Univerzita Karlova

Filozofická fakulta

Ústav informačních studií a knihovnictví

Studia nových médií

Diplomová práce

Bc. Jan Kryšpín

Vzorce pohybu obyvatel ve městech

Human Urban Mobility Patterns

Praha 2019

Vedoucí práce: Mgr. Josef Šlerka, Ph.D.

Poděkování

Děkuji Mgr. Josefu Šlerkovi, Ph.D. za vedení mé diplomové práce a za dobré rady. Dále děkuji rodině za pomoc a trpělivost během mých dlouhých studií.

Prohlášení:

Prohlašuji, že jsem diplomovou práci vypracoval samostatně, že jsem řádně citoval všechny použité prameny a literaturu a že práce nebyla využita v rámci jiného vysokoškolského studia či k získání jiného nebo stejného titulu.

V Praze, dne 16. prosince 2019

Jan Kryšpín

Klíčová slova (česky)

Mobilita, migrace, vzorce, geolokační sociální sítě, Foursquare, Rekola

Klíčová slova (anglicky):

Mobility, migration, patterns, Location Based Social Networks, Foursquare, Rekola

Abstrakt (česky)

Cílem této diplomové práce je za pomoci moderních nástrojů definovat vzory pohybu obyvatel ve městech a zjistit, co ovlivňuje lidskou migraci. V první polovině rozebírá historii výzkumu a jednotlivé metody a modelování. Druhá polovina práce se věnuje výzkumu dat z aplikace Foursquare a z projektu Rekola. Tyto data zpracovává na základě gravitačního modelu a teorie intervenujících příležitostí.

Abstract (in English):

The aim of this thesis is to define human urban mobility patterns using modern tools and find out what influence human migration. First part deals with history of research and introduces different methods of modeling. Second part consists of own research based on data from application Foursquare and project Rekola. This data is processed using gravitation model and theory of intervening opportunities.

Obsah

1 Úvod 8

2 Mobilita 9

2.1 Historie studia mobility 9

2.2 Typy mobilit 12

2.3 Každodenní mobilita 14

2.4 Virtuální mobilita 15

3 Výzkumy vzorů pohybu 17

3.1 Vzorce pohybu u zvířat 17

3.2 Vzorce pohybu z běžného života 17

3.3 Vzorce pohybu z nových médií 19

3.4 Vzorce pohybu z dotazníků 22

3.5 Shrnutí 23

4 Modelování mobility 26

4.1 Modelování mobility na základě vzdálenosti 26

4.1.1 Gravitační model migrace 26

4.1.2 Lévyho model 27

4.1.3 Gyrační poloměr 28

4.2 Modelování mobility na základě hodnocení 28

4.2.1 Radiační model lidské mobility 28

4.2.2 Teorie intervenujících příležitostí 29

5 Vlastní výzkum 30

5.1 Foursquare 30

5.2 Metoda výzkumu *A Tale of Prague* 31

5.2.1 Výzkum 31

5.2.2 Zpracování dat 33

5.2.3 Městský pohyb a zákon síly 35

5.2.4 Vliv hustoty oblasti....37
5.2.5 Porovnání gravitačního modelu....42
5.3 Metoda shlukové analýzy....43
5.3.1 Arts & Entertainment – Kulturní zábava....45
5.3.2 Travel & Transport – Doprava a cestování....45
5.3.3 Food – Jídlo a restaurace....45
5.3.4 Outdoor & Recreation – Příroda a rekreace....46
5.3.5 Nightlife Spot – Noční život....46
5.3.6 Shop & Service – Obchody a služby....46
5.3.7 Výsledky výzkumu shlukové analýzy....47
5.4 Porovnání výzkumu s daty z aplikace Rekola....47
5.4.1 Rekola....47
5.4.2 Dataset Rekola....48
5.4.3 Výzkum....49
6 Závěr....52
7 Bibliografie....53
8 Seznam obrázků....62
9 Seznam grafů....63
10 Seznam tabulek....64
11 Použité zkratky....65

1 Úvod

Vědecké práce věnující se mobilitě se rozchází ve dvou názorových pólech, a to v tom, jak se lidé rozhodují, jestli dle vzdálenosti míst či dle jejich hodnocení. První skupina, inspirována Newtonovými pohybovými zákony, tvrdí, že lidé vyhodnocují energii a čas strávený pohybem, a tak čím větší je vzdálenost mezi dvěma místy, tím se tok migrace snižuje. Tyty gravitační modely popsal v roce 1950 John Q. Stewart ve své práci *The Development of Social Physics*, ale první definice a zákony určil již v roce 1885 Ernst Georg Ravenstein v *The Laws of Migration*. Na základě těchto zákonů a modelů jsou plánovány například městské a meziměstské komunikace. Druhá skupina tvrdí, že není žádná přímá vazba mezi mobilitou a vzdáleností. Pohyb lidí z jednoho místa na druhé je přímo závislý na počtu alternativních míst blíže, než je místo cílové. Lidé tak vyhledávají místa pouze dle jejich ohodnocení a nehledí na samotnou délku trasy. Ta je tak irrelevantní. Posuny jsou tedy dány prostorovým rozložením míst, tedy spíše nabídkou příležitostí než dopravní vzdáleností jako v gravitačních modelech. První názorový pól je tedy preferován především na základě výpočtové jednoduchosti, a to i navzdory tomu, že několik statistických studií prokázalo, že konceptem intervenčních příležitostí je možné vysvětlit větší škálu mobility. V této je nejprve rozebráno, co je to mobilita. Dále jaké jsou metody hledání vzorů pohybu a jejich modelování. Následuje vlastní výzkum, ve kterém jsou rozebrány dva dataset, jeden z geolokační sociální sítě Foursquare a druhý z bikesharingového projektu Rekola. Cílem výzkumu je najít v datech metodu či teorii, kterou lze jednotlivé pohyby popsat.

2 Mobilita

Mobilita, tedy schopnost lidí přemisťovat se, je klíčovou součástí života. Je jedním z nejdůležitějších znaků moderní společnosti. S rostoucí velikostí měst a vzdáleností mezi bydlištěm a pracovištěm jejich obyvatel roste i městská mobilita. K ní přispěl i technologický vývoj dopravy, který umožnil rychlejší propojení míst a umožnil lidem přesouvat se po větších vzdálenostech každý den. Mobilitu můžeme rozdělit na fyzickou a virtuální. Fyzická mobilita je klasický pohyb lidí a věcí, kdežto virtuální mobilita je ovlivněna informačními technologiemi. (Kellerman, 2006) „Mobilita je nevyhnutelná součást současného života. Jak v hospodářství, tak ve volném čase. Její negativní dopady poznamenávají náš život. Stále více dopravních prostředků ucpává městské ulice. Přeplněné dopravní tepny stejně jako kongesce se stávají součástí každodenního života. Vědci již před delší dobou odhalili vztah mezi hospodářským růstem země a potřebou mobility. Jednoduše řečeno se zvyšováním životní úrovně roste doprava a naopak. Globalizace, zvyšující se internacionalizace hospodářství a integrace středoevropských a východoevropských zemí vytváří nové nároky na výkonnost transportního systému. Aby bylo možné udržet zvyšující se dopravní výkony, jsou nutné změny v dopravní politice, v řízení dopravy i její infrastruktuře. Je také důležité snižovat vstup energií při zvyšujících se přepravních objemech, což se příznivě objeví na životním prostředí.” (Udržitelná mobilita, 2007)

2.1 Historie studia mobility

Pohybem za zabýval již například Aristoteles, Tomáš Akvinský nebo Ibn Rašíd. Přímo mobilitou pak i francouzský filozof Henri Bergson, který zkoumal vztah mezi pohybem a jeho vnímáním. Z jeho děl mimo jiné vyplývá, že vnímání imobility je pouhá iluze, jelikož se ve skutečnosti všechno pohybuje a mění. Tvrdil tedy, že imobilita vůbec neexistuje imobilita. (Adey, 2010)

Růst v popularitě studia mobility se objevil v 19. století. Jedním z prvních průkopníků byl například Ernst Georg Ravenstein, který ve svém díle *The Laws of Migration* popsal 8 zákonů, které se vztahují i k mobilitě:

1. Většina migrantů se pohybuje jen po krátkých vzdálenostech.

2. Existuje proces absorpce, kdy lidé okamžitě obsazují okolí rychle rostoucích měst a vzniklé mezery vyplňují migranti ze vzdálenějších míst, do doby, dokud působí tzv. *pull faktory*.
3. Existuje proces disperze, který je inverzní absorpce.
4. Každý proud migrace vytváří kompenzační protiproud.
5. Migranti ze vzdálených míst míří do největších obchodních a průmyslových center.
6. Lidé z měst migrují méně než lidé z venkovských oblastí.
7. Ženy migrují více než muži.
8. Ekonomické faktory jsou hlavním důvodem migrace. (Ravenstein, 1885)

Dalším významným vědcem v oboru mobility byl Georg Simmel, který ve svých sociologických pracích definoval tzv. nové mobilitní paradigma („*new mobilities paradigm*“, zkráceně „*mobilities*“), ty zpracovali Urry a Sheller. Paradigma přesouvá pozornost od studia míst původu a cílových destinací na významy skryté za pouhými přesuny. (Urry, 2006) Vychází z šesti hlavních východisek:

1. Prvním je fenomén mobility zdůrazňující roli „tempa“ (pulz města, který utváří městský život po stránce sociální, ekonomické i infrastrukturní).
2. Dále vztah lidí a technologií a jejich vliv na fungování společnosti. A tvorba hybridní geografie.
3. Další východisko je, že dochází k mobilizaci „prostorového obratu“ v sociálních vědách.
4. Samotný pohyb a emocionální aspekty pohybu, prostřednictvím nichž si každý vytváří své emocionální geografie.
5. Topologií sociálních sítí, především se slabými vazbami, které mohou generovat „malé světy“ ve zdánlivě nespojených místech.
6. A poslední východisko začleňuje i analýzy komplexních systémů, které jsou v tzv. „spořádaném nepořádku“.

Pro studium každodenní mobility jsou tedy nejvíce relevantní východiska 1, 2 a 4, která zohledňují význam času a emociálních aspektů při pohybu.

Vzestup ve studiu mobility nastal zejména v šedesátých a sedmdesátých letech minulého století. V této době se do popředí dostala geografie dopravy a nejčastějším cílem výzkumu se stalo vytvoření obecně platných teorií a modelů. Šlo především o gravitační modely, teorie prostorových interakcí, lokalizační i dopravní modely a další. Vznikl také pojem *homo*

economicus („*rational-mobile-person*“) čili člověk zakládající své chování a volby čistě na racionálním uvažování. (Cresswell, 2011)

Základy studia denní mobility a každodenního života položil švédský geograf Torsten Hägerstrand. Ten publikoval článek *What about People in Regional Science?* v periodiku *Papers in Regional Science*, ve kterém představil geografii času. (Hägerstrand, 1970) Ta poskytuje rozsáhlý soubor metod a konceptů pro zachycení, popis a interpretaci každodenního života na úrovni jednotlivců a domácností. Zaměřuje se především na vnější, materializovanou sféru akce. Jejím často zmiňovaným nedostatkem je proto její slabost v objasňování způsobů rozhodování jedinců a strategií, které lidé uplatňují při realizaci každodenních povinností a dalších aktivit. (Hallin, 2017)

Geografie času také poskytuje systematický jazyk a metodiku pro popis každodenního prostorového chování a umožňuje zachytit časoprostorovou choreografii lidské existence nejen v každodenním, ale i v celoživotním měřítku sledování. (Thrift, 1977) (Pred, 1977) Dle Pratta se zaměřuje na sledování časování a umístění lidských aktivit v průběhu dne (celého života), na možnosti a problémy, které s sebou časování a umístění přinášejí, a na otázku, nakolik toto poznání může přispět k pochopení společenských vztahů. (Pratt, 2016) Zaměření je především na jednotlivce a jeho samostatné chování. V geografii času lze identifikovat dva vůdčí principy korporeality a kontinuity. (Hagerstrand, 1982)

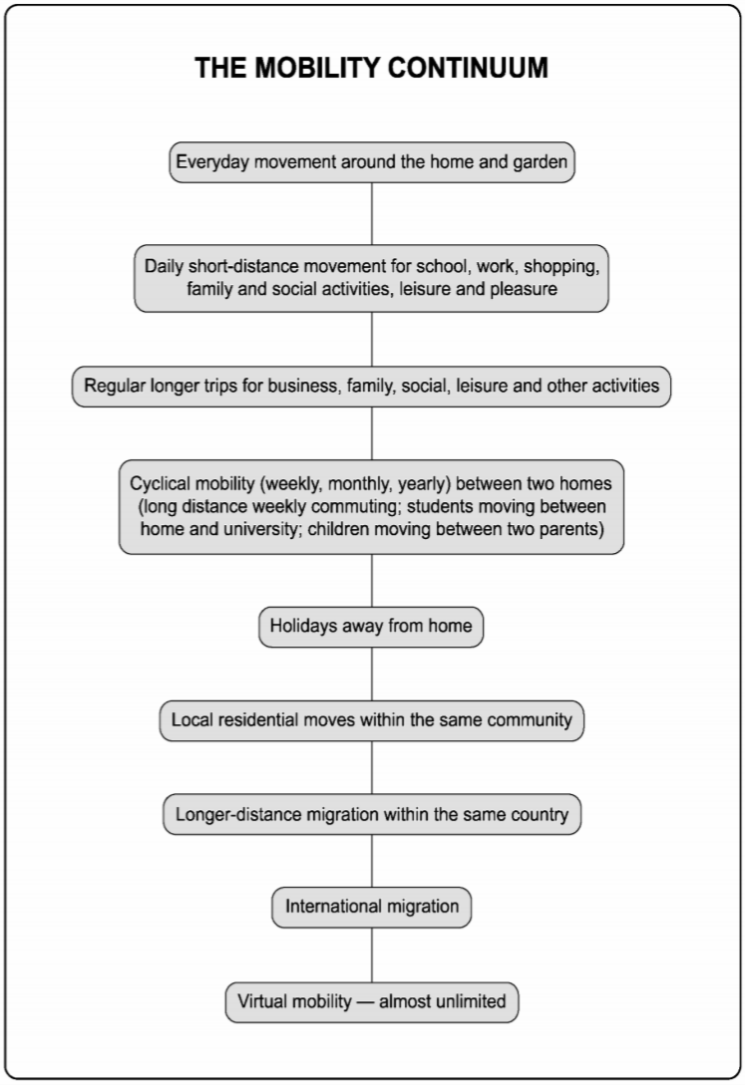
Kontinuitou je myšlené propojení minulosti, současností a budoucnosti. Jejím fundamentálním vyjádřením je koncept cesty. Život člověka si tak můžeme představit jako kontinuální linku v čase a prostoru, kdy současný okamžik je pevně navázán na předchozí událostí a budoucnost je zase pevně navázána na něj. V každý moment této dráhy se lidé aktivně rozhodují co dál, ale jejich rozhodnutí jsou omezena vnějšími přírodními a společenskými podmínkami.

Korporealita se zabývá geografickým a materiálním kontextem, v němž se odehrává aktuální událost. Vyšší prioritu tak má situační povaha lidského jednání, v jejímž rámci se propojuje kontext místa i času (kontinuita a korporealita). Všechny lidské události jsou tak umístěny na konkrétní území a do souvislosti s předcházejícími událostmi. Tento důraz na situační analýzu najdeme v mnoha raných pracích vědců zabývajících se geografií času. (Parkes, 1980) (Hagerstrand, 1982) (Pred, 1981)

2.2 Typy mobilit

Mobilitu lze dělit mnoha způsoby. Jeden z nich je například podle hlediska časového, na dlouhodobé a krátkodobé pohyby a na opakující se, cyklické pohyby, ku příklady cesty do zaměstnání či školy. Další způsob dělení je dle směru. Jsou pohyby jednosměrné, typicky migrační pohyby či stěhování a obousměrné, mezi než nejčastěji bývá řazen turismus. Vědecké studie se rozcházejí v tom, zda do mobility spadá pouze migrace (Kaufmann, 2002) či všechny zmíněné pohyby (Kellerman, 2006) (Urry, 2006).

Podle Pooleyho a kol. není možné mobilitu lidí jednoduše rozdělit. Snaží se ji tak pochopit skrze tzv. kontinuum mobility. Na jedné straně spektra najdeme residenční migraci, kde jednotlivec změní své obvyklé bydliště a má v úmyslu v dohledné době v nové lokalitě zůstat. K takové migraci může dojít na různých úrovních, s různými dopady, od mezinárodní migrace přes státní hranice až po stěhování z jedné ulice do druhé ve stejné městské oblasti. Zatímco v prvním případě migrace vyžaduje úplnou změnu všech aspektů každodenního života (práce, škola, přátelé) a každodenní mobilitu spojenou s těmito činnostmi, v druhém případě vyžaduje pouze drobné úpravy vzorců každodenního života a souvisejících cest. (Pooley, 2005)



Obrázek 1: Kontinuum mobility. Zdroj: (Pooley, 2005)

Někteří lidé se navíc běžně pohybují mezi dvěma a více domovy, a tím tak rozmazávají rozdíly mezi rezidenční migrací a denní mobilitou. Tento fenomén má mnoho příkladů a jejich počet výrazně vzrostl během 20. století. Mohou to být například lidé pracující v oblasti agrikultury či stavebnictví stěhující se za zrovna dostupnou prací, zaměstnanci služeb přesouvající se mezi pobočkami, studenti na kolejích či internátech anebo děti z rozpadlých rodin, které střídavě žijí u svých rodičů. Pro ně je stěhování mezi více domovy víceméně denní činností.

Na druhé straně spektra je každodenní mobilita (viz kapitola 2.3) a virtuální mobilita (viz kapitola 2.4).

Tim Cresswell definuje ve své knize On the move tři přístupy k mobilitě. První říká, že pokud hovoříme o lidské mobilitě, mluvíme o ní jako o hrubém faktu. Je pozorovatelná a empiricky měřitelná. Tato mobilita je analyzována například dopravními plánovači či vědci zabývajícími se migrací. Zpracovává jí výkonný hardware a software v animačních studiích a laboratořích sportovní vědy. Je to pohyb zkoumaný biometrickými systémy na letištích a mnoho dalšího. Tato mobilita je nejabstraktnější čistý pohyb.

Druhý přístup k mobilitě je vyjádřený skrze rozmanité soubory reprezentací od filmů až po zákony, od literatury k filozofii. Mobilita je tak zachycena v reprezentacích, které jí dávají smysl skrze produkci významů, které jsou často ideologické. Mobilita znamená toto, mobilita znamená tamto. A tak se samotný přesunu z bodu A do bodu B stává synonymem svobody, odporu, kreativity a života samotného.

