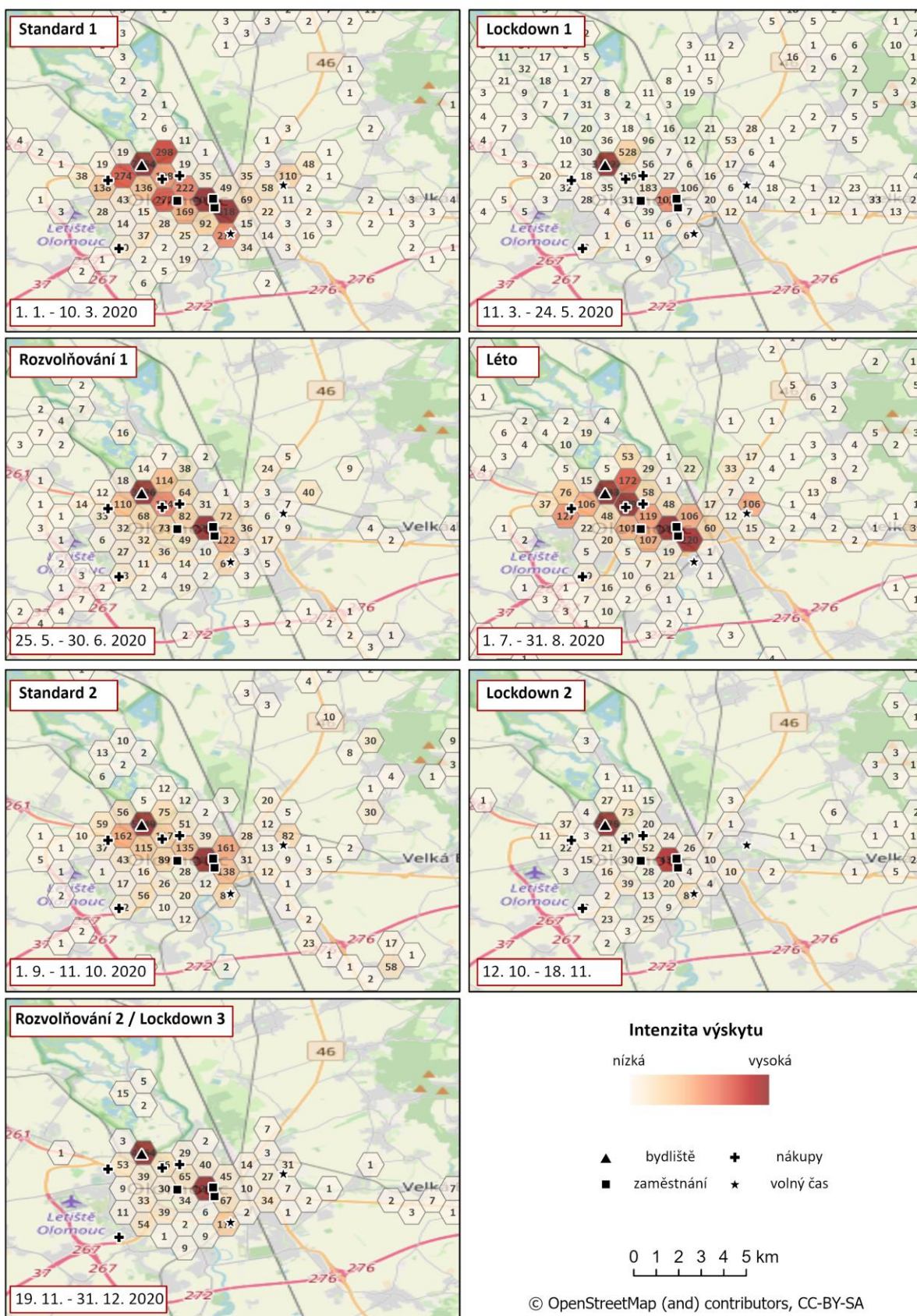


Obrázek 5: Intenzita výskytu osoby B



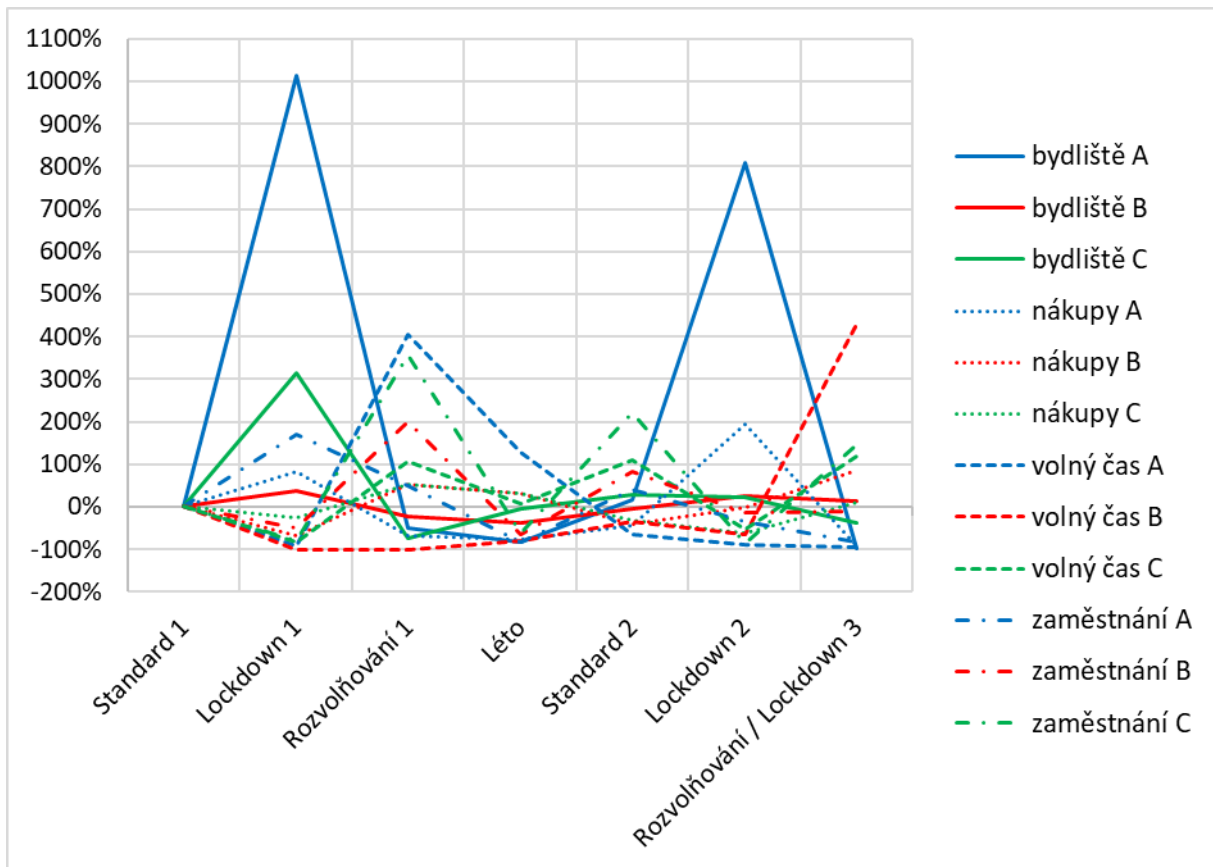
Zdroj: vlastní zpracování

**Obrázek 6: Intenzita výskytu osoby C**



Zdroj: vlastní zpracování

Obrázek 7: Změna ve výskytu osob A, B a C ve sledovaných obdobích



Zdroj: vlastní zpracování

**Tabulka 2: Počet zaznamenaných bodů s polohou ve sledovaných obdobích a lokalitách normovaný dle počtu dnů**

Fáze	Označení	Bydliště			Nákupy			Volný čas			Zaměstnání			Počet dnů
		A	B	C	A	B	C	A	B	C	A	B	C	
1. 1. - 10. 3.	Standard 1	99	172	99	5	10	5	11	1	5	75	68	69	69
11. 3. - 24. 5.	Lockdown 1	1107	236	410	9	3	3	1	0	1	202	35	14	75
25. 5. - 30. 6.	Rozvolňování 1	544	178	102	3	5	5	5	2	2	303	103	65	37
1. 7. - 31. 8.	Léto	94	110	98	1	6	7	11	0	2	52	35	25	62
1. 9. - 11. 10.	Standard 2	109	105	127	0	4	4	4	0	4	72	63	79	41
12. 10. - 18. 11.	Lockdown 2	992	130	155	1	4	2	0	0	2	47	54	12	38
19. 11. - 31. 12.	Rozvolňování / Lockdown 3	18	146	97	0	7	2	0	1	4	9	48	30	43