Poslední, třetí, přístup popisuje mobilitu praktikovanou, zakoušenou a ztělesněnou. Mobilitu jako bytí světa. Způsob, jakým chodíme, například vypovídá o tom, jak se cítíme. Naše nálada ovlivňuje naše vnímání mobility. Často jsou prostředky, jakými osobně zakoušíme mobilitu a jak se pohybujeme, propojeny s významy přiřazenými mobilitě skrze reprezentaci. (Cresswell, 2006)

2.3 Každodenní mobilita

Každodenní mobilita zahrnuje všechny opakující se činnosti běžného života, které bereme často jako samozřejmé a přehlížíme je. (Gregory, 2009) Proto její první studie najdeme až na počátku 20. století. Jedním z prvních vědců, který se problematice každodenní mobility věnoval byl například Sigmund Freud, který zkoumal běžné procesy lidského chování v publikaci *Psychopatologie všedního života*. Sociologie, geografie, etnologie i sociokulturní antropologie se postupem času více a více zabývaly každodenním životem a každodenností až se koncem 70. let staly pevnou součástí těchto oborů. (Machová, 2015)

Francouzský marxisticky orientovaný filozof a sociolog Henri Lefebvre definoval každodennost („everydayness“) jako: „soubor funkcí propojujících do jednoho celku systémy, které se mohou zdát odlišné a oddělené...“. Do těchto systémů dle něj patří práce, rodina, soukromý život a volný čas. (Mulíček, 2013) A právě oddělení těchto složek vede ke každodennímu přemisťování. S tímto tvrzením souhlasí i Urry, dle něhož je tato mobilita tvořena vzorci a funkcemi. (Freudendal-Pedersen, 2009)

Toto každodenní přemisťování zůstalo v současnosti víceméně stejné, mění se především způsoby dopravy, které lidé používají. Ty jsou úzce spjaty s vývojem města samotného. Dopravní cesty město formují, pomáhají mu růst a profilovat se. Z tohoto důvodu je doprava ve městech často přirovnávána k lidskému krevnímu oběhu, protože bez ní město nemůže fungovat a její kvalita definuje kvalitu života ve městě. (Cresswell, 2006)

Každodenní mobilita samozřejmě ovlivňuje i jiné typy mobilit. Tomu se podrobně věnovali ve výzkumu Jonas Larsen, John Urry a Kay Axhausen. Jejich studie britských občanů potvrdila, že oproti začátku 20. století, kdy lidem trvala cesta do práce alespoň 30 minut, protože chodili především pěšky či jezdili na kole, nyní je toto omezení značně nižší a lidé tak nemusí bydlet blízko pracovišť. (Larsen, 2006)

Díky snazší dostupnosti každodenní dopravy nyní mohou lidé měnit zaměstnání bez toho, aby se museli stěhovat do jeho blízkosti či se mohou přestěhovat a stále dojíždět do stejného zaměstnání. Například ve Velké Británii byla dle Larsena a kol. na konci 19. století průměrná vzdálenost dojíždění 3,6 kilometru a cesta trvala 17,7 minut. Ve 30. a 40. letech dvacátého století, v době, kdy dominovala veřejná doprava, narostla průměrná vzdálenost na 7 až 8 kilometrů a doba dojíždění na 34 minut. A v 90. letech minulého století, kdy každé druhé dojíždění proběhlo za pomoci osobního automobilu, dosáhla průměrná vzdálenost na 14,6 kilometrů, ale průměrná délka zůstala na 34 minutách. Samozřejmě ale potvrzují, že stále převládají krátké každodenní cesty a lidé nejvíce pohybu udělají v lokálních oblastech, takže chůze zůstává i nadále významnou složkou každodenní mobility, obzvláště pro děti. (Larsen, 2006)

Díky tomu, že lidé nemusí bydlet blízko zaměstnání, mohou si vybrat místo k životu dle jiných prioritních bodů, které chtějí mít v blízkosti. Podle Knoxe a Pinche se dají rozdělit na tři skupiny. První skupinou jsou lidé s rodinou, kteří chtějí bydlet blízko škol či nemocnic, dále kariéristé, kteří jsou velmi mobilní a chtějí žít v oblastech odpovídajících jejich společenskému statutu. A třetí skupinu tvoři konzumně založení lidé, kteří vyhledávají například blízkost uměleckého života nebo restaurací. (Knox, 2010)

2.4 Virtuální mobilita

Již zmíněná virtuální mobilita se objevila díky vzniku nových médií. V dnešní době je možné určité interakce provádět skrze technologie. Lze tak komunikovat, nakupovat, prodávat, vzdělávat se či bavit se bez nutnosti fyzického setkání, a tak tato mobilita není téměř distančně

omezena. (Pooley, 2005) Virtuální mobilita tak neutlačuje fyzickou mobilitu, ale naopak ji doplňuje takovým způsobem, že se fyzická mobilita stává činností, během níž lze stále komunikovat a čas strávený přesunem z místa na místo dostává nový význam. V některých případech ji ale i nahrazuje, například u lidí, kteří nejsou schopní se fyzicky přesouvat, jimž virtuální mobilita umožňuje dostat se do kontaktu se světem. (Kellerman, 2006)

3 Výzkumy vzorů pohybu

Navzdory jejich významu pro městské plánování, odhady dopravních situací a šíření biologických či mobilních virusů, naše porozumění základním zákonům upravujícím lidský pohyb zůstává omezené, a to především z důvodu nedostatku nástrojů pro sledování časově rozlišené polohy jednotlivců. (Eubank, 2004) (Kleinberg, 2007) Přitom nalezení přirozené podobnosti v cestovních vzorcích by mohlo mít dopad na všechny jevy, které jsou způsobeny lidskou mobilitou, od prevence epidemií až po reakce na mimořádné události, urbanistické plánování a další. (González, 2008)

3.1 Vzorce pohybu u zvířat

Lidské vzorce pohybu ovlivňuje mnoho faktorů, od způsobů dopravy po pracovní a rodinné dispozice a priority. Proto mnoho studií mobility populace vychází z pohybu zvířat, a tak jsou odhady lidské trajektorie postavené na *náhodné procházce* či difúzních modelech. Tyto studie se prvně zaměřovaly na vzorce pohybu albatrosů, později například na pohyb opic nebo mořských predátorů. Autoři studie pohybu albatrosů umístili na 5 ptáků elektronická zařízení, která sledovala jejich veškerý pohyb a zda je pták ve vodě či na souši. Toto měření provedli devatenáctkrát během 3 letních měsíců na britských ostrovech Jižní Georgie. (Viswanathan, 1996) Z jejich výsledků vyplynulo, že trajektorie zvířecích pohybů se dá odhadnout pomocí Lévyho modelu (znám také jako Lévyho lety, více v kapitole 4.1.2), který můžeme definovat jako *náhodnou procházku*, která je vytvářena náhodnými kroky podle funkce hustoty pravděpodobnosti s *heavy-tail*. Velikost kroku je označena jako α. Začíná jako vysoká hodnota, postupně se snižuje a vrcholí v řešení směrem k poslední generaci. (Ong, 2014) (González, 2008)

3.2 Vzorce pohybu z běžného života

Další možností, jak sledovat pohyb obyvatel a vytvářet z něj vzorce, které by ho popisovaly je sledování pohybu bankovek. Tomu se věnuje studie *The scaling laws of human travel,* za niž stojí autoři D. Brockmann, L. Hufnagel a T. Geisel. Ti využili projekt *Where's George?*, který založil v roce 1998 Hank Eskin a který je zaměřený na sledování pohybu amerických bankovek. Eskin začal uvádět na internetové stránce http://www.wheresgeorge.com sériová čísla bankovek, jež mu prošly rukama a samotné

bankovky označovat razítkem projektu. Mohou se tak přidat i další lidé a doplňovat, kde na označené bankovky narazili. Aktuálně je v tomto projektu evidováno již skoro 300 milionů bankovek. Z již zmíněné vědecké studie, která se tomuto projektu věnuje vyplynulo, že i tento pohyb lze popsat Lévyho modelem, protože pohyb bankovek probíhá většinou na jednom místě, přerušený delším přesunem či delší nehybností. Autoři tvrdí, že jejich práce jako první přinesla empirický důkaz tohoto pohybu u lidí. (Brockmann, 2006) (Noulas, 2012)

Studie *Understanding individual human mobility patterns* od amerických vědců okolo Marty C. González analyzuje data z mobilních telefonů. K dispozici měli dvě sady dat, první zaznamenávala trajektorii pohybu 100 tisíc anonymizovaných uživatelů mobilních telefonů za šestiměsíční období. Tito uživatelé byli náhodně vybráni z celkových 6 milionů uživatelů. Pokaždé, když držitel telefonu přijal nebo odeslal textovou zprávu či když přijal nebo vyvolal hovor, zaznamenala se jeho pozice podle lokací vysílacích věží, na které byl připojen, a tak bylo možné vytvořit trajektorie pro každého uživatele. Protože uživatelé často komunikovali jen v určité časy a poté delší dobu telefon nepoužívali, zaměřili se autoři ve studii i na druhý datový set. Ten byl vytvořen z lokací 206 uživatelů mobilních telefonů měřených každé dvě hodiny po dobu jednoho celého týdne. Z výsledků zjistili, že jejich modely nekorespondují s Lévyho modelem, protože lidské trajektorie vykazují vysokou míru časové a prostorové pravidelnosti, přičemž každý jednotlivec je charakterizován časově nezávislou charakteristickou cestovní vzdáleností a vysokou pravděpodobností návratu na několik vysoce frekventovaných míst. Po korekci rozdílů v cestovních vzdálenostech a vrozené anizotropii každé trajektorie se jednotlivé modely cestování zhroutily do jediného rozdělení prostorové pravděpodobnosti, což naznačuje, že navzdory různorodosti jejich historie cestování lidé dodržují jednoduché reprodukovatelné vzorce. (González, 2008)

Další výzkum postavený na datech z mobilních telefonů je *Analyzing and Geo-visualizing Individual Human Mobility Patterns Using Mobile Call Records*. Provedli ho čínští vědci na Pekingské univerzitě v roce 2010. Za cíl práce si dali připravit několik způsobů preprocesování a prostorovou analýzu pro zpracování ICT dat pro těžbu individuálních lidských pohybových vzorců. Datový set pro tuto studii byl postaven na vzorku skoro devíti set tisíc uživatelů mobilních telefonů, kteří provedli alespoň jeden telefonní hovor ve vybraném termínu devíti dnů. Tato data měli autoři od čínského operátora a získaný dataset byl vyfiltrován na neuvedené devíti milionové čínské město. O hovorech vždy měli uvedeno jejich datum, čas, délku a lokace vysílací věží, ke kterým byli účastnící hovoru připojeni. Pro modelování zvolili 2D a 3D *space-time framework*. Pro simulování pohybů jednotlivce v časoprostoru adoptovali ortogonální 3D

souřadnicový systém, 2D systém pro prostor a 1D pro čas. Jednotlivá data zaznamenávali jako časoprostorové body vyjádřené jako $\langle x, y, z \rangle$, kde x a y určují lokaci ve 2D prostoru a z indikuje konkrétní čas. Další částí jejich výzkumu byla zaměřena na studii rozdílu pohybu u mužů a žen a u různých věkových skupin. Z toho měření zjistili, že mezi muži a ženami nepozorují žádné definovatelné rozdíly v pohybu a u věkových skupin pozorovali očekávaný rozdíl mezi lidmi v důchodovém a dospívajícím věku, kteří se pohybují méně a kratší vzdálenosti, naproti tomu mladí lidé a lidé ve středním věku mají průměrnou vzdálenost delší a také frekvenci mobility vyšší. (Kang, 2010)

3.3 Vzorce pohybu z nových médií

Noulas a kolektiv v práci *An Empirical Study of Geographic User Activity Patterns in Foursquare* pracuje s menší datovou sadou monitorující 111 dní v roce 2011, kdy zaznamenávali *check-iny* uživatelů sociální sítě Foursquare, které sami uživatelé sdíleli na sociální síti Twitter. Autoři se v této studii zaměřují na geotemporální dynamiku kolektivní aktivity uživatelů (*geotemporal dynamics of collective user activity*) Foursquare a snaží se v ní odhalit denní a týdenní vzorce, vlastnosti městských sousedství a opakující se přechody mezi různými činnostmi lidí. Z jejich statistické analýzy vyplynulo, že přibližně 20 % uživatelů vytvoří pouze jeden *check-in*[1], 40 % udělá více než deset a přibližně 70 000 aktivních uživatelů (okolo 10 %) udělá více než 100 označení. Dále zjistili, že počet *check-inů* na *venues* [2]některých kategorií není proporční k počtu skutečných návštěv. Například většina uživatelů se na svých domácích *venues* označuje výrazně méně častěji, než se na nich skutečně vyskytuje. Z toho vyplývá, že díky tomu, že tato data nejsou zaznamenána, je nemožné vytvoři celkový přehled celodenního pohybu uživatelů. Naproti tomu zjistili, že čím je uživatel aktivnější, tím je větší pravděpodobnost, že se na těchto méně častých typech *venues* označí. Z toho usuzují, že čím aktivnější uživatel, tím spíše se označuje na všech místech, kde se nachází. Na základě této heterogenity se dle nich zdá, že Foursquare je vhodný zdroj spíše pro datovou analýzu pohybu jednotlivců než populace jako celku. (Noulas, 2011) (Sun, 2015)

S touto myšlenkou pracují v další studii nazvané *A Tale of Many Cities: Universal Patterns in Human Urban Mobility*. V ní pracují s datovou sadou o 2 měsíce delší než ve studii poslední a tvrdí v ní, že univerzálnost zjištěná Brockmannem a spol. je v rozporu s

[1] Check-in – uživatelské označení své polohy na venue
[2] Venue – označení aplikace Foursquare pro jedno konkrétní místo

pozorováním, že posuny silně závisí na tom, kde se odehrávají. Jako příklad uvádějí studii o stovkách tisíc mobilních telefonů v Los Angeles a New Yorku, která demonstruje různě charakteristické délky cest ve dvou městech. Výsledky této studie například uvádějí, že medián vzdálenosti, kterou denně urazí obyvatelé Los Angeles je dvakrát větší než u obyvatel New Yorku, ale 25 % Newyorčanů cestuje 6krát dále než jejich protivníci z Los Angeles. (Isaacman, 2010) Tento výzkum dle nich naznačuje buď úplnou absenci univerzálních vzorců v lidské mobilitě, nebo skutečnost, že fyzická vzdálenost není správná proměnná k jejich vyjádření. (Noulas, 2012)

Další studie zaměřená na hledání vzorů pohybu z geolokačních sociálních sítí je italská práce *Mining human mobility patterns from social geo-tagged data*. V té autoři představují novou metodologii pro extrahovaní a analýzu časových a geo referencí spojených s daty ze sociálních sítí. A pomocí ní těží informace o lidské dynamice a chování uvnitř měst. Jako zdroj dat použili geolokované příspěvky z platformy Twitter a v něm provedli analýzu nepředvídatelných a nepravidelných informací. Zejména se zaměřili na získání sad denních trajektorií a použili sekvenční algoritmus dolování za účelem zisku tras častých cest. Poté definovali množinu prostorových a časových rysů nad těmito cestami, a provedli statistickou charakterizaci vzorů, pravidel a zákonitostí daných trajektorií. Jejich datová sada byla zaměřen na oblast Londýn, kde zaznamenali přibližně 7 a půl milionu tweetů v období šesti měsíců od června 2013 do listopadu včetně. (Comito, 2016)

Práce *Discovering Geographic Regions in the City Using Social Multimedia and Open Data* se zaměřuje na město Amsterdam. Cílem autorů bylo pokusit se za pomoci veřejně dostupných dat identifikovat různé oblasti ve městě. Dle jejich domněnky mohou být tyto oblasti charakterizovány specifickými vzorci, které souvisejí s jejich vizuálním vzhledem, se způsobem, jakým je popisují uživatelé sociálních médií a se vzorci lidské mobility. Aby toho dosáhli využili data získaná ze sociálních sítí Foursquare a Flickr. Z Foursquare získali místa zájmu a jejich fotografie nahrané uživateli a doplnili je fotkami, anotacemi a uživateli ze služby Flickr. Tato data doplnili o velké množství statistik od lokálních státních institucí, přidali tak například informace o demografii, bydlení či službách. Poté je zpracovali a vizualizovali za použití velké sady sémantických konceptů, které byly výsledkem konvoluční neuronové sítě[3] a extrahovaly latentní Dirichletova témata[4] z jejich anotací. Uživatele, textové a vizuální informace a lokální statistiky dále agregovali na úrovni oblastí dle poštovního směrovacího

[3] Neuronové sítě s pevnou strukturu propojení, vhodné například na klasifikaci obrázků.
[4] Metodu identifikace skrytých témat v textech.

čísla, které použili jako základ pro detekování větších oblastí ve městě. K identifikaci těchto oblastí provedli klastrování založené na jednotlivých modalitách i jejich celku. Závěrem díky experimentální analýze zjistili, že jimi automaticky detekované oblasti jsou smysluplné a mají potenciál pro lepší pochopení dynamiky a složitosti města. (Rudinac, 2017)

Vědci z Pekingské univerzity získali obrovskou datovou sadu od čínské nejmenované geolokační sociální sítě. Mohli tak pracovat s více než 23 miliony označení polohy od více než půl milionu uživatelů. Svůj výzkum popsali v práci nazvané *Uncovering Patterns of Inter-Urban Trip and Spatial Interaction from Social Media Check-In Data*. Data byla z období září 2011 až září 2012 a pokryla téměř celou Čínskou lidovou republiku. Dohromady vymezili 370 klasterů, které se nacházeli ve městech. Pro jejich analýzu se rozhodli použít gravitační model a také vytvořili prostorovou síť. Ze své práce vyvodili 4 aspekty. Za prvé, městské rozmístění bodů je popsatelné exponenciálním rozdělením a nepozorovali u něj výskyt *heavy-tail*[5]. Toto rozdělení je možné pozorovat i u městského pohybu, u něj ale nalezli *thin-tail*. Pokud by všechna jimi pozorovaná města byla identická a měla stejný počet obyvatel, poté by pohyb v nich následoval mocninné rozdělení. Zjistili, že velikost a místní charakteristiky měst jsou to, co diametrálně mění výsledky mocninného rozdělení a exponenciální funkce rozdělení. Za druhé zjistili, že jimi sledované prostorové interakce reflektované *check-iny* sedí velmi dobře na gravitační model. Tím potvrdili mocninné rozdělení u těchto prostorových interakcí, které dle nich bylo pozorováno i na mnoha jiných datových sadách. Argumentovali, že ač jiné výzkumy tvrdí, že gravitační model nemůže dobře odhadovat budoucí prostorové interakce, pokud je populace místa použita přímo jako masa v modelu, v jejich výzkumu je to naopak. Tvrdí, že využití gravitačního modelu pro odhadnutí jak rozsahu dané oblasti, tak funkce vzdálenosti (*distance decay function*), je správný přístup. Za třetí, jejich výzkum ukázal vazbu mezi prostorovými interakcemi a vzorci pohybu obyvatel. Funkce vzdálenosti může být také použita pro interpretování individuálních pohybů. Vzdáleností parametr $b = 0.8$ je nižší než ty, které odhadli v městské mobilitě a z toho dle autorů značí slabší efekt funkce vzdálenosti (*decay function effect*). Tvrdí ale, že nemohou vyvrátit, že jejich výpočet pro trajektorií nejde použít i na trasy jednotlivce. Různé individuální úrovně pohybových vzorů mohou vést ke stejné kolektivní statistice. A ukazují to na drobném výzkumu pohybu čtyř osob, kde pohyb jich všech je ovlivněn vzdáleností *decay effect*, ale zároveň není tento efekt známý pro každou osobu. Závěrem tohoto výzkumu tvrdí, že pro vytvoření precizních individuálních modelů by potřebovali více dlouhodobých a detailních dat. Poslední zjištěný aspekt je, že konstrukcí

[5] Graf funkce je strmější a okrajové a střední hodnoty mají vyšší pravděpodobnost.

prostorové sítě z dat *check-inů* mohli rozdělit oblast Čínské lidové republiky na regiony pomocí detekce komunit. Jejich výsledky oplývají stejným vzorem jako předchozí studie, ve kterých jsou komunity prostorově soudržné a ve shodě s geografickými jednotkami – provinciemi. Tyto vzorce mohou být také způsobeny efektem funkce vzdáleností, který obecně ovlivňuje přilehlá města k tvorbě pevnějších vazeb a uzavírání do klastrů. (Liu, 2014)

3.4 Vzorce pohybu z dotazníků

Jednou z mála českých studií věnujících se vzorům pohybu je *A City in Motion: Time-Space Activity and Mobility Patterns of Suburban Inhabitants and the Structuration of the Spatial Organization of the Prague Metropolitan Area* od autorů Jakuba Nováka a Luďka Sýkory. Zaměřuje se na hledání vztahů mezi příměstskými oblastmi a dalšími částmi pražské metropolitní oblasti zkoumáním časoprostorové aktivity a vzorců mobility obyvatel nově postavených příměstských čtvrtí. Zkoumáním každodenních aspektů chování obyvatel těchto nových čtvrtí identifikují vliv suburbanizace na měnící se geografii metropolitní oblasti a snaží se tak lépe porozumět, jak je prostorová organizace pražské metropolitní oblasti vytvářena, reprodukována a transformována. Zdroj jejich empirických dat je primárně složen z 262 deníků od osmdesáti osmi osob z třiceti domácností, doplněných dotazníky pro domácnosti a rozhovory s hlavami domácností. Pro samotný výzkum použili několik vzájemně propojených konceptů, kterými se zaměřili především na geografii času, teorii strukturace a studii postkomunistického města. Výsledky jejich práce potvrdily očekávané, že nové pražské příměstské čtvrti v metropolitní oblasti jsou naprosto závislé na nabídce pracovních míst a služeb ze středu metropole. Nově vybudovaná příměstská nákupní zařízení však do jisté míry tuto strukturu narušují a zachovávají určité denní aktivity obyvatel v příměstské zóně. Kromě empirických pozorování bylo klíčovým cílem jejich práce diskutovat a aplikovat koncepty a metody geografie času na výzkum restrukturalizace měst a pochopit strukturu metropolitní prostorové organizace. (Novák, 2016)

Několik studií také kombinuje různé datové sady, jednou z nich je například práce *Unravelling daily human mobility motifs*. Ve výzkumu se zaměřili na tři různé datové sady: dotazník o mobilitě z Paříže, údaje od francouzského operátora se seznamem uživatelů z Paříže a dotazník pro osoby pohybující se v americkém městě Chicago[6]. Z dotazníkových šetření autoři vybrali pro Chicago 23 764 a pro Paříž 23 429 popsaných dnů v týdnu od lidí, kteří byli

[6] https://www.cmap.illinois.gov/data/transportation/travel-survey

vybráni tak, aby data byla reprezentativní pro celou populaci města. V Chicagu museli účastníci vyplnit do dotazníku informace o své mobilitě během jednoho nebo dvou dní včetně detailů:

- Den v týdnu, kdy pohyb probíhal
- Délka pohybu
- Lokace, kde pohyb probíhal
- Důvod pohybu
- Zvolený způsob dopravy

Na základě těchto informací bylo možné vytvořit celodenní vzorce mobility pro anonymizované jednotlivce. Pařížský dotazník požadoval stejné informace, ale místo geografické lokace pouze délku trasy. Autoři se v této studii rozhodli věnovat pouze všedním dnům, protože diverzita mezi víkendy a pracovními dny byla příliš vysoká. Francouzský poskytovatel telefonního připojení poskytl autorům studie informace o příchozích a odchozích hovorech a o textových zprávách. Datová sada tak obsahovala lokace vysílacích věží zapojených v konverzaci, časové určení a anonymizované unikátní číslo uživatele. Z těchto informací byli vědci schopni rekonstruovat denní mapu mobility uživatelů během šesti měsíční periody. Po začištění dat jim zbylo 39 820 aktivních uživatelů, které podrobili výzkumu. Ten se zaměřil na analýzu časových a prostorových trajektorií tisíců osob jako jednotlivých sítí. Pomocí konceptu motivů ze síťové teorie, nalezli v každodenní mobilitě přítomných pouze 17 unikátních sítí, které se řídily jednoduchými pravidly. Tyto sítě, které nazývali motivy, jsou dostačující pro zachycení až 90% obyvatelstva v průzkumech a datových sadách z údajů o mobilních telefonech pro různé země. Každý jedinec vykazuje charakteristický motiv, který se zdá být stabilní po dobu několika měsíců. Denní lidská mobilita může být proto reprodukována analyticky sledovatelným frameworkem pro Markovovy řetězce[7] modelováním období častých cest následovaných obdobími nižší. Další modelování provedli pomocí Perturbationového modelu a pomocí radiačního modelu, o něm více v kapitole 4.2.1. (Schneider, 2013)

3.5 Shrnutí

Z vybraných studií je jasné, že nejlépe můžeme vzorce pohybu modelovat dle dat z dotazníků přímo od obyvatel měst, tak jako tomu bylo v historii. Jsme schopni tak zjistit body

[7] Náhodný pravděpodobnostní proces, pro který platí, že pravděpodobnost přechodu do následujícího stavu závisí pouze na současnosti, a ne na minulosti

trasy, důvod pohybu, délku pohybu a zvolený způsob dopravy. Data z dotazníků ale často trpí některými problémy:

- Náročnost získávání dat – při studii *Measuring Personal Travel and Goods Movement: A Review of the Bureau of Transportation Statistics' Surveys* byl použit telemarketing a úspěšnost hovorů byla pouze 41 %. (Measuring Personal Travel and Goods Movement, 2003)
- Finanční náročnost
- Malý vzorek – málo účastníků výzkumu zvyšuje chybovost a nepřesnost výsledků. (Richardson, 1996)
- Nepřiznané cesty
- Chybně přiznané – například chyby v časovém určení a jiné. (McGowen, 2007)

Proto se v dnešní době pro modelování mobility používají větší datové sady, často založené na GPS lokaci. K tomu se nabízí geolokační sociální sítě, které oplývající velkým množstvím dat, která pokud se dobře zpracují, tak opravdu mohou nabídnout vzorce pohybu obyvatel. Jak je znát z jednotlivých studií, způsobů, jak data zpracovat a modelovat je mnoho. V tabulce 1 je výběr známých výzkumů, jejich zdrojových dat a metod výzkumu z posledních let.

Název studie	**Zdroj dat**	**Metody výzkumu**
Lévy flight search patterns of wandering albatrosses	Měření polohy albatrosů	Uplatnění Lévyho modelu
The scaling laws of human travel	Pohyb bankovek	Lévyho model
Understanding individual human mobility patterns	Používaní mobilních telefonů	Lévyho model a gyrační poloměr, hledání mocninné funkce
Analyzing and Geo-visualizing Individual Human Mobility Patterns Using Mobile Call Records	Používaní mobilních telefonů	2D a 3D space-time framework

An Empirical Study of Geographic User Activity Patterns in Foursquare	Foursquare data stažena přes Twitter	Hledání mocninné funkce, rozdělení pravděpodobnosti
A Tale of Many Cities: Universal Patterns in Human Urban Mobility	Foursquare data stažena přes Twitter	Hledání mocninné funkce, intervenující příležitosti, rozdělení pravděpodobnosti
Mining human mobility patterns from social geo-tagged data	Geolokované Twitter příspěvky	Rozdělení pravděpodobnosti
Discovering Geographic Regions in the City Using Social Multimedia and Open Data	Foursquare, Flickr, Open data	Klastrování, Topic modelování
Uncovering Patterns of Inter-Urban Trip and Spatial Interaction from Social Media Check-In Data	Čínská geolokační sociální síť	Funkce vzdálenosti, exponenciální funkce rozdělení
A City in Motion: Time-Space Activity and Mobility Patterns of Suburban Inhabitants and the Structuration of the Spatial Organization of the Prague Metropolitan Area	Dotazník	Empirické pozorování, geografie času
Unravelling daily human mobility motifs	Dotazník, používaní mobilních telefonů	Síťová teorie, Perturbation model a radiační model

Tabulka 1: Přehled výzkumů vzorců pohybu. Zdroj: vlastní

4 Modelování mobility

Při hledání vzorů mobility se v dnešní době vychází především z pěti základních modelů. Těmi jsou: Gravitační model migrace, Lévyho model, Gyrační poloměr, Radiační model lidské mobility a Teorie intervenujících příležitostí. (Brockmann, 2006) (González, 2008) (Song, 2010) První dva modely staví na vzdálenosti a jejich použití má již dlouhou historii. Naproti tomu druhé dva modely sice staví na teoretických pracích starých i několik desítek let, ale aktivně se používají až v posledních letech. To je způsobeno mimo jiné rostoucí velikostí datových sad, které mají vědci v současné době k dispozici. (Simini, 2012) (Viswanathan, 1996)

4.1 Modelování mobility na základě vzdálenosti

4.1.1 Gravitační model migrace

Jednou z nejrozšířenějších empirických zákonitostí v regionálních studiích je, že jakákoliv forma prostorové interakce (migrace, dojíždění, obchodování, výměna informací, aj.) má vlastnosti svých toků vázané ke svým zásobám a nepřímo také ke vzdálenosti. A tak můžeme pozorovat, že zákony této prostorové interakce připomínají Newtonův zákon gravitační síly z roku 1687. (Anderson, 2011) S touto myšlenkou aplikování fyzikálních zákonů na pohyb populace mezi dvěma lokacemi přišel John Q. Stewart. (Stewart, 1950) Že vlastnosti migračních toků mají gravitační tendence ale potvrdil už o mnoho let dříve Ernst Georg Ravenstein. (Ravenstein, 1885)

V jeho obecně aplikované formě gravitační model migrace definujeme jako

$$M_{ij} = G\frac{P_i^{\alpha} \times P_j^{\beta}}{D_{ij}^{\gamma}},$$

kde M_{ij} značí počet rezidentů v oblasti j, kteří dříve žili v oblasti i; P_i a P_j jsou celkové populace v oblastech i a j (většinou měřeno v době migrace); D_{ij} je vzdálenost mezi i a j; α, β, γ jsou parametry, které se musí odhadnout a G je proporční konstanta, která závisí na kontextu (geografie, časová dispozice atd.).

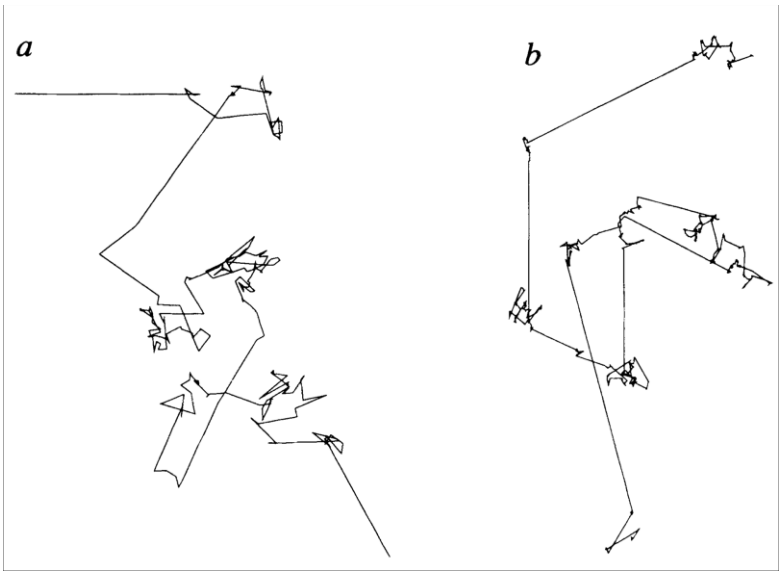
Užití gravitačních modelů výrazně vzrostlo v 60. letech 20. století, kdy Tinbergen publikoval analýzu toků mezinárodního obchodu právě prostřednictvím gravitačního modelu a na něj pak navazovali další autoři, jako Anderson, Bergstrand, Deardorff, Frankel a Romer a další. (Letouzé, 2009)

4.1.2 Lévyho model

Dalším modelem používaným pro modelování migrace lidí a v tomto případě i často zvířat je Lévyho model (či Lévyho let). Ten pojmenoval po francouzském matematikovi Paulu Lévym francouzsko-americký matematik Benoît Mandelbrot. (Mandelbrot, 1982) Lévyho model umožňuje náhodný posun, jehož krok je brán z Lévyho stabilní distribuce s nekonečnou variancí. Pro tento algoritmu je charakteristické střídání krátkých úseků s dlouhými skoky a často je pozorovatelný u pohybu zvířat. (Viswanathan, 1996) (Yang, 2010)

$$L(c,\lambda) = \frac{\lambda\Gamma(\lambda) \times \sin\left(\frac{\pi\lambda}{2}\right)}{\pi} \times \frac{1}{c^{1+\lambda}}$$

Kde $\Gamma(\lambda)$ reprezentuje gamma funkci, $c \in R$ je váha a $\lambda\epsilon[0,2]$ je exponent funkce hustoty. V případě nastavení parametru α = 1 se jedná o Cauchyho distribuci, při $\alpha = 2$ o Gaussovu distribuci. (Yang, 2013)



Obrázek 2: Pohyb albatrosů, který popisuje Lévyho model. Zdroj: (Viswanathan, 1996)

4.1.3 Gyrační poloměr

Někteří lidé pravidelně cestují na delší vzdálenosti, než je průměr, naproti tomu jsou lidé, kteří se stále pohybují na malé oblasti. Proto vzniknul Gyrační poloměr, který má zachytit právě tento fenomén a indikovat charakteristickou vzdálenost cestování pro uživatele za určitý časový úsek. Každý člověk si tak v rámci svého poloměru gyrace vybere svou destinaci. (Yan, 2011)

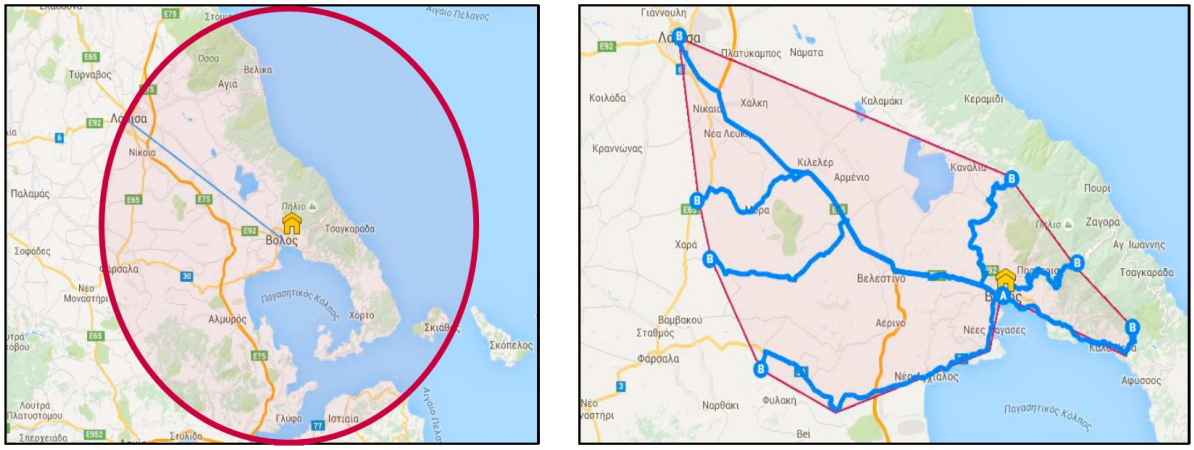
4.2 Modelování mobility na základě hodnocení

4.2.1 Radiační model lidské mobility

Radiační model se prvně se objevil ve studii *The System of Intermetropolitan Population Flows* od amerického ekonoma Williama Alonsa publikované v roce 1971. Nejpřesněji ho ale definoval Filippo Simini a kolektiv o čtyřicet let později. Ti vyčítají gravitačnímu modelu jeho parametry, které se liší podle oblasti použití, a tak znemožňují univerzální použití tohoto modelu. Proto ve vzorci radiačního modelu žádné parametry nejsou, jeho vzorec je

$$T_{ij} = T_i \frac{m_i n_j}{(m_i + s_{ij})(m_i + n_j + s_{ij})},$$

kde T_{ij} značí celkový počet lidí cestujících z i do j; m a n jsou populace v daných oblastech a s je veškerá populace v rádiusu od bodu i. Výsledný radiační model predikuje vzorce mobility ve shodě s dopravními vzorci pozorovatelnými v řadě jiných výzkumů. (Simini, 2012) (Zipf, 1946)



Obrázek 3: Porovnání gravitačního a radiačního modelu. Zdroj: (Stefanouli, 2017)

4.2.2 Teorie intervenujících příležitostí

V roce 1940 vyšel v *American Sociological Review* článek *Intervening Opportunities: A Theory Relating Mobility and Distance* od amerického sociologa Samuela A. Stouffera. V něm autor představil teorii intervenujících příležitostí, kterou se snažil popsat pravděpodobnost migrace lidí. Jeho hypotéza tvrdí, že větší vliv na mobilitu populace má oproti samotné vzdálenosti či populačním tlaku v počátečním místě migrace nabídka příležitostí v jejich cílové destinaci. Tato hypotéza je v souladu s druhým, třetím a čtvrtým Ravenstainovým zákonem migrace (Stouffer, 1940)

Stoufferem definovaný zákon intervenčních příležitostí je „Počet lidí, kteří se přesouvají danou vzdálenost je přímo úměrný počtu příležitostí na daném vzdáleném místě a nepřímo úměrný počtu intervenujících příležitostí.“. Podle autora intervenující příležitosti přesvědčují migranty usadit se již na některém z míst během cesty a nedorazit tak do cílového místa. Protikladem této myšlenky je Zipfův zákon z roku 1946. (Stouffer, 1940) (Zipf, 1946)

Kořeny Zipfova zákona jsou v gravitačním modelu a v práci Gasparda Mongeho v 18. století. Jak již bylo řečeno, gravitační model závisí na nastavitelných parametrech, které se liší oblast od oblasti a trpí tak analytickou nekonzistencí.