Zdroj: vlastní zpracování

**Tabulka 3: Změna ve výskytu osob ve sledovaných lokalitách oproti výchozímu období**

Fáze	Označení	Bydliště			Nákupy			Volný čas			Zaměstnání		
		A	B	C	A	B	C	A	B	C	A	B	C
1. 1. - 10. 3.	Standard 1	0 %	0 %	0 %	0 %	0 %	0 %	0 %	0 %	0 %	0 %	0 %	0 %
11. 3. - 24. 5.	Lockdown 1	1015 %	37 %	314 %	82 %	-69 %	-27 %	-91 %	-100 %	-82 %	170 %	-50 %	-79 %
25. 5. - 30. 6.	Rozvolňování 1	-51 %	-24 %	-75 %	-67 %	53 %	54 %	404 %	-100 %	106 %	50 %	199 %	356 %
1. 7. - 31. 8.	Léto	-83 %	-39 %	-4 %	-78 %	31 %	31 %	128 %	-80 %	7 %	-83 %	-66 %	-62 %
1. 9. - 11. 10.	Standard 2	16 %	-5 %	30 %	-45 %	-41 %	-33 %	-65 %	-35 %	109 %	39 %	82 %	221 %
12. 10. - 18. 11.	Lockdown 2	808 %	25 %	22 %	195 %	0 %	-61 %	-90 %	-64 %	-54 %	-35 %	-15 %	-85 %
19. 11. - 31. 12.	Rozvolňování / Lockdown 3	-98 %	12 %	-37 %	-94 %	87 %	11 %	-94 %	430 %	118 %	-82 %	-11 %	145 %

Zdroj: vlastní zpracování

#### 4 Diskuse

Výsledky popisované v předchozí kapitole poměrně přesně a věrohodně popisují skutečné chování sledovaných osob. Oproti běžným přístupům ke sběru dat o pohybu obyvatel se jedná o ukázkou velmi personalizovaného přístupu, který umožňuje retrospektivní časoprostorové analýzy vhodné pro identifikaci prostorových vzorů a změn v mobilitě i jiném chování osob. Ačkoliv se jedná o velmi citlivá data osobního charakteru, tak v případě jejich anonymizace a alespoň částečné agregace, by mohla být do budoucna významným a kvalitativně hodnotným zdrojem informací. Tato data jsou ostatně v agregované podobě poskytována v prostřednictvím služby Google Mobility Reports za administrativní celky (blíže např. Pászto, Burian, & Macků, 2020). Nevýhodou však je poměrně omezená podoba dat z pohledu atributů, kdy je k dispozici pouze srovnání vůči referenčnímu období za vybrané typy zájmových míst (např. domov, parky, obchod) bez podrobnějších informací např. o počtu zahrnutých osob. Z těchto dat tedy nelze ani usuzovat o pohybu jednotlivých osob nebo skupin mezi více lokalitami. Data použitá od tří osob v rámci této studie jsou do jisté míry odlišná. Každý z uživatelů disponoval jiným zařízením (jiný mobilní telefon) a data tak byla sbírána s různou úrovní přesnosti a také s různou frekvencí (vlivem různým metod určení polohy). Z tohoto důvodu může být analýza částečně zkreslující, a proto byl kladen důraz především na interpretaci jednotlivých osob a nikoliv na souhrnné hodnocení chování celé skupiny.

Důležitým aspektem provedené analýzy je také ochrana osobních údajů. Z vizualizace a analýzy použitých dat je velmi dobře patrná celá řada osobních informací, např. místo bydliště, pracoviště, doba pobytu v jednotlivých lokalitách, trasy přesunu, atd. Takovéto informace jsou právem většinou uživatelů vnímány jako velmi osobní a citlivé, a takto je k nim také nutno přistupovat. Pokud by takováto data byla běžně využívána za individuální osoby, jednalo by se nepochybně o výrazné zásahy do soukromí. Proto je nutné v případě práce s takovými daty přistoupit nejen k anonymizaci, ale také k agregaci dat za větší skupinu obyvatel, ze které již nebude možné usuzovat o chování jednotlivce.