Proto se vhodnější využít zmíněný radiační model, stochastický proces zaznamenávající lokální pohybové rozhodování, který nám pomáhá analyticky odvodit a definovat dojíždění a mobilitu a jeho vstupními hodnotami jsou jen informace o distribuci obyvatelstva. (Simini, 2012)

5 Vlastní výzkum

V rámci této práce byl proveden výzkum nad datasetem označení polohy z aplikace Foursquare a nad seznamem jízd z projektu Rekola. Cílem výzkumu bylo potvrdit či vyvrátit použití teorie intervenujících příležitostí pro určování vzorů pohybu obyvatel ve městech, na příkladu Hlavního města Prahy. Pro porovnání výsledků byly použity jiné studie, které zkoumají stejnou problematiku, ale nad jinými daty a jinými městy. První fáze výzkumu proběhla v přímém porovnání se studií *A Tale of Many Cities: Universal Patterns in Human Urban Mobility*. Další fází byla shluková analýza kategoriálních dat ze stejného datasetu. A poslední fází byl stejný výzkum jako v první části, ale s daty z bikesharingového projektu Rekola.

5.1 Foursquare

Geolokační sociální službu Foursquare spoluzaložili v roce 2009 Dennis Crowley a Naveen Selvadurai. Jejím předchůdcem byla služba Dodgeball, kterou Dennis Crowley vytvořil v roce 2000 jako svou absolventskou práci na New York University. Dodgeball koupil v roce 2005 Google a poté v roce 2009 ji nahradil službou Google Latitude, kterou ukončil v roce 2013 a její funkcionality integroval do produktu Google Maps. (Foursquare, 2019)

Foursquare byl od začátku vytvářen jako aplikace pro mobilní telefony. Prvních 5 let své existence zahrnoval v jedné aplikaci *check-iny* (označení a sdílení polohy uživatele na konkrétním místě) své a svých přátel, hledání *venues* (místa pro *check-in*), vyhledávání a hodnocení míst aj. Poté se v roce 2014 Foursquare rozdělil na dvě aplikace: Foursquare City Guide a Swarm. Swarm obsahuje vše okolo konkrétních *check-inů* a Foursquare je nyní pouze vyhledávač a rozhraní pro hodnocení *venues*. Foursquare měl na konci minulého roku alespoň 50 milionů aktivních uživatelů měsíčně. Jedná se tak o nejrozšířenější geolokační sociální službu. Celkový počet *check-inů* v srpnu 2017 překročil 12 miliard od začátku služby a ve Spojených státech amerických ho používalo přibližně 13 % uživatelů starších 18 let v roce 2018. (Foursquare, 2019) (30 Foursquare Statistics to Help You Optimize the Platform in 2019, 2019)

5.2 Metoda výzkumu *A Tale of Prague*

Zvolená metoda vychází z výzkumu *A Tale of Many Cities: Universal Patterns in Human Urban Mobility*. Jeho autoři měli k dispozici datovou sadu z geolokační sítě Foursquare o 35 289 629 *check-inech* (z toho okolo 10 milionů ve městech) od 925 030 uživatelů na 4 960 496 místech, sesbíranou od 27. května 2010 do 3. listopadu 2010. Tu získali pomocí sociální sítě Twitter, na kterou uživatelé sdíleli své *check-iny* veřejně. Nejprve provedli test shody modelu s pozorováními pomocí hustoty rozdělení pravděpodobnosti, poté Kolmogorovův–Smirnovův test, a nakonec metodu maximální věrohodnosti pro získání mocninných funkcí.

Kolmogorovův–Smirnovův test je způsob testování normality, ale implementace v R umožňuje porovnávat i více než 5000 vzorků. A pro takovéto velikosti už by měl být tento test také více než dostatečně silný. Kolmogorovův–Smirnovův test je test obecnější, umožňuje porovnávat s libovolným rozdělením, případně porovnávat, zda dva konkrétní výběrové soubory pocházejí ze stejného rozdělení, aniž bychom konkrétně specifikovali, z jakého. (Bořil, 2015)

Z prvních částí výzkumu nevyplynuly žádné univerzální vzorce, a tak dalším krokem bylo, že určili každé cestě mezi dvěma body hodnotu (*rank*). Tu určili na základě intervenujících příležitostí, konkrétně tak, že hodnota každé cesty je součet destinací, které jsou blíže startovní pozici než konečná destinace.

Závěrem sérií simulací potvrdili hypotézy, že hustota zájmových míst, je řídícím prvkem městské mobility. Jen z informací o distribuci míst ve městě a z ohodnocení pohybu obyvatel byli schopni reprodukovat skutečnou distribuci mobility pozorovanou na reálných datech. Tyto výsledky dle autorů studie otevřely novu cestu dalším výzkumům a ukázaly, jak můžeme těžit z geolokačních služeb, pomocí nichž jsme schopni získat mnoho praktických dat. (Noulas, 2012)

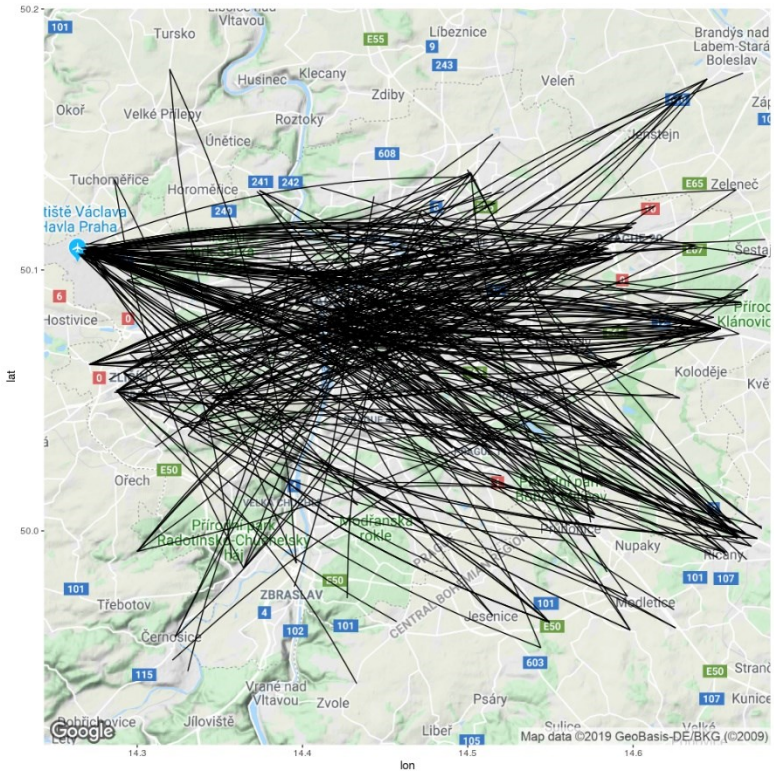
5.2.1 Výzkum

V této práci je snaha tento výzkum reprodukovat nad jinou datovou sadou s cílem zkoumat mobilitu v hlavním městě České republiky, Praze. Praha má rozlohu 496,2 $km^{2,}$ počet obyvatel 1 268 796 a hustotu zalidnění 2 557 obyvatel na kilometr čtvereční. (Integrovaná strategie pro ITI Pražské metropolitní oblasti, 2016)

Pro tento výzkum byla použita datová sada z práce *NationTelescope: Monitoring and Visualizing Large-Scale Collective Behavior in LBSNs* od autorů Dingqi Yang, Daqing Zhang, Longbiao Chen a Bingqing Qu. Ti vytvořili platformu *NationTelescope*, která slouží k monitoringu, porovnávání a vizualizaci dat z geolokačních sociálních sítí (*Location Based Social Networks*). Pomocí této služby získali ze sociální sítě Foursquare data za 18 měsíců od dubna 2012 do září 2013. Během této doby zaznamenali 33 278 683 označení (*check-inů*) po celém světě dohromady na 3 680 126 místech (*venues*) od 266 909 aktivních uživatelů. Tuto datovou sadu zveřejnili pro další studie ve formě tři souborů, jeden seznam všech *check-inů*, jeden seznam všech *venues* a jeden seznam měst. Seznam *check-inů* obsahuje anonymizované ID uživatele, ID *venue* ze sociální sítě Foursquare, na které si udělal *check-in*, čas ve formátu UTC a odchylku času *check-inu* od UTC času. V seznamu *venues* je 7 sloupců: ID *venue*, zeměpisná šířka a zeměpisná délka, kategorie *venue* dle Foursquare kategorií a kód státu dle ISO 3166-1 alpha-2. V posledním souboru měst nalezneme název města, zeměpisnou šířku a délku středu města, kód státu a celé jméno státu a typ města (hlavní město státu, provincie atd.). (Yang, 2015)

venue_id	user_id	timestamp	timediff	venue_lat	venue_lon	category	country
4b825bfaf964a52077d030e3	159493	Fri Apr 05 04:54:19 2013	120	50.08017	14.50109	Hospital	CZ
4b75302df964a52074002ee3	25870	Fri Apr 05 05:24:50 2013	120	50.11902	14.48728	High School	CZ
4dadb50743a18e54131bf0a8	206687	Fri Apr 05 06:01:02 2013	120	50.10471	14.26628	Airport Gate	CZ
4af3c836f964a52056ef21e3	62275	Fri Apr 05 06:53:27 2013	120	50.08435	14.44114	University	CZ
4b0bc619f964a520873323e3	89104	Fri Apr 05 06:56:39 2013	120	50.09976	14.35970	High School	CZ
4c987f0d05a1b1f704c48d53	66309	Fri Apr 05 06:57:36 2013	120	50.10103	14.38606	College Residence Hall	CZ
4af3c836f964a52056ef21e3	62275	Fri Apr 05 06:59:47 2013	120	50.08435	14.44114	University	CZ
4b75302df964a52074002ee3	52365	Fri Apr 05 07:04:44 2013	120	50.11902	14.48728	High School	CZ
4af1dc86f964a520e1e321e3	93636	Fri Apr 05 07:51:57 2013	120	50.08843	14.47796	Dog Run	CZ
515e85a7e4b03ca04216156b	24789	Fri Apr 05 08:05:15 2013	120	50.10764	14.26919	Rental Car Location	CZ
4ce3b357c36ba35d722171f6	159402	Fri Apr 05 08:08:49 2013	120	50.07258	14.40883	Tattoo Parlor	CZ

Tabulka 2: Ukázka z datasetu Foursquare. Zdroj: vlastní a (Yang, 2015)



Obrázek 4: Hrubá vizualizace tras *check-inů* uživatelů v Praze. Dobře rozpoznatelná je nejnavštěvovanější *venue* – Letiště. Zdroj: vlastní

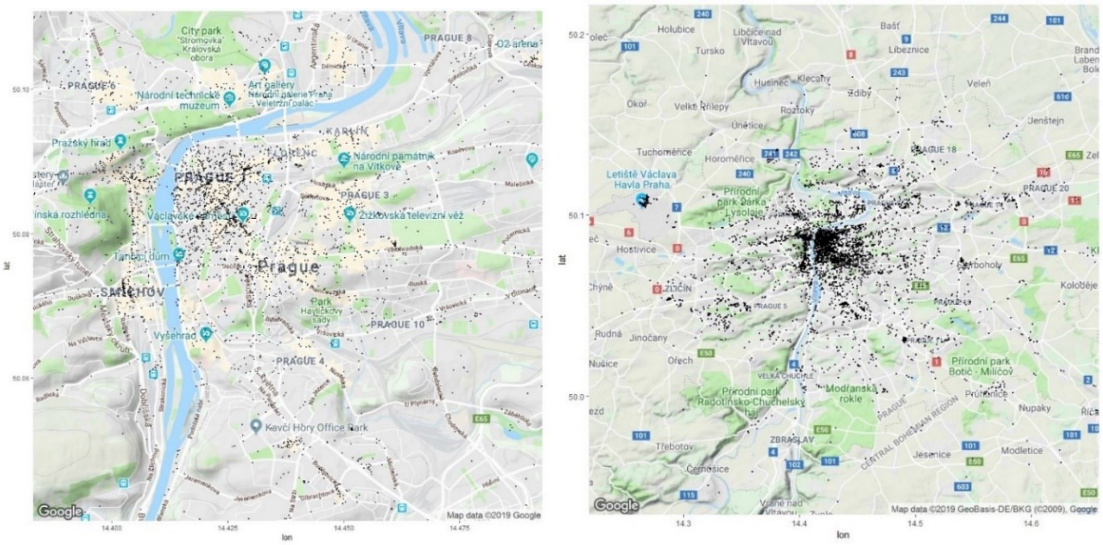
5.2.2 Zpracování dat

Pro zpracování dat byl použit programovací jazyk R[8] určený pro statistickou analýzu dat a jejich grafické zobrazení, který je zdarma dostupný pod licencí GNU GPL v2. Pro práci bylo použito grafické prostředí jazyka *RStudio*[9], které je pro nekomerční účely také zdarma.

Pro modelování a vizualizace byl použit buď tabulkový procesor Microsoft Excel nebo knihovny jazyka R: *dplyr*, *ggplot*, *ggmap* a další. (Wickham, 2015) (Wickham, 2016) (Kahle, 2013)

Prvním krokem bylo z dat vybrat ta týkající se *venues* v Praze. K tomu byly, pro zjednodušení, použity čtyři zeměpisné souřadnice ohraničující město ze čtyř světových stran a vytvářející tak čtyřúhelník přes celé město. Ty jsou 14.217 až 14.709 stupňů východní délky a 49.941 až 50.178 stupňů severní délky, tato data byla získána z datového modulu Wikipedie[10]. Pomocí těchto souřadnic bylo možné omezit seznam *venues* na 5 622 míst.

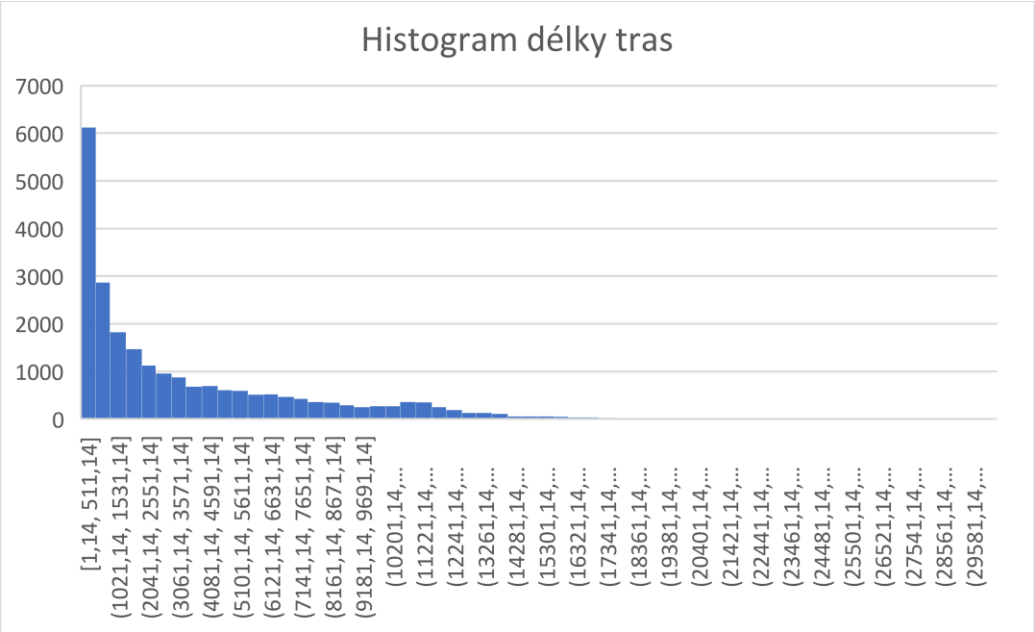
[8] https://www.r-project.org/
[9] https://www.rstudio.com/
[10] https://en.wikipedia.org/wiki/Module:Location_map/data/Czech_Republic_Prague



Obrázek 5: Vizualizace *venues* v Praze. Vlevo *venues* jen s více *checkiny*. Zdroj: vlastní

Dalším krokem bylo omezit seznam *check-inů* na ty, které se staly na *venues* v Praze, a to pomocí porovnání ID místa v seznamu pražských *venues* a v seznam *check-inů*. Tak se seznam vyfiltroval na 25 512 označení během měřených 18 měsíců, a to od celkem 1 321 unikátních uživatelů. Dále byl použit časový údaj každého označení a byla vytvořena tabulka jednotlivých cest z bodu A do bodu B, která obsahovala 23 613 řádků, tedy jednotlivých přesunů. U těchto jednotlivých přesunů byla změřena vzdálenost pomocí R knihovny *geosphere*, která obsahuje funkci *distm*, která výměnou za dva parametry souřadnic vrátí jejich vzájemnou vzdálenost v metrech. (Hijmans, 2019)

Nejdelší naměřená trasa mezi dvěma *check-iny* byla 30 590 metrů. Pokud prozkoumáme kvantily těchto vzdáleností, zjistíme, že dolní kvantil je na 482 metrech a horní na 5 479 metrech, to vidíme i na histogramu, kde jsou vidět prvky mocninového rozdělení. Průměr všech vzdáleností byl sice 3 607 metrů, ale medián pouze 1 857. Je tedy jasné, že většina pohybu je na krátké trase do 2 kilometrů.



Graf 1: Histogram délky tras z Foursquare datasetu. Zdroj: vlastní

5.2.3 Městský pohyb a zákon síly

Noulas a kolektiv použili pro jejich výzkum výpočty mocninných funkcí na základě zákona síly. S ním se můžeme setkat často pod názvem jako Long Tail – Dlouhý chvost či jako Pareto distribution – Paretovo rozdělení, které může způsobovat tzv. Paretův princip (známé také jako pravidlo 80-20). Ten se nachází všude okolo nás, v populaci měst, Richterově škále zemětřesení, rozdělení bohatství či v pohybu populace. (Newberry, 2019)

Během hledání, a hlavně popisování zákonu síly vzniká mnoho chyb, například pří použití metody nejmenších čtverců. Proto byl použit bezpečnější a již zmíněný postup kombinující Kolmogorovův–Smirnovův test s maximální věrohodností pro získání mocninných funkcí. (Clauset, 2009)

Pro výpočet zákonu síly a mocninných funkcí je použit tento vzorec

$$p(x) \propto x^{-\alpha},$$

kde α je konstantní parametr rozdělení, známý také jako exponent nebo škálovací parametr. Tento parametr se většinou nachází v intervalu $2 < \alpha < 3$, ale výjimky jsou možné. (Clauset, 2009)

City Name	Movements	Places	Density (Places/km)	Area (km^2)	Mean transport
Amsterdam	2934	8847	275,61	21,63	2,29
Atlanta	3220	10090	214,72	19,94	5,37
Austin	296	9492	199,32	14,06	5,82
Bangkok	5860	7574	248,32	10,81	3,97
Boston	2196	6795	366,94	13,25	1,57
Chicago	85496	23050	315,16	41,94	4,02
Columbus	2388	7463	181,18	8,88	5,42
Dallas	9380	8177	200,8	13,06	5,21
Denver	695	6123	215,26	12,81	4,67
Houston	7996	11808	168,68	14,63	7,57
Indianapolis	382	6417	213,02	5,38	6,99
Kuala Lumpur	2595	14223	268,44	30,88	3,18
Las Vegas	2437	11910	260,39	16,63	4,76
London	2837	15760	290,92	30,5	3,32
Los Angeles	6092	18508	220,92	31,5	4,86
Milwaukee	8697	5318	218,77	9,56	3,15
Minneapolis	9572	5482	228,04	11,13	3,1
New York	71502	43681	715,02	58	2,24
Orlando	7783	8060	224,56	8,88	5,44
Paris	8392	12648	261,98	35,94	2,77
Philadelphia	4545	10270	293,2	17,31	2,86
Phoenix	4436	8689	183,1	9,44	6,27
Portland	8409	8413	238,34	15,63	3,08
Rio de Janeiro	5808	6788	248,2	12,31	5,99
San Antonio	3516	8237	144,17	6	8,35
San Diego	9152	13365	227,26	22,38	5,7
San Francisco	12168	15970	377,64	32,25	2,36
Santiago	6743	10636	235,17	20,69	4,94
Seattle	6423	10410	294,6	20,75	3,61
Seoul	4303	9271	250,76	18,31	4,8
Singapore	9624	15617	316,67	21,31	5,26
Sao Paulo	2855	14291	224,68	32,56	4,31
Toronto	7548	13870	322,26	24,81	3,59
Washington	1557	10279	325,11	21,31	1,92

Tabulka 3: Tabulka s přehledem pohybu, míst, hustoty míst, plochy a průměrné délky trasy v porovnávaných městech. Zdroj: (Noulas, 2012)

Pro porovnání s ostatními výzkumy, byl zjištěn exponent mocninného rozdělení délky trasy. A to za pomocí knihoven *igraph* a *poweRlaw*. (Csárdi, 2019) (Gillespie, 2019) Exponent α vyšel s *poweRlaw* 4,594297 a s *igraph* 3,065155, to je na hranici obvyklých hodnot, které jsou $2 < \alpha < 3$ dle *POWER-LAW DISTRIBUTIONS IN EMPIRICAL DATA*. (Clauset, 2009) Hodnoty jsou různé, protože každá knihovna používá jiný algoritmus. Knihovna poweRlaw

pracuje s metodou maximální věrohodnosti, tu použili také ve studii *A Tale of Many Cities* a dosáhli srovnatelného výsledku $\alpha = 4.67$.

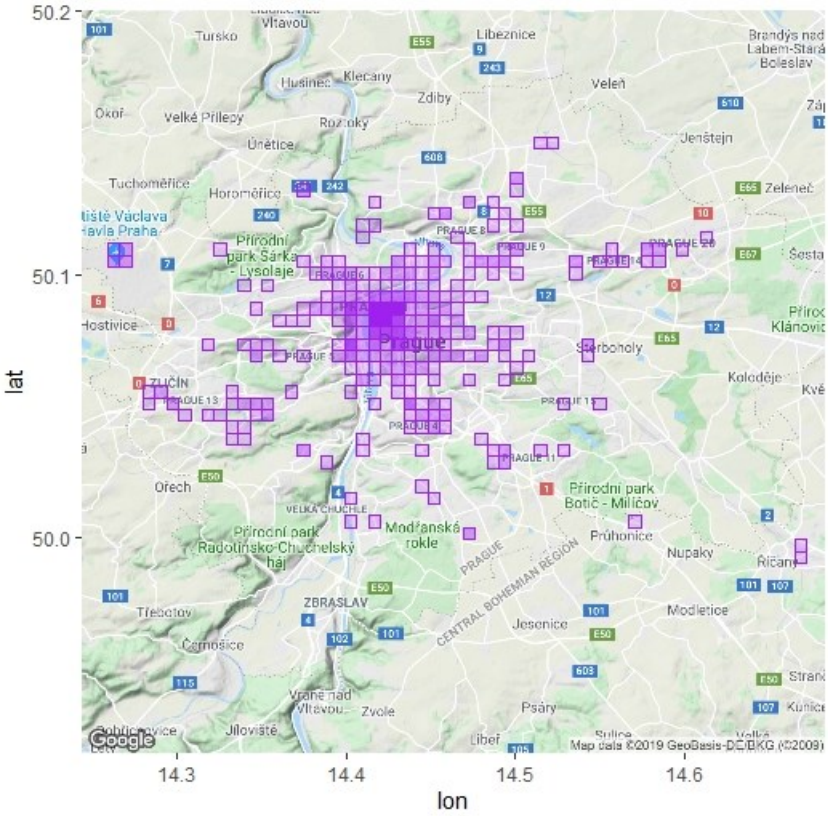
S cílem potvrdit výsledky ze studií *The scaling laws of human travel* a *Understanding individual human mobility patterns* nejprve vypočítali distribuci přesunu v celém datasetu. Vypozorovali, že tato distribuce odpovídá silovému zákonu s exponentem 1,5 a p-hodnotou 0,494. Tyty hodnoty se téměř shodují s exponentem změřeným u pohybu dolarových bankovek (1,59) a také s hodnotami z analýzy hovorů z mobilních telefonů (1,75). V kontrastu těchto studií dále pozorovali, jak díky technologii GPS, je možné používat tyto geolokační sociální sítě a mít tak zaznamenané přesné polohy na jednotky metrů. Tomu odpovídají také jejich data, ze kterých vyplívá, že do 100 metrů je pravděpodobnost pohybu jednotná.

V námi zkoumaném vzorku toto není a funkce hustoty pravděpodobnosti má stále klesající tendenci. To je nejspíš způsobeno menším vzorkem dat nebo použitím většího rozptylu mezi *check-iny* uživatelů.

Jak již bylo zmíněno, námi naměřený exponent hustoty pravděpodobnosti délky všech cest dosáhl 4,594297, který je téměř stejný jako naměřený ve studii $\alpha = 4.67$. S tímto exponentem byla zkoumána metoda maximální věrohodnosti a výsledkem bylo, že stejně jako v původní studii není shoda silového zákonu dostatečná, a tak není možné pomocí těchto hodnot definovat vzory pohybu.

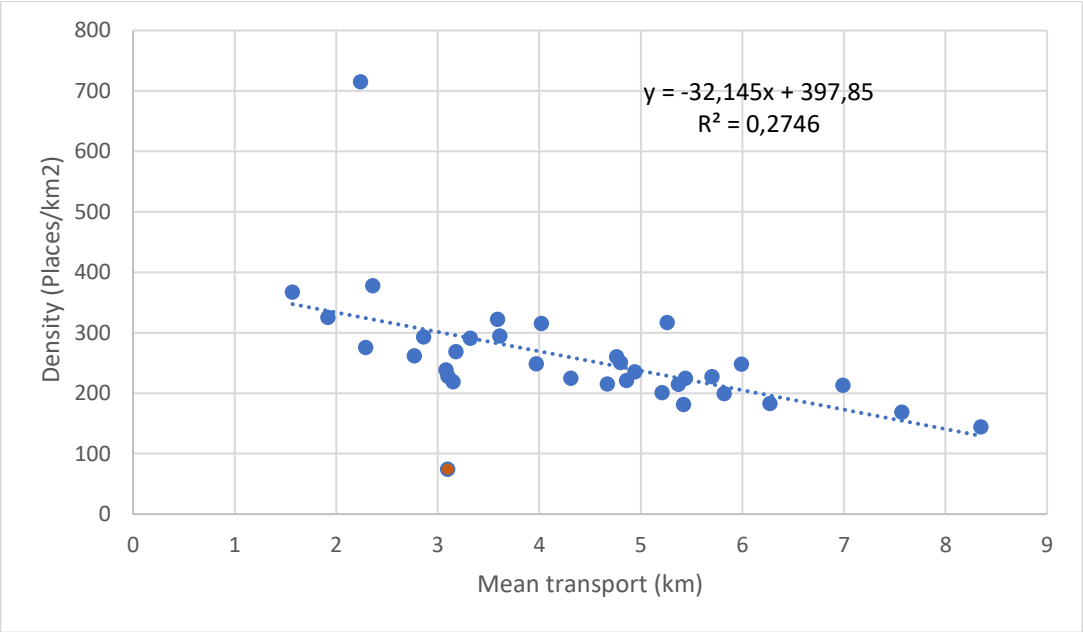
5.2.4 Vliv hustoty oblasti

Pokud chceme vyzkoušet platnost Stoufferova zákona, je nutné se nejdříve zaměřit na hustotu míst. Je pak možné zjistit do jaké míry je jejich hustota navázána na lidský pohyb ve městech. (Stouffer, 1940) Prvním krokem k získáním těchto informací bylo rozdělení města na menší pole oblastí a zjištění počtu zájmových míst v nich. Pro zjednodušení bylo použito čtyřúhelníkové souřadnice Prahy a byla rozdělena po 0.007004 stupních šířky a po 0.004497 stupních délky, tedy na pole o rozloze přibližně 0.25 km². Pole, ve kterých se nacházelo alespoň 5 *venues* byla zaznamenána a zanesena na mapu, dohromady jich bylo 244 oblastí. Očekávaně je nejzahuštěnější oblast okolí dolní část Václavského náměstí (50.081,14.422-50.086,14.427) s 211 *venues* na 0.25 km².

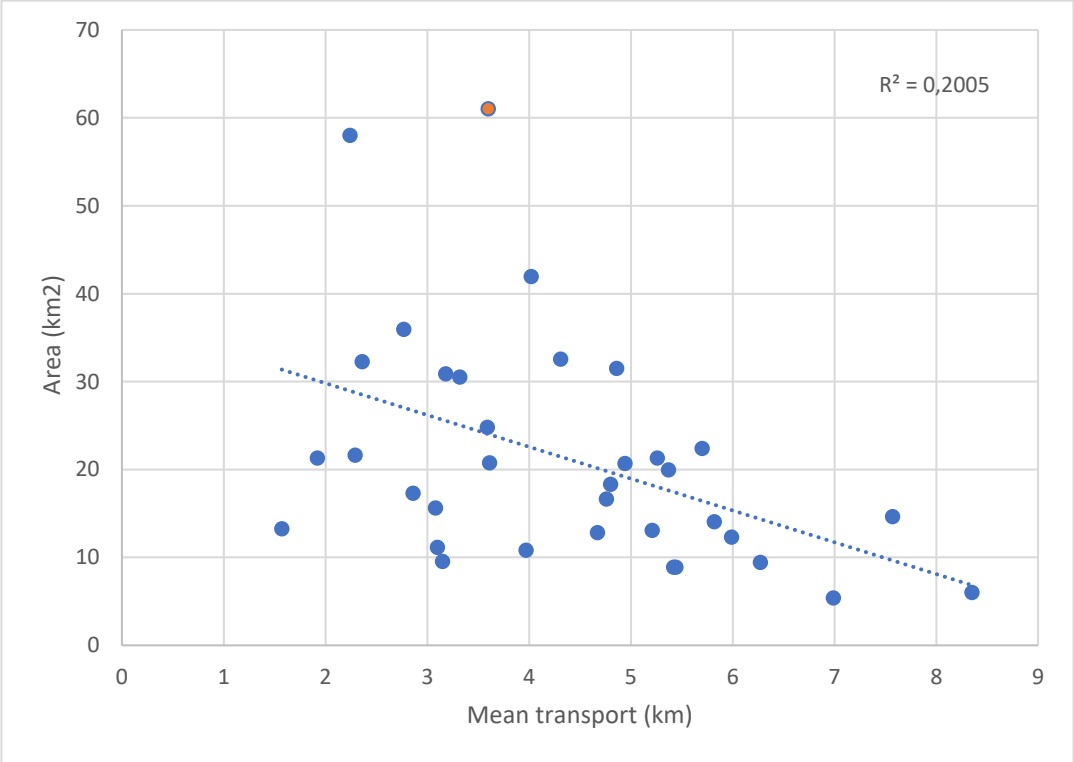


Obrázek 6: Vizualizace oblastí o ploše 0,25km^2, které měly alespoň 5 *venues*. Zdroj: vlastní

Celková rozloha oblastí s alespoň 5 *venues* je 61 kilometrů čtverečních a je na nich 4 512 míst. Hustota je tak přibližně 74 *venues* na kilometr čtvereční a s průměrnou délkou pohybu přibližně 3 kilometry na jednu trasu vidíme, že v porovnání s jinými městy ze studie *A tale of many cities* jsou tyto hodnoty o mnoho nižší. Z jejich viditelné korelace je odvoditelné, že lidé počítají s nákladností větších vzdáleností, dostupnost konkrétních míst je taky důležitá. Za příklad uvádějí například kvalitní trhy, na které jsou lidé ochotni jezdit větší vzdálenosti. Také se nám přidáním Prahy do grafu výrazně snížil Koeficient determinace, který nově dosahoval pouze $R^2 = 0{,}2746$ (rovnice trendu je nově $y = -32{,}145x + 397{,}85$) oproti původním $R^2 = 0{,}59$. To značí ještě nižší výpovědní hodnotu grafu. Pokud Prahu umístíme i na druhý graf, potvrdíme zjištění Noulase a kolektivu, že mezi průměrnou vzdáleností trasy a oblastí je ještě menší korelace než mezi průměrnou vzdáleností a hustotou míst. Koeficient determinace sice po rozšíření grafu vzroste z $R^2 = 0{,}19$ na $R^2 = 0{,}2746$, ale to je stále příliš nízká hodnota, aby měly tyto grafy pro nás váhu.



Graf 2: Graf závislosti hustoty *venues* a průměrné délky trasy. V dolní části oranžově Praha. Zdroj: vlastní a (Noulas, 2012)



Graf 3: Graf závislostí plochy a průměrné délky trasy. V horní části oranžově Praha. Zdroj: vlastní a (Noulas, 2012)

Dalším krokem při zjišťování, zda je hustota faktorem ovlivňující lidskou mobilitu bylo určení hodnoty každé cesty mezi dvěma místy. Ta se určovala tak, že se spočetl počet míst w, která byla blíže, myšleno vzdáleností, k počáteční *venue* u než cílová *venue* v. Tato hodnota má výhodu, protože i když se zakládá na absolutních číslech je v konečných hodnotách relativní, a tak nezávislá na oblasti či městě.

$$rank_u(v) = |\{w : d(u,w) < d(u,v)\}|$$

Na tu to část výzkumu byla použita knihovna RANN[11], která obaluje C++ knihovnu *ANN: A Library for Approximate Nearest Neighbor Searching* a umožňuje tak velmi rychlé (O(N log N)) hledání vzdáleností mezi body. (Arya, 2019) Bylo tak možné během pár desítek minut spočítat kolik *venues* je blíž a určit tak hodnotu za pomocí tohoto kódu:

```
library(RANN)

intv_ops <- list()

for(i in 1:nrow(trips)) {
  intv_ops[[i]] <- nn2(dest, trips[i, 2:3], searchtype = "radius", radius = (trips$distance[i]/1.11) *
0.00001)$nn.idx
}

trips$rank <- sapply(intv_ops, function(x) sum(x != 0))
```

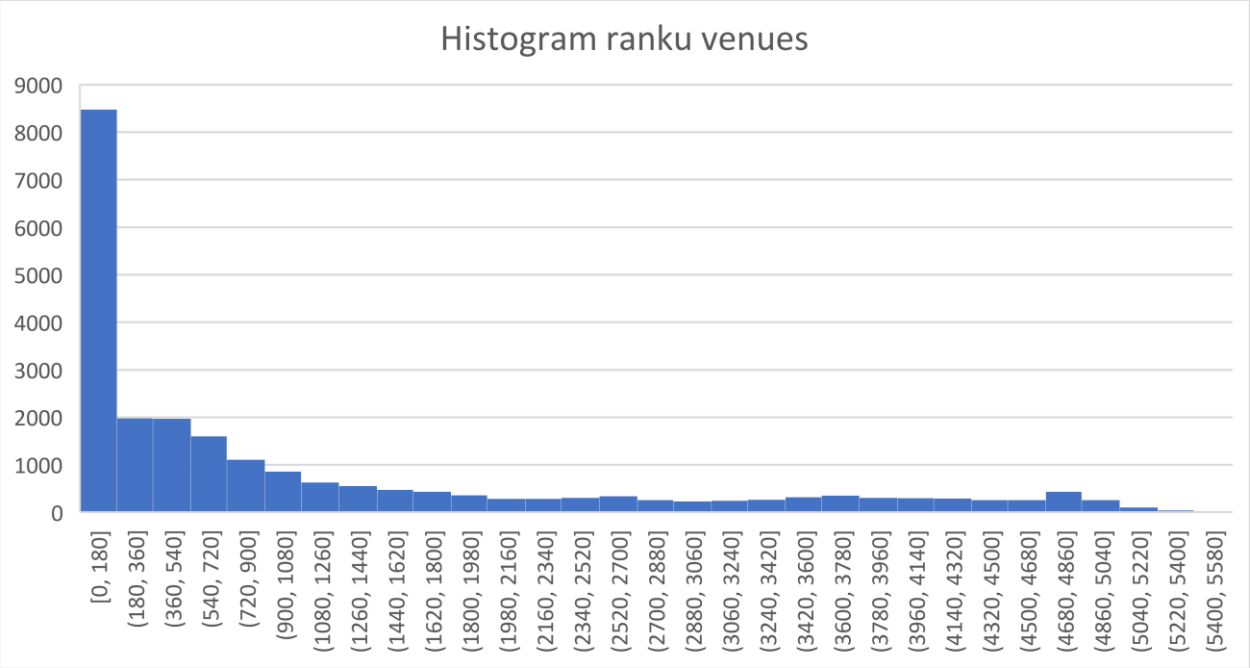
Pokud získané hodnoty umístíme na log-log graf, kde na jedné ose jsou hodnoty cest a na druhé je hustota rozdělení pravděpodobnosti zjistíme, že rozdělení hodnot má lineární trend podobný tomu pravděpodobnostnímu. Toto pozorování naznačuje, že pravděpodobnost pohybu na určité místo klesá, když počet míst bližších, než je cílové, stoupá.