Výsledky námi představené analýzy jsou částečně ovlivněny přesností určení polohy pomocí BTS stanic nebo GPS. Značná část vstupních dat ze služby Google Location je nepřesných, a je tedy nutná poměrně náročná filtrace a jejich zpracování z hlediska jejich konečné validity. Část zobrazených dat je ovlivněna chybným, respektive nepřesným určením polohy, například vlivem nižšího počtu viditelných BTS stanic, a tedy nižší úrovní přesnosti polohy. Na druhou stranu agregace do hexagonální sítě tyto nedostatky stírá a problémové je "pouze" zařazení dat na hranici hexagonů. Velikost hexagonů, případně volba jiného pravidelného gridu pro přepočítání dat může také do jisté míry ovlivňovat dosažené výsledky. Posledním aspektem, který ovlivňuje výsledky je omezený prostorový výběr analýzy na město Olomouc. V datech od jednotlivých osob tak nejsou zahrnuty další aktivity mimo Olomouc, které zejména v letním období a víkendech zahrnují větší množství bodů. I přes tyto zmíněné nedostatky lze však možné výsledky (zejména v rámci města) považovat za velmi věrohodné a relevantní.

Ve srovnání s daty, kterými disponují mobilní operátoři, jsou data Google Location kvalitativně výrazně lepší, zejména s ohledem na polohovou přesnost, která je u těchto dat vyšší než při určování polohy pouze na základě interpolace z BTS stanic. Určení polohy v rámci služby Google Location využívá kromě signálu z BTS stanic také GPS signál a informace o připojení k internetu (zejména Wifi). Principiálně je tedy prakticky vždy poloha určena s mnohem vyšší přesností. V obecné rovině mají však i samotná data od mobilních operátorů velký význam. Vlastnictví mobilního telefonu je ve většině zemí považováno takřka za samozřejmé, a proto mohou data mobilních operátorů sloužit jako velmi relevantní zdroj pro identifikaci skutečné polohy obyvatel v čase. Problémem těchto dat je však jejich obtížná dostupnost, ať už z důvodu ceny nebo ochrany osobních údajů. Možnosti zpracování dat a jejich anonymizace je však již na takové úrovni, která nebrání jejich intenzivnějšímu využívání. Je tedy velmi žádoucí tento zdroj dat i v budoucnu využívat, neboť by mohl nahradit část dat, která jsou dnes pořizována velmi nákladně, nepravidelně, nebo neexistují vůbec.

## 5 Závěr

Výše zmíněné analýzy a jejich interpretace přesně a věrohodně popisují skutečné chování sledovaných osob. Oproti běžným přístupům ke sběru dat o pohybu obyvatel se jedná o ukázkou velmi personalizovaného přístupu, který umožňuje retrospektivní časoprostorové analýzy vhodné pro identifikaci prostorových vzorů a změn v mobilitě i jiném chování osob. Ačkoliv se jedná o velmi citlivá data osobního charakteru, tak v případě jejich anonymizace a alespoň částečné agregace, by mohla být do budoucna významným a kvalitativně hodnotným zdrojem informací.

Využití detailních a individuálních dat pro analýzu a hodnocení chování uživatelů v krizové situaci pandemie COVID-19 je velmi jedinečný přístup, který dosud nebyl v takovémto rozsahu (pokryt celý jeden rok) nikde publikován. Jsme si plně vědomi toho, že z těchto dat jsou patrné mnohé další informace o prostorovém chování jednotlivých osob a jde současně o osobní údaje, které je velmi důležité chránit. Je vhodné upozornit na fakt, že taková data (v případě našeho souhlasu) o nás společnost Google uchovává a dále využívá. O to důležitější je veřejnosti představit o jaká data jde a k jakým účelům mohou být využívána. Současná situace ve světě názorně ukazuje, že polohová a individuální data mohou být prospěšná také v případě pandemií nebo jakýchkoliv jiných krizových situacích, kdy se jedná o bezpečnost a zdraví obyvatel. Současně je však nutné k těmto datům přistupovat velmi citlivě a důsledně dbát ochrany osobních údajů.