Noulas a kolektiv pozorovali tento jev na všech jimi zkoumaných městech, jejich společný exponent byl $\alpha = 0.84 \pm 0.07$. Z toho usoudili, že existuje univerzální vzor pohybu napříč testovanými městy, kde hustota míst je řídícím faktorem lidské mobility. (Noulas, 2012)

Tento výsledek lze přirovnat k výsledkům ze studie *Geographic routing in social networks*, ze které vyplynulo, že pravděpodobnost nalezení uživatelova přítele v určité

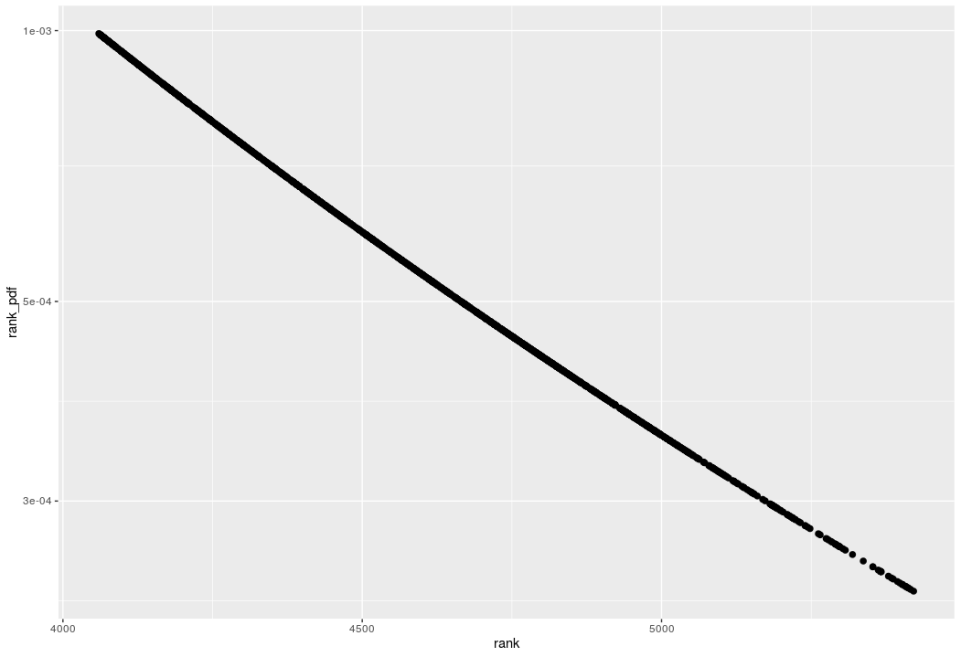
[11] https://cran.r-project.org/web/packages/RANN/index.html

vzdálenosti v geografické sociální síti je nepřímo úměrná počtu osob geograficky blíže uživateli. (Liben-Nowell, 2005)



Graf 4: Histogram ranku *venues* z Foursquare datasetu. Zdroj: vlastní

Z námi zkoumaného datasetu vyplynul exponent silového zákona ranku $\alpha = 1{,}03$. Jeho hodnota je tak blízká hodnotám ze studie *A Tale of Many Cities,* že můžeme usoudit, že i na základě něho jde vytvářet a modelovat vzory pohybu obyvatel. Tento exponent jsme ověřili pomocí Kolmogorovo–Smirnově testu. Námi zjištěný exponent je vyšší než 1 a tak lépe definuje funkci silového zákona. (Clauset, 2009)



Graf 5: Graf vztahu Rank a funkce hustoty pravděpodobnosti dat z aplikace Foursquare. Zdroj: vlastní

5.2.5 Porovnání gravitačního modelu

Kromě zkoumání vztahu mobility, ranku a vzdálenosti se Noulas a kolektiv pokusili vytvořit také gravitační model pohybu obyvatel ve městech. Rozhodli se k tomu ze dvou důvodů, za prvé kvůli atraktivnosti a přitažlivosti jednotlivých míst a za druhé z důvodu obavy o efekt větší vzdálenosti na samotný pohyb. Druhý důvod je určen přímo tou vzdáleností mezi body trasy, který značíme v tomto případě $\mathrm{d(u, v)}$, mezi body u a v. Pro první důvod je třeba zjistit gravitační masu obou bodů trasy. Tu jsme změřili tak, že se určil rádius r a sečetly se všechny *venues,* které byly v tomto rádiusu v okolí bodů trasy. Čím hustější oblast, ve které se *venue* nacházela, tím více jsme jich nalezli a tím byla *venue* atraktivnější. Na základě těchto dat jsme zjistili masu m každého bodu. Výpočet pro gravitační model je poté definován tímto vzorcem

$$\mathrm{P}_g[u \to v] \propto \frac{m_u . m_v}{d(u,v)^b},$$

kromě rádiusu bylo třeba zvolit ještě exponent b, který kontroluje vliv vzdálenosti na výsledek výpočtu. Ten ve studii *A Tale of Many Cities* vyzkoušeli v rozsahu 0,5 až 2,5 a nakonec zvolili 1. Samotný rádius zkoušeli také ve větším rozsahu a nakonec použili 100 metrů. V našem případě bylo lepší, vzhledem k výrazně nižší hustotě *venues,* zvolit vyšší

hodnotu, například 1 kilometr. Pro exponent vzdálenosti byla po několika pokusech zvolena hodnota 1,5, protože jen tak bylo možné srovnat hodnoty ranku a gravitačního modelu. Na výpočet byla opět použita knihovna *RANN*.

Již ze změřených hodnot, ale i z následné vizualizace bylo zřejmé, že gravitační model není popsatelný v kratších vzdálenostech, a tak není možné ho využít pro obecné popsání. Noules a kolektiv navíc argumentují, že jeho použití je třeba upravit pro každé město a rozdíly mezi velkoměsty a maloměsty budou příliš velké.

5.3 Metoda shlukové analýzy

Postup zvolený v metodě výzkumu *A Tale of Prague* zkoumá všechny uživatele a nehledí na konkrétní kategorie *venues* – míst, kde se označují. Tím mohou vznikat nepřesnosti ve výsledcích, pokud se například lidé označují jen na místech spadajících do podobných kategorií. Abychom předešli tomuto problému bylo třeba udělat shlukovou analýzu kategorií označení.

První pokus byl za pomoci R, knihovny *Cluster* [12]a její funkce *Gower*[13]. Ta umožňuje vytvářet hierarchické shlukové analýzy na základě kategorií. Nejprve byly zjištěny jednotlivé kategorie navštívených *venues* a poté vytvořena matice, kde na jedné straně byl seznam kategorií – dohromady 337 položek a na druhé seznam uživatelů. Obsah matice byl poměr jejich označení v dané kategorii. Pokud se uživatel označoval pouze v nemocnicích, měl ve sloupci *Hospital* číslo jedna a ve zbylých sloupcích nula. Bohužel výsledky za pomoci knihovny *Cluster* nebyly vyhovující. Knihovna vytvářela shluky, které obsahovaly jen 1 kategorii.

Další zvolený přístup využil API služby Foursquare. Nejprve bylo třeba stáhnout seznam všech kategorií *venues* v jejich přesné hierarchii. Ta se skládá z 10 základních kategorií, které zahrnují zbývajících 327 podkategorií. Dále na základě těchto dat připravit novou matici, která obsahovala již pouze 10 sloupců, a to se základními kategoriemi a u nich poměry zastoupení označení uživatele.

[12] https://cran.r-project.org/web/packages/cluster/index.html
[13] https://www.rdocumentation.org/packages/cluster/versions/2.1.0/topics/daisy

userid	cat_arts	cat_college	cat_event	cat_outdoor	cat_professional	cat_residence	cat_shop	cat_travel	cat_nightlife	cat_food
184	0,25	0	0	0,25	0,083333	0	0	0,166667	0	0,166667
397	0	0	0	0,2	0	0	0	0,2	0	0,6
450	0,12	0	0	0,16	0,08	0	0,04	0,2	0	0,16
517	0	0	0	0	0	0	0	1	0	0
539	1	0	0	0	0	0	0	0	0	0
689	0	0	0	0	0	0	0	1	0	0
755	0,125	0	0	0,375	0	0	0,125	0,125	0,125	0
813	0	0	0	0	0	0	0	1	0	0
1123	0,014866	0,000991	0	0,185332	0,042616	0,080278	0,048563	0,120912	0,020813	0,199207
1156	0,111111	0	0	0	0,111111	0	0,111111	0,222222	0,222222	0,222222

Tabulka 4: Ukázka shlukové analýzy matice kategorií *venues* a uživatelů a poměry jejich označení v daných kategoriích. Zdroj: vlastní

Na tuto matici byl poté použit algoritmus shlukové analýzy *k-means* – k průměru – ze stejnojmenné knihovny. Tomuto algoritmu byl předán parametr *k* s počtem požadovaných shluků. Byla zvolena hodnota 6 a na základě té vzniklo 6 skupin uživatelů.

- Nejpočetnější skupina *Arts & Entertainment* se 731 uživateli, kteří provedli dohromady 18 098 označení na území Hlavního města Prahy.
- Další skupina byla *Travel & Transport*, ve které bylo 464 uživatelů, kteří se označili na 1 017 místech.
- Třetí skupinou bylo 281 uživatelů, kteří se 1 034krát označili na *venues* spadající do kategorie *Food*.
- Čtvrtá skupina *Outdoor & Recreation* s 199 uživateli a 554 označeními.
- Předposlední *Nightlife Spot* s 151 uživateli a 346 označeními.
- A konečně poslední *Shop & Service*, ve které bylo jen 67 uživatelů, ale udělali 1 738 označení.

V základních kategoriích *College & University*, *Event*, *Professional & Other places* a *Residence* neměla žádná skupina uživatelů dostatek označení, aby to vytvořilo samostatný shluk.

V následujících kapitolách je rozbor jednotlivých shluků – klasterů. Pro každou skupina byla zopakována stejná analýza jako na celek, za účelem zjištění, zda se teorie intervenčních příležitostí neobjevuje pouze v určité skupině lidí.

5.3.1 Arts & Entertainment – Kulturní zábava

Do této skupiny mimo jiné patří kategorie muzea, kina, divadla, kluby, zoo či stadiony. Vzhledem k velikosti skupiny bylo očekáváno, že budou výsledky podobné jako u celku. Z poměru označení a uživatelů je zřejmé, že tito lidé jsou aktivními uživateli Foursquare.

Nejdelší naměřená trasa mezi dvěma *check-iny* byla také 30 590 metrů. Pokud prozkoumáme kvantily těchto vzdáleností, zjistíme, že dolní kvantil je na 592 metrech a horní na 5 568 metrech. Průměr všech vzdáleností byl sice 3 688 metrů, ale medián pouze 2 042. Je tedy jasné, že většina pohybu je na krátké trase do 2 kilometrů. Tyto výsledky jsou tedy velmi podobné celkovým výsledkům.

Pomocí knihovny *poweRlaw* jsme zjistili exponent vzdálenosti α = 4,394891 a $x_{min} =$ 14 440. Exponent ranku jsme zjistili, že se rovná α = 6.98256 a jeho $x_{min} = 271$.

5.3.2 Travel & Transport – Doprava a cestování

Dopravní shluk zahrnuje například kategorie letiště, zastávky, spoje, taxi či hotely a hostely. V této skupině tak bude více turistů. To se projevilo na průměrné délce trasy, která byla 4 557 metrů a medián byl 2 189 metrů. Nejdelší trasa byla skoro 29 kilometrů, a to z letiště na jihovýchodní okraj Prahy. Kvantily byly na 388 a na 9 265 metrech. Zde je vidět odlišnost oproti celku a první skupině, kdy je zřejmé, že tito lidé cestují delší vzdálenosti.

Dále jsme zjistili exponent silového zákonu pro vzdálenost α = 4,261106 a $x_{min} =$ 12 429. Exponent ranku vychází na α = 2.024882 a $x_{min} = 96$.

5.3.3 Food – Jídlo a restaurace

Třetí skupina s ale druhým nejvyšším počtem označení byla stravovací. V tomto shluku je také nejvíce subkategorií, okolo 350, ale v Praze jich najdeme méně. Do této kategorie totiž patří různé národnostní kuchyně, kterých jsou desítky. Dále sem patří fast foody, kavárny, food courty či pekárny. Lidé, kteří se nejvíce označují v této kategorii průměrně nacestovali 3 093

metrů a medián jejich tras byl 1 586 metrů. Kvantily byly 509 a 3 806 metrů a nejdelší trasa téměř 19 kilometrů. Tato skupina má oproti celku nižší průměrnou trasu, ale jinak jsou hodnoty srovnatelné.

Exponent silového zákonu vzdáleností byl u kategorie Jídla α = 1,921371 a $x_{min} = 1221$. Silový zákon ranku je exponent α = 3.059494 a $x_{min} = 406$.

5.3.4 Outdoor & Recreation – Příroda a rekreace

Shluk rekreace obsahuje mimo jiné posilovny, různá sportoviště a hřiště, hrady a zámky a přírodní místa jako řeky, lesy, pláže, jezera či stromy. Průměrná trasa byla 3 605 metrů a medián 1 703 metrů. Kvantily této skupiny byly 576 a 3 619 metrů a nejdelší trasa 22,5 kilometru. Zde opět vzrostla, a to nejspíše většími vzdálenostmi samotných míst.

Zjištěný exponent vzdálenosti byl α = 3.059494 a $x_{min} = 406$. A pro rank byl exponent α = 3.059494 a $x_{min} = 406$.

5.3.5 Nightlife Spot – Noční život

Předpolední shluk zahrnuje uživatele co nejvíce navštěvovali *venues* patřící do nočního života. Byly to například různé bary, pivovary, hospody a noční či strip kluby.

Průměrná vzdálenost je 3159.15 metrů a nejdelší 16 580 metrů. Střední hodnota vyšla na 1369 metrů a kvantily na 591 a 3 652 metrů.

Z knihovny *poweRlaw* jsme zjistili exponent vzdálenosti α = 1.77004 a jeho $x_{min} = 790$. A také tyto hodnoty pro rank α = 1,986392 a $x_{min} = 55$.

5.3.6 Shop & Service – Obchody a služby

Poslední shluk je složen z lidí, co se nejvíce označuji v obchodech a různých službách. Mezi ty mimo jiné patří bankomaty, nákupní centra, tržiště, řeznictví, knihkupectví, banky a mnoho dalších, po jídle je toto druhá největší rodičovská kategorie.

Nejdelší trasa byla přes 24 kilometrů a průměrná 4463.136 metrů. Kvantily odpovídaly 530 a 6 969 metrů a střední hodnota 3 071 metrů.

Nalezli jsme exponent vzdálenosti α = 4.397543 a x_{min} = 11 452. A také exponent ranku α = 5.612412 a x_{min} = 421.

5.3.7 Výsledky výzkumu shlukové analýzy

Rozdělením uživatelů podle jejich nejčastěji navštěvované kategorie *venues* jsme zjistili, že navzdory tomu, že celkový dataset lze pomocí ranku definovat u jednotlivých *klasterů* to není již tak jednoznačné. Pozorovali jsme, že se například pohyby lidí, kteří primárně označují svou polohu na místech, které spadají do kategorie Obchody a služba a do kategorie Kulturní zábava, nedají dobře popsat definovat teorií intervenujících příležitostí, respektive silovým zákonem.

Naproti tomu ostatní jsou dobře popisné, především lidé navštěvující hlavně *Nighlife Spots*. U této skupiny je jasně patrné, že vzdálenost pro ně není tolik rozhodující a upřednostňují konkrétní, i klidně vzdálená, místa.

5.4 Porovnání výzkumu s daty z aplikace Rekola

Aby bylo možné zkontrolovat výsledky, které jsme získali z výzkumu dat z geolokační aplikace Foursquare, porovnali jsme je se sadou dat jízd z aplikace Rekola.