## 6 Literatura

- Batty, M. (2013). Big data, smart cities and city planning. *Dialogues in Human Geography*, 3(3), 274–279. <https://doi.org/10.1177/2043820613513390>
- Batty, M. (2016). Big data and the city. *Built Environment*, 42(3), 321–337. <https://doi.org/10.2148/benv.42.3.321>
- Birch, C. P. D., Oom, S. P., & Beecham, J. A. (2007). Rectangular and hexagonal grids used for observation, experiment and simulation in ecology. *Ecological Modelling*, 206(3–4), 347–359. <https://doi.org/10.1016/j.ecolmodel.2007.03.041>
- Boates, I. (2020). *Android\_location\_converter*.
- Burian, J., Stachova, M., & Vondrakova, A. (2018). Land suitability assessment of the Olomouc region: an application of an Urban Planner model. *Journal of Maps*, 14(1), 73–80. <https://doi.org/10.1080/17445647.2018.1493407>
- Calabrese, F., Diao, M., Di Lorenzo, G., Ferreira, J., & Ratti, C. (2013). Understanding individual mobility patterns from urban sensing data: A mobile phone trace example. *Transportation Research Part C: Emerging Technologies*, 26, 301–313. <https://doi.org/10.1016/j.trc.2012.09.009>
- Castro, P. S., & Zhang, D. (2013). From Taxi GPS Traces to Social and Community Dynamics: A Survey. *ACM Computing Surveys*, 46(2), 34. <https://doi.org/10.1145/2543581.2543584>
- Cox, M., & Ellsworth, D. (1997). *Application-Controlled Demand Paging for Out-of-Core Visualization*. Získáno z <https://www.nas.nasa.gov/assets/pdf/techreports/1997/nas-97-010.pdf>
- Čtyrský, J., Novotný, V., & Soukup, M. (2016). Big Data Praha: Využití lokalizačních dat mobilních operátorů v hl. m Praze. *Sborník příspěvků Symposium GIS Ostrava 2016 – Geoinformatika pro společnost*. Ostrava.
- Frothingham, S. (2014). Will Strava Metro turn data into profit? (cover story) (Roč. 23). <https://doi.org/10.1038/2201000c0>
- Hao, J., Zhu, J., & Zhong, R. (2015). The rise of big data on urban studies and planning practices in China: Review and open research issues. *Journal of Urban Management*, 4(2), 92–124. <https://doi.org/10.1016/j.jum.2015.11.002>
- Hawelka, B., Sitko, I., Beinart, E., Sobolevsky, S., Kazakopoulos, P., & Ratti, C. (2014). Geo-located Twitter as proxy for global mobility patterns. *Cartography and Geographic Information Science*, 41(3), 260–271. <https://doi.org/10.1080/15230406.2014.890072>