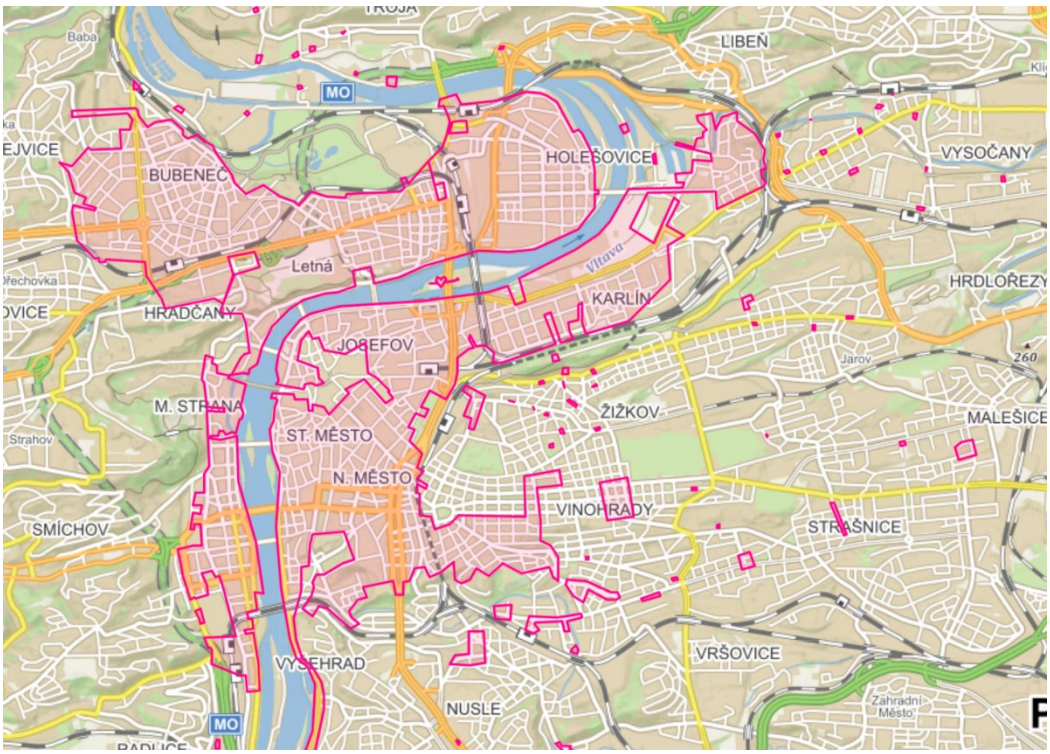
5.4.1 Rekola

Rekola je český projekt sdílení jízdních kol – takzvaný Bikesharing. Založili ho v Praze v roce 2013 bratři Filip a Vít Ježkovi jako první *bikesharing*[14] v České republice. Systém repasovaných růžových kol se v následujícím roce rozšířil do Brna a Olomouce. V dalších letech se rozšířil mezi desetitisíce uživatelů a do dalších měst. V současné době také působí v Českých Budějovicích, Frýdku-Místku, Kladně, Liberci, Mladé Boleslaví, Ostravě, Písku a ve finském Vaasa. (Rekola, 2019)

Aktuálně je v systému přes 2000 kol a na podzim 2019 v něm bylo zaregistrováno přes 220 tisíc uživatelů. (Rekola, 2019) Podle statistiky v roce 2019 je průměrnému uživateli Rekol v Praze mezi dvaceti až třiceti lety, přičemž z této skupiny Rekola v životě vyzkoušel každý

[14] Sdílení kol

2,7. člověk. (Brejčák, 2019) V sezóně 2018 dosáhl systém počtu 761 731 výpůjček. Kola se nemusí vracet do stojanů a uživatelé je mohou nechat, kdekoliv zamknuté v zóně, kterou vidí v mobilní aplikaci. (Rekola, 2019)



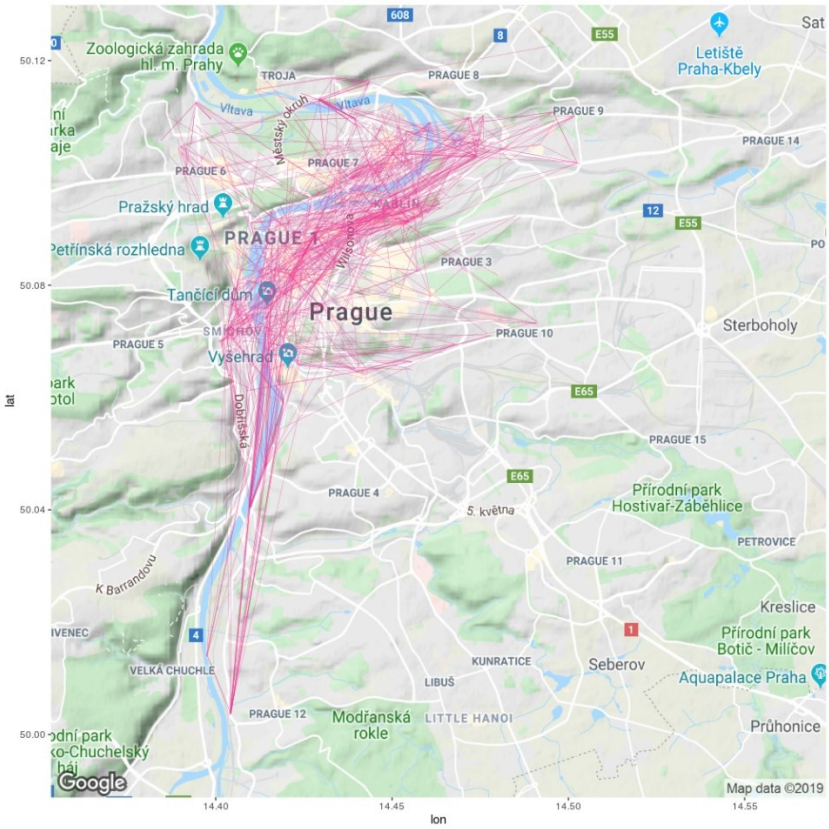
Obrázek 7: Mapa zón v Praze bikesharingového projektu Rekola. Zdroj: (Rekola, 2019)

5.4.2 Dataset Rekola

Zkoumaný dataset pochází přímo od firmy Rekola Bikesharing s.r.o., ta předává svá anonymizovaná data vedením měst a městských částí jako jsou například IPR v Praze či BKom v Brně. Pro tento výzkum byl speciálně poskytnutý dataset z roku v 2018, který obsahuje seznam jízd od 1. června do 31. července na území Hlavního města Prahy. Seznam obsahuje pro každou jízdu anonymizované ID uživatele, ID jízdy, datum a čas začátku jízdy, datum a čas konce jízdy a zeměpisné šířky a délky polohy startu a konce jízdy. Jednotlivé jízdy mají anonymní ID uživatele, aby nebylo možné spárovat více jízd, a tak například určit polohu jeho bydliště či zaměstnání. V získaném období proběhlo na území Prahy 118 217 výpůjček.

id	loan_id	started_at	arrived_at	latitude_start	longitude_start	latitude_end	longitude_end
6848444	490041	01.06.2018 0:00	01.06.2018 0:07	50.069503784179	14.467863082885	50.068653106689	14.457106590270
6848458	490042	01.06.2018 0:02	01.06.2018 1:05	50.095458984375	14.427111625671	50.058990478515	14.419227600097
6848486	490044	01.06.2018 0:10	01.06.2018 0:38	50.102779388427	14.473775863647	50.085441589355	14.427412033081
6848500	490045	01.06.2018 0:11	01.06.2018 0:26	50.073177337646	14.491091728210	50.074317932128	14.438890457153
6848514	490046	01.06.2018 0:11	01.06.2018 0:39	50.102779388427	14.473775863647	50.085491180419	14.427758216857
6848528	490047	01.06.2018 0:11	01.06.2018 0:37	50.089550018310	14.422050476074	50.084121704101	14.414181709289

Tabulka 5: Ukázka datasetu Rekola. Zdroj: (Rekola, 2019)



Obrázek 8: Ukázka 1000 náhodných jízd z datasetu Rekola. Zdroj: vlastní

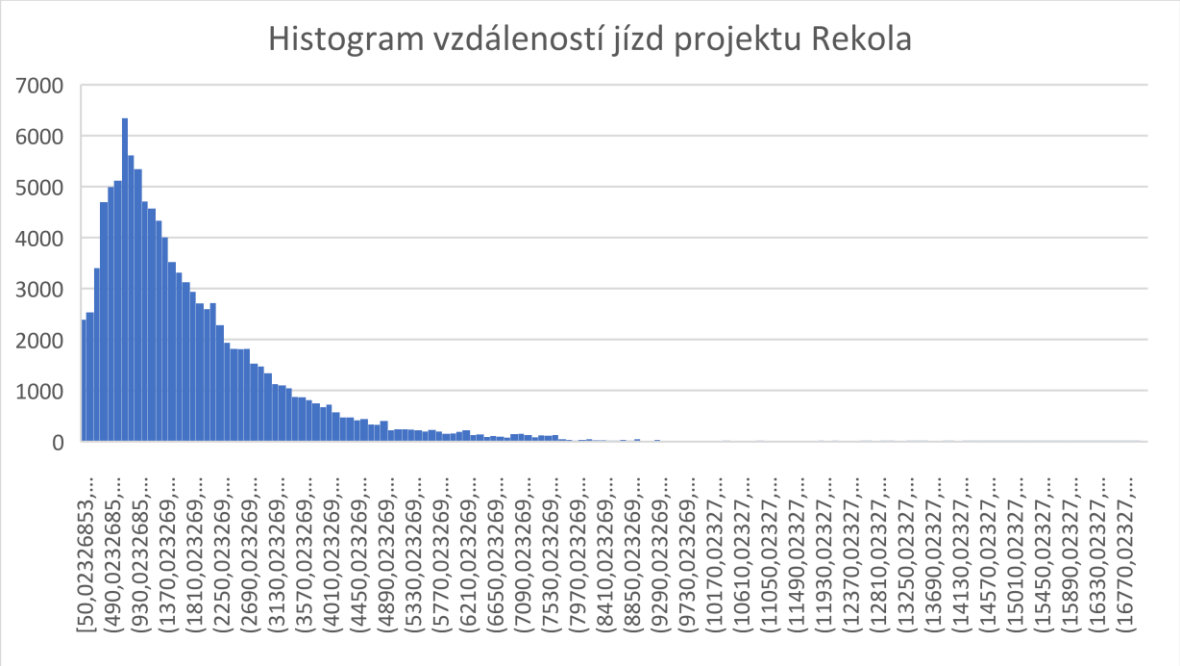
5.4.3 Výzkum

Stejným výpočtem jako u Foursquare datasetu byla zjištěna průměrná vzdálenost trasy, přesněji vzdálenost mezi počátečním a cílovým bodem, protože o průběhu trasy nemáme data.

Poté byla data začištěna o jízdy, které urazily méně než 50 metrů, protože takové jízdy pro nás nemají výpovědní hodnotu, poté byly odebrány jízdy, které pravděpodobně byly pouze vypůjčení a okamžité vrácení kola. Vzorek se tak zmenšil na 109 152 výpůjček. Průměrná jízda dosáhla 1 812,57 metrů a její průměrný čas byl 24 minuty. Největší vzdálenost trasy byla 17,175 km a nejdelší výpůjčka trvala přes 20 hodin.

Pokud prozkoumáme kvantily vzdáleností, zjistíme, že dolní kvantil je na 788 metrech a horní na 2 393 metrech, to vidíme i na histogramu, kde jsou vidět prvky mocninového rozdělení. Medián všech vzdáleností byl pouze 1 384 metrů. Je tedy jasné, že většina pohybu i z těchto dat, tedy jízd, je na krátké trase a to do 1,5 kilometru.

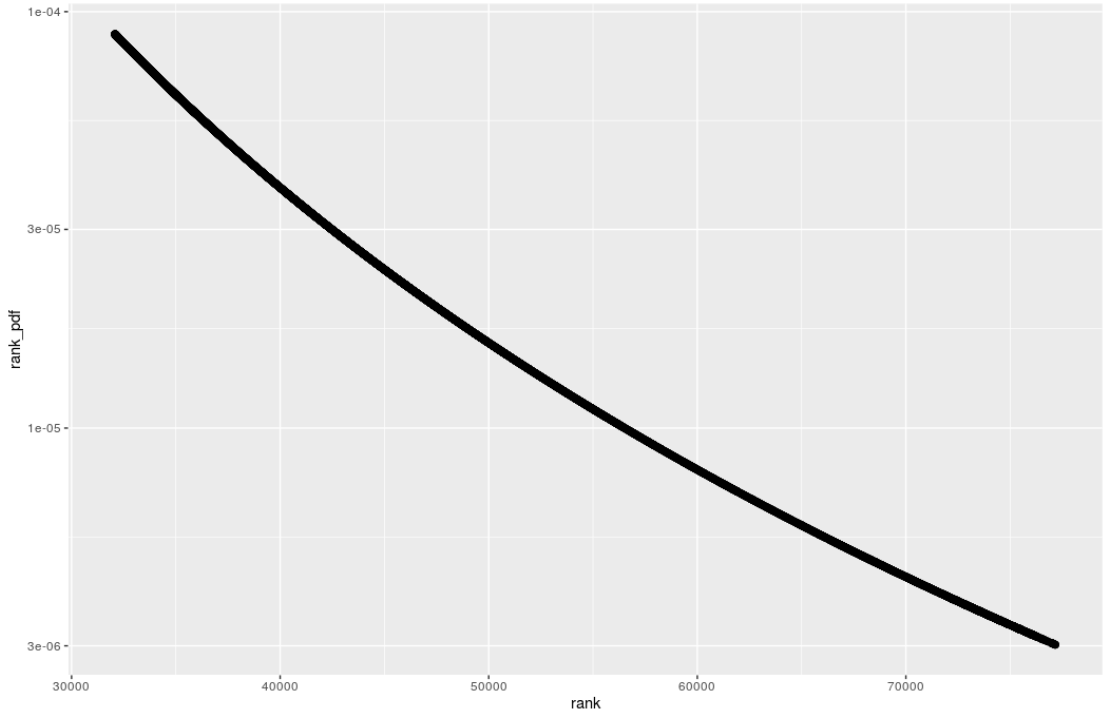
Pomocí knihovny *poweRlaw* jsme zjistili exponent hustoty pravděpodobnosti vzdálenosti α = 4,029619 a x_{min} = 3 565.



Graf 6: Histogram vzdáleností jízd projektu Rekola. Zdroj: vlastní

Po vypočtení vzdáleností tras byl dopočítán také stejným způsobem rank jednotlivých míst. Ten byl nejvyšší 77 162 a průměr 13 247 a kvantily 2 506 a 18 585. Střední hodnota jsme naměřili 6 928.

Dále jsme zjistili exponent hustoty pravděpodobnosti pro rank, ten dle knihovny *poweRlaw* vyšel na α = 1,77004 a jeho $x_{min} = 790$. Tato hodnota je také přijatelná a v akceptovatelných mezích, a tak je možné na tento dataset aplikovat teorii intervenujících příležitostí. Vyplývá tak, že pro uživatele toho projektu není tolik důležitá vzdálenost, jako samotná destinace a její okolí.



Graf 7: Graf vztahu Rank a funkce hustoty pravděpodobnosti dat z projektu Rekola. Zdroj: vlastní

6 Závěr

Cílem této práce bylo zjistit, zda se dá určit obecný vzorec popisující pohyb obyvatel ve městech. A pokud ano, jestli je v současné době stále aktuální využívat gravitační modely postavené na studiích z 19. století či jestli současná doba nenabízí lepší metody pro tuto analýzu. Nejprve byla představena mobilita a její různé druhy. Poté varianty výzkumů vzorů pohybu a jejich modelování.

Poté proběhl samotný vlastní výzkum zaměřený na Hlavní město Prahu, který primárně vycházel ze studie provedené nad desítkami amerických měst a nad jiným datasetem. V této práci byl nejprve zkoumán dataset z geolokační sociální sítě Foursquare, který zahrnoval osmnáct měsíců *check-inů* uživatelů. Bylo nad ním provedeno několik operací a porovnání jak různých metod, tak i výsledků s jinými studiemi. Nakonec bylo zjištěno, že na základě těchto dat lze najít vzor pohybu postavený na Teorii intervenujících příležitostí. Byl také proveden výzkum gravitačního modelu, který se ukázal, že by nejspíše fungoval jen na větší města a jen na delší vzdáleností, tudíž ho nejde považovat za obecný vzor pohybu obyvatel ve městech. Výsledek byl poté ještě otestován na menších klastrech, které ukázaly, že při konkrétnějších datech nemusí model vždy odpovídat. Poslední částí byla analýza dat z bikesharingového projektu Rekola, který poskytl data ze dvou měsíců sezóny 2018. Výsledkem této analýzy byl stejný výsledek jako u Foursquare dat, tedy že je možné v těchto datech najít vzory dle Teorie intervenujících příležitostí.

Odpovědí na základní otázku práce je tedy, že pokud chceme dnes získat co nejpřesnější a nejuniverzálnější modely a vzory pohybu obyvatel ve městech, tak máme vycházet z jiných dat než ze vzdálenosti, například z ranku na základě teorie intervenujících příležitostí.

7 Bibliografie

30 Foursquare Statistics to Help *You* Optimize the Platform in 2019, 2019. In: *99firms.com* [online]. -: 99firms.com [cit. 2019-04-21]. Dostupné z: https://99firms.com/blog/foursquare-statistics/

ADEY, Peter, 2010. Mobility. 1. New York: Routledge. ISBN 9780415433990.

ANDERSON, James, 2011. The Gravity Model. Annual Review of Economics [online]. 3(1), 133-160 [cit. 2019-04-25]. DOI: 10.1146/annurev-economics-111809-125114. ISSN 1941-1383. Dostupné z: http://www.annualreviews.org/doi/10.1146/annurev-economics-111809-125114

ARYA, Sunil, David MOUNT, Samuel KEMP a Gregory JEFFERIS, 2019. RANN: Fast Nearest Neighbour Search (Wraps ANN Library) Using L2 Metric. *R: The R Project for Statistical Computing* [online]. -: The R Foundation [cit. 2019-05-02]. Dostupné z: https://CRAN.R-project.org/package=poweRlaw

BOŘIL, Tomáš, 2015. Testování rozdělení. In: Statistika ve fonetickém výzkumu [online]. Praha: FF UK [cit. 2019-12-16]**.** Dostupné z: https://fu.ff.cuni.cz/STAT/14_testy_rozdeleni.html

BREJČÁK, Peter, 2019. Vítek Ježek (Rekola): „V budoucnu vznikne další evoluční stupeň nad elektrokoly, chceme být při tom“. *StartupJobs.cz*. Praha: StartupJobs.cz. Dostupné také z: https://www.startupjobs.cz/newsroom/rozhovor-vitek-jezek-rekola

BROCKMANN, D., L. HUFNAGEL a T. GEISEL, 2006. The scaling laws of human travel. *Nature* [online]. **439**(7075), 462-465 [cit. 2019-04-20]. DOI: 10.1038/nature04292. ISSN 0028-0836. Dostupné z: http://www.nature.com/articles/nature04292

CLAUSET, Aaron, Cosma SHALIZI a M. *NEWMAN, 2009. Power-Law Distributions in Empirical* Data. SIAM Review [online]. 51*(*4), 661-703 **[**cit. 2019-04-28]. DOI: 10.1137/070710111. ISSN 0036-1445. Dostupné z: http://epubs.siam.org/doi/10.1137/070710111

COMITO, Carmela, Deborah FALCONE a Domenico TALIA, 2016. Mining human mobility patterns from social geo-tagged data. *Pervasive and Mobile Computing* [online]. **33**(-

), 91-107 [cit. 2019-05-01]. DOI: 10.1016/j.pmcj.2016.06.005. ISSN 15741192. Dostupné z: https://linkinghub.elsevier.com/retrieve/pii/S1574119216300700

CRESSWELL, Tim, 2006. On the move: *mobility in the modern Western world.* New York: Routledge. ISBN 978-0-415-952-56-9.