- Haworth, J. (2016). Investigating the potential of activity tracking app data to estimate cycle flows in urban areas. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 41, 515–519. <https://doi.org/10.5194/isprsarchives-XLI-B2-515-2016>
- Hochmair, H. H., Bardin, E., & Ahmouda, A. (2017). Estimating Bicycle Trip Volume for Miami-Dade County from Strava Tracking Data. *Transportation Research Board 96th Annual Meeting*, 1–17.
- Kahrik, A., Novák, J., Temelová, J., Kadarik, K., & Tammaru, T. (2015). Patterns and Drivers of Inner City Social Differentiation in Prague and Tallinn. *Geografie*, 120(2), 275–295. Získáno z [https://apps.webofknowledge.com/full\\_record.do?product=UA&search\\_mode=GeneralSearch&qid=6&SID=V2DKdt9s61tiunvCHKq&page=1&doc=1](https://apps.webofknowledge.com/full_record.do?product=UA&search_mode=GeneralSearch&qid=6&SID=V2DKdt9s61tiunvCHKq&page=1&doc=1)
- Kocich, D. (2018). *Multilingual Sentiment Mapping Using Twitter, Open Source Tools, and Dictionary Based Machine Translation Approach*. [https://doi.org/10.1007/978-3-319-61297-3\\_16](https://doi.org/10.1007/978-3-319-61297-3_16)
- Kurian, S., Bhatti, A., Ting, H., Storlie, C., Shah, N., & Bydon, M. (2020). Utilization of Mobility Data in the Fight Against COVID-19. *Mayo Clinic proceedings. Innovations, quality & outcomes*, 4(6), 733–735. <https://doi.org/10.1016/J.MAYOCPIQO.2020.10.003>
- Löchtefeld, M. (2019). Detournavigator - Using google location history to generate unfamiliar personal routes. *Conference on Human Factors in Computing Systems - Proceedings*, 1–6. <https://doi.org/10.1145/3290607.3313047>
- Long, X., Jin, L., & Joshi, J. (2012). Exploring trajectory-driven local geographic topics in foursquare. *Proceedings of the 2012 ACM Conference on Ubiquitous Computing - UbiComp '12*, 927. <https://doi.org/10.1145/2370216.2370423>
- Musakwa, W., & Selala, K. M. (2016). Mapping cycling patterns and trends using Strava Metro data in the city of Johannesburg, South Africa. *Data in Brief*, 9, 898–905. <https://doi.org/10.1016/j.dib.2016.11.002>
- Noulas, A., Scellato, S., Lambiotte, R., Pontil, M., & Mascolo, C. (2012). A tale of many cities: Universal patterns in human urban mobility. *PLoS ONE*, 7(5). <https://doi.org/10.1371/journal.pone.0037027>
- Noulas, A., Scellato, S., Mascolo, C., & Pontil, M. (2011). An Empirical Study of Geographic User Activity Patterns in Foursquare. *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*, (January), 570–573. <https://doi.org/papers3://publication/uuid/557455DB-AC4A-4C73-968A-31E7A663BC4E>
- Novák, J., & Temelová, J. (2012). Everyday Life and Spatial Mobility of Young People in Prague: A Pilot Study Using Mobile Phone Location Data. *SOCIOLOGICKY CASOPIS-CZECH SOCIOLOGICAL REVIEW*, 48(5), 911–938.
- Pászto, V., Burian, J., & Macků, K. (2020). COVID-19 data sources : evaluation of map applications and analysis of behavior changes in Europe ' s population. *GEOGRAFIE*, 125(2), 39.
- Pászto, V., Burian, J., & Macků, K. (2021). Changing Mobility Lifestyle : A Case Study on the Impact of COVID-19 Using Personal Google Locations Data. *International Journal of E-Planning Research*, 10(2), 66–79. <https://doi.org/10.4018/IJEPR.20210401.oa6>
- Romanillos, G., Zaltz Austwick, M., Ettema, D., & De Kruijf, J. (2016). Big Data and Cycling. *Transport Reviews*, 36(1), 114–133. <https://doi.org/10.1080/01441647.2015.1084067>
- Romero, F. M. (2019). Creation of images in virtual map location services. The case of Google Maps. *Hipertext.net: Revista Académica sobre Documentación Digital y Comunicación Interactiva*, 0(18), 66–76. <https://doi.org/10.31009/hipertext.net.2019.i18.07>
- Ruktanonchai, N. W., Ruktanonchai, C. W., Floyd, J. R., & Tatem, A. J. (2018). Using Google Location History data to quantify fine-scale human mobility. *International Journal of Health Geographics*, 17(1), 1–13. <https://doi.org/10.1186/s12942-018-0150-z>
- Rybníkář, L. (2019). *Analýza a vizualizace dat služby Google Location*. Univerzita Palackého v Olomouci.
- Sapiezynski, P., Stopczynski, A., Gatej, R., & Lehmann, S. (2015). Tracking human mobility using WiFi signals. *PLoS ONE*, 10(7). <https://doi.org/10.1371/journal.pone.0130824>

- Selala, M. K., & Musakwa, W. (2016). The potential of strava data to contribute in non-motorised transport (NMT) planning in johannesburg. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 41, 587–594. <https://doi.org/10.5194/isprsarchives-XLI-B2-587-2016>
- Skendzic, A., Kovacic, B., & Tijan, E. (2018). General data protection regulation - Protection of personal data in an organisation. *2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics, MIPRO 2018 - Proceedings*, 1370–1375. <https://doi.org/10.23919/MIPRO.2018.8400247>
- Soukup, M. (2017). Co nám prozradí data z mobilů? *ArcRevue, ARCDATA PRAHA*, (2), 10–13.
- Štarchoň, P., & Pikulík, T. (2019). GDPR principles in Data protection encourage pseudonymization through most popular and full-personalized devices - mobile phones. *Procedia Computer Science*, 151, 303–312. <https://doi.org/10.1016/J.PROCS.2019.04.043>
- Sun, Y. (2017). Exploring potential of crowdsourced geographic information in studies of active travel and health: Strava data and cycling behaviour. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 42(2W7), 1357–1361. <https://doi.org/10.5194/isprs-archives-XLII-2-W7-1357-2017>
- Wang, X. (2013). Intelligent multi-camera video surveillance: A review. *Pattern Recognition Letters*, Roč. 34, s. 3–19. <https://doi.org/10.1016/j.patrec.2012.07.005>
- Zděrad, M. (2016). Využití lokalizačních dat mobilních operátorů pro městské plánování. *Sborník příspěvků z konference GIS Esri v ČR*. Praha: ARCDATA PRAHA.