CRESSWELL, Tim a Peter MERRIMAN, 2011. Geographies of mobilities: practices, spaces, subjects. 1. Burlington, VT: Ashgate. ISBN 14-094-1782-4.

CSÁRDI, Gábor, 2019. Igraph: Network Analysis and Visualization. *R: The R Project for Statistical Computing* [online]. -: The R Foundation [cit. 2019-04-28]. Dostupné z: https://CRAN.R-project.org/package=igraph

EUBANK, Stephen, Hasan GUCLU, V. ANIL *KUMAR, Madhav MARATHE, Aravind SRINIVASAN,* Zoltán TOROCZKAI a Nan WANG, 2004. Modelling disease outbreaks in realistic urban social networks. *Nature* [online]. **429**(6988), 180-184 [cit. 2019-04-20]. DOI: 10.1038/nature02541. ISSN 0028-0836. Dostupné z: http://www.nature.com/articles/nature02541

Foursquare [online], 2019. NYC, SF *& Chicago: Foursquare [cit. 2019-04-15].* Dostupné z: https://foursquare.com

FREUDENDAL-PEDERSEN, Malene, 2009. *Mobility in daily life: between freedom and unfreedom*. 1. Burlington, VT: Ashgate. Transport and society. ISBN 978-0-7546-7490-0.

GILLESPIE, Colin, 2019. *PoweRlaw:* Analysis of Heavy Tailed Distributions. R: The R Project for Statistical Computing [online]. -: The R Foundation [cit. 2019-04-28]. Dostupné z: https*.*//CRAN.R-project**.org**/package=poweRlaw

GONZÁLEZ, Marta, César HIDALGO a Albert-László BARABÁSI, 2008. Understanding individual human mobility patterns. Nature [online]. 453(7196), 779-782 [cit. 2019-04-19]. DOI: 10.1038/nature06958. ISSN 0028-0836**.** Dostupné z: http://www.nature.com/doifinder/10.1038/nature06958

GREGORY, Derek, 2009. The dictionary of human geography. 5th ed. Malden, MA: Blackwell. ISBN 978-1-4051-3288-6.

HAGERSTRAND, Torsten, 1982. *Diorama, path and project. Tijdschrift voor Economische* en Sociale Geografie [online]. 73(6), 323-339 [cit. 2019-04-15]. DOI:

10.1111/j.1467-9663.1982.tb01647.x. ISSN 0040-747X. Dostupné z: http://doi.wiley.com/10.1111/j.1467-9663.1982.tb01647.x

HÄGERSTRAND, Torsten, 1970. *What about People in Regional Science?. Papers* in Regional Science. 1970(24), 6–21. ISSN 1056-8190.

HALLIN, Per, 2017. New Paths for Time-Geography?. *Geografiska Annaler: Series B, Human Geography* [online]. **73**(3), 199-207 [cit. 2019-04-14]. DOI: 10.1080/04353684.1991.11879625. ISSN 0435-3684. Dostupné z: https://www.tandfonline.com/doi/full/10.1080/04353684.1991.11879625

HIJMANS, Robert, Ed WILLIAMS *a* Chris VENNES, 2019. Geosphere: Spherical Trigonometry. R: The R Project for Statistical Computing [**online**]. -: The R Foundation [cit. 2019-04-28]. Dostupné z: https://CRAN.R-project.org/package=geosphere

Integrovaná strategie pro ITI Pražské metropolitní oblasti, 2016. In: Institut plánování a rozvoje hlavního *města* Prahy [online]**.** Praha 2: Institut plánování a rozvoje hlavního města Prahy [cit. 2019-04-26]. Dostupné z: http://www.iprpraha.cz/uploads/assets/dokumenty/ssp/ITI/integrovana-strategie-pro-iti-prazske-metropolitni-oblasti_verze_1_5.pdf

ISAACMAN, Sibren, Richard BECKER, Ramón CÁCERES, Stephen KOBOUROV, James ROWLAND a Alexander VARSHAVSKY, 2010. A tale of two cities. In: *Proceedings of the Eleventh Workshop on Mobile Computing Systems & Applications - HotMobile '10* [online]. New York, New York, USA: ACM Press, s. 19- [cit. 2019-04-21]. DOI: 10.1145/1734583.1734589. ISBN 9781450300056. Dostupné z: http://portal.acm.org/citation.cfm?doid=1734583.1734589

KAHLE, David a Hadley WICKHAM, 2013. Ggmap: Spatial Visualization with ggplot2. KAHLE, David a Hadley WICKHAM. The R Journal [online]. 5(1) [cit. 2019-12-16]. Dostupné z: https://journal.r-project.org/archive/2013-1/kahle-wickham.pdf

KANG, Chaogui, Song GAO, Xing LIN, Yu XIAO, Yihong YUAN, Yu LIU a Xiujun MA, 2010. Analyzing and geo-visualizing individual human mobility patterns using mobile call records. In: *2010 18th International Conference on Geoinformatics* [online]. Beijing, China: IEEE, s. 1-7 [cit. 2019-05-02]. DOI: 10.1109/GEOINFORMATICS.2010.5567857. ISBN 978-1-4244-7301-4. ISSN 2161-0258. Dostupné z: http://ieeexplore.ieee.org/document/5567857/

KAUFMANN, Vincent, 2002. Re-thinking *mobility: contemporary sociology.* 1. Burlington, VT: Ashgate. ISBN 9780754618423.

KELLERMAN, Aharon, 2006. Personal mobilities. 1. New York: Routledge, Taylor & Francis Group. ISBN 0415391598.

KLEINBERG, Jon, 2007. Computing: The wireless epidemic. Nature [online]. 449(7160), 287-288 [cit. 2019-04-20]. DOI: 10.1038/449287a. ISSN 0028-0836. Dostupné z: http://www.nature.com/doifinder/10.1038/449287a

KNOX, Paul a Steven PINCH, 2010. Urban Social Geography: *An Introduction*. 6. Velká Británie: Pearson Education Limited. ISBN 978-0-273-71763-8.

LARSEN, Jonas, John URRY *a K. AXHAUSEN,* 2006. Mobilities, networks, geographies. 1. Burlington, VT: Ashgate. ISBN 0-7546-4882-6.

LETOUZÉ, *Emmanuel, Mark PURSER, Francisco RODRÍGUEZ a Matthew CUMMINS, 2009. Revisiting the* Migration-Development Nexus: A Gravity Model Approach. *Human Development Research Paper (HDRP) Series* [online]. **2009**(44), 1–42 [cit. 2019-04-25]. Dostupné z: https://mpra.ub.uni-muenchen.de/19227/

LIBEN-NOWELL, D., J. NOVAK, R. KUMAR, P. *RAGHAVAN a A. TOMKINS, 2005.* Geographic routing in social networks. Proceedings of the National Academy of Sciences [online]. 102(33), 11623-11628 [cit. 2019-04-30]. DOI: 10.1073/pnas.0503018102. ISSN 0027-8424. Dostupné z: http://www.pnas.org/cgi/doi/10.1073/pnas.0503018102

LIU, Yu, Zhengwei SUI, Chaogui KANG, Yong GAO a Peter CSERMELY, 2014. Uncovering Patterns of Inter-Urban Trip and Spatial Interaction from Social Media Check-In *Data. PLoS ONE [online]. 9(1), 1–11* [cit. 2019-05-02]. DOI: 10.1371/journal.pone.0086026. ISSN 1932-6203. Dostupné z: https://dx.plos.org/10.1371/journal.pone.0086026

MACHOVÁ, Barbora, 2015. Žitá každodennost obyvatel horského Poreče v Makedonii. Brno. Disertační práce. Masarykova univerzita, Filozofická fakulta.

MANDELBROT, Benoit, 1982. The fractal geometry of nature. 1. San *Francisco: W.H.* **Freeman**. ISBN isbn0-7167-1186-9.

MCGOWEN, Patrick a Michael MCNALLY, 2007. Evaluating the potential to predict *activity types from GPS and GIS* data. In: TRB 86th Annual Meeting Compendium of Papers

[CD-ROM]. Washington, DC, USA: Transportation Research Board 86th Annual Meeting, s. 22 [cit. 2019-05-02]. ISBN -. ISSN -.

Measuring *Personal Travel and Goods Movement* [online], 2003. 1. Washington, D.C: National Academies Press [cit. 2019-05-02]. DOI: 10.17226/10837. ISBN 978-0-309-08599-1. Dostupné z: https://www.nap.edu/catalog/10837/measuring-personal-travel-and-goods-movement-a-review-of-the

MULÍČEK, Ondřej, Robert OSMAN *a* Daniel SEIDENGLANZ, 2013. Imaginace a reprezentace prostoru v každodenní zkušenosti. Sociologický *časopis*. Praha: Sociologický ústav AV ČR, **49**(5), 781–810. ISSN 0038-0288.

NEWBERRY, Mitchell G. a *Van* M. SAVAGE, 2019. Self-Similar Processes Follow a Power Law in Discrete Logarithmic Space. Physical Review Letters [online]. 122(15), *1 [*cit. 2019-12-15]. DOI: 10.1103/PhysRevLett.122.158303. ISSN 0031-9007. Dostupné z: https://link.aps.org/doi/10.1103/PhysRevLett.122.158303

NOULAS, Anastasios, Salvatore SCELLATO, Renaud LAMBIOTTE, Massimiliano PONTIL, Cecilia MASCOLO a Juan AñEL, 2012. A Tale of Many Cities: Universal Patterns in Human Urban Mobility. *PLoS ONE* [online]. **7**(5) [cit. 2019-04-20]. DOI: 10.1371/journal.pone.0037027. ISSN 1932-6203. Dostupné z: https://dx.plos.org/10.1371/journal.pone.0037027

NOULAS**,** Anastasios, Salvatore SCELLATO, Cecilia MASCOLO a Massimiliano *PONTIL*, 2011. An Empirical Study of Geographic User Activity Patterns in Foursquare. AAAI Publications, Fifth International AAAI Conference on Weblogs and Social Media [online]. The AAAI Press, **-**(-), 570–573 [cit. 2019-04-30]. Dostupné z: http://dblp.uni-trier.de/db/conf/icwsm/icwsm2011.html#NoulasSMP11

NOVÁK, Jakub a Luděk SÝKORA, 2016. A city in motion: time-space activity and mobility patterns of suburban inhabitants and the structuration of the spatial organization of the prague metropolitan area. *Geografiska Annaler: Series B, Human Geography* [online]. **89**(2), 147-168 [cit. 2019-04-30]. DOI: 10.1111/j.1468-0467.2007.00245.x. ISSN 0435-3684. Dostupné z: https://www.tandfonline.com/doi/full/10.1111/j.1468-0467.2007.00245.x

ONG, Pauline, 2014. Adaptive Cuckoo Search Algorithm for Unconstrained Optimization. The Scientific World Journal [online]. 2014, 1-8 [cit. 2019-04-20]. DOI:

10.1155/2014/943403. ISSN 2356-6140. Dostupné z: http://www.hindawi.com/journals/tswj/2014/943403/

PARKES, Don a N. *THRIFT, 1980. Times, spaces,* and places: a chronogeographic perspective. 1. New York: J. *Wiley. ISBN 978-0471276166.*

POOLEY, Colin, Jean TURNBULL a Mags ADAMS, 2005. A mobile century?: changes in everyday mobility in Britain in the twentieth century. 1. Burlington, VT: Ashgate. ISBN 978-0-7546-4181-0.

PRATT, Andy, 2016. Coordinating Employment, Transport and Housing in Cities: An Institutional Perspective. Urban Studies *[online*]. 33(8), *1357-1375 [cit. 2019-04-14]. DOI:* 10.1080/0042098966691**.** ISSN 0042-0980. Dostupné z: http://journals.sagepub.com/doi/10.1080/0042098966691

PRED, Allan, 1977. The Choreography of Existence: Comments on Hagerstrand's Time-Geography and Its Usefulness. *Economic Geography*. **53**(2), 207–221. DOI: 10.2307/142726. ISSN 00130095. Dostupné také z: https://www.jstor.org/stable/142726?origin=crossref

PRED, Allan, 1981. Social Reproduction and the Time-Geography of Everyday Life. Geografiska Annaler. Series B, *Human* Geography [online]. 63(1), 5–22 [cit*. 2019-04-15*]. DOI: **10**.2307/490994. ISSN 04353684. Dostupné z: http://www.jstor.org/stable/490994?origin=crossref

RAVENSTEIN, E., 1885. The Laws of Migration. Journal of the Statistical *Society of London* [online]. **48**(2), 167–235 [cit. 2019-04-25]. DOI: 10.2307/2979181. ISSN 09595341. Dostupné z: https://www.jstor.org/stable/2979181?origin=crossref

REKOLA, , 2019. Rekola. Rekola - *Sdílená kola pro* rychlé přesuny po městě. Dostupné také z: https://www.*rekola.cz/pribeh-rekol*

REKOLA, Facebook, 2019. Rekola. Facebook. Dostupné také z: https://www.facebook.com/RekolaEU/posts/1887575931345388

RICHARDSON, Anthony, Elizabeth AMPT a Arnim MEYBURG, 1996. Non Response in Household Travel Surveys. In: *Conference on household travel surveys: new concepts and research needs* [online]. 10. Washington, D.C, USA: National Academy Press, 79–114 [cit. 2019-05-02]. ISBN 978-0309062107. ISSN 1073-1652.

RUDINAC, Stevan, Jan ZAHÁLKA a *Marcel WORRING, 2017. Discovering* Geographic Regions in the City Using Social Multimedia and Open Data. AMSALEG, Laurent, ed., Gylfi Þór GUðMUNDSSON, ed., Cathal GURRIN, ed., Björn Þór JÓNSSON, ed. a Shin'ichi SATOH, ed., Laurent AMSALEG, Gylfi GUðMUNDSSON, Cathal GURRIN, Björn JÓNSSON, Shin'ichi SATOH. *MultiMedia Modeling* [online]. 10133. Cham: Springer International Publishing, s. 148-159 [cit. 2019-05-01]. Lecture Notes in Computer Science. DOI: 10.1007/978-3-319-51814-5_13. ISBN 978-3-319-51813-8. Dostupné z: http://link.springer.com/10.1007/978-3-319-51814-5_13

SCHNEIDER, C., V**.** BELIK, T. COURONNE, Z. SMOREDA a M. GONZALEZ, 2013. *Unravelling daily* human mobility motifs. Journal of The **Royal** Society Interface [online]. 10(84), 20130246-20130246 [cit. 2019-05-02]. DOI: 10.1098/rsif.2013.0246. ISSN 1742-5689. Dostupné z: http://rsif.royalsocietypublishing.org/cgi/doi/10.1098/rsif.2013.0246

SIMINI, Filippo, Marta GONZÁLEZ, *Amos* MARITAN a Albert-László BARABÁSI, 2012. A universal model for mobility and migration patterns. Nature [online]. 484(7392), 96-100 [cit. 2019-04-26]. DOI: 10.1038/nature10856. ISSN 0028-0836. Dostupné z: http://www.nature.com/doifinder/10.1038/nature10856

SONG, Chaoming, Tal KOREN, Pu WANG a Albert-László BARABÁSI, 2010. Modelling the scaling properties of human mobility. Nature Physics [online]. 6(10), 818-823 [cit. 2019-05-01*].* DOI: 10.1038/nphys1760. ISSN 1745-2473. Dostupné z: http://www.nature.com/articles/nphys1760

STEFANOULI, **Maria** a Serafeim POLYZOS, 2017. Gravity vs radiation model: two approaches on commuting in Greece. Transportation Research Procedia [online]. 24, 65-72 [cit. 2019-05-01]. DOI: *10.*1016/j.trpro.**2017**.05.069. ISSN 23521465. Dostupné z: https://linkinghub.elsevier.com/retrieve/pii/S2352146517303502

STEWART, John, 1950. The Development of Social Physics. American Journal *of Physics* [online]. **18**(5), 239-253 [cit. 2019-04-25]. DOI: 10.1119/1.1932559. ISSN 0002-9505. Dostupné z: http://aapt.scitation.org/doi/10.1119/1.1932559

STOUFFER, Samuel, 1940. Intervening Opportunities: A Theory Relating Mobility and *Distance. American Sociological* Review [online]. 5(6), 845–867 [cit. 2019-04-28]. DOI: 10.2307/2084520. ISSN 00031224. Dostupné z: http://www.jstor.org/stable/2084520?origin=crossref

SUN, Yeran a Ming LI, 2015. Investigation of Travel and Activity Patterns Using Location-based Social Network Data: A Case Study of Active Mobile Social Media Users. *ISPRS International Journal of Geo-Information* [online]. **4**(3), 1512-1529 [cit. 2019-04-30]. DOI: 10.3390/ijgi4031512. ISSN 2220-9964. Dostupné z: http://www.mdpi.com/2220-9964/4/3/1512

THRIFT, N., 1977. An introduction to time-geography. 1. Norwich [Eng.*]: Geo Abstracts, University of East Anglia. ISBN 0902246674*.

Udržitelná mobilita, 2007. In: Observatoř bezpečnosti silničního provozu [online]. Brno: Observatoř bezpečnosti silničního provozu [cit. 2019-04-02]. Dostupné z: https://www.czrso.cz/clanek/udrzitelna-mobilita/?id=1038

URRY, John a Mimi SHELLER, 2006. The new mobilities paradigm. Environment and Planning A. 2006(38), 207–226. DOI: 10.1068/a37268.

VISWANATHAN, G., V. AFANASYEV, S. BULDYREV, E. *MURPHY, P. PRINCE a H. STANLEY, 1996. Lévy flight search* patterns of wandering albatrosses. Nature [online]. 381(6581), 413-415 [cit. 2019-04-20]. DOI: 10.1038/381413a0. ISSN 0028-0836. Dostupné z: http://www.nature.com/articles/381413a0

WICKHAM, Hadley, 2015. Dplyr. A Grammar of Data *Manipulation [online]. -: RStudio [cit. 2019-12-16]. Dostupné z: https:*//dplyr.tidyverse.*org/*

WICKHAM, Hadley, **2016**. Ggplot2: Elegant Graphics for Data Analysis. Create Elegant Data Visualisations Using the Grammar of Graphics [online]. New York: Springer-Verlag [cit. 2019-12-16]. Dostupné z: https://ggplot2.tidyverse.org/

YANG, Dingqi, Daqing ZHANG, Longbiao CHEN a Bingqing QU*, 2015. NationTelescope: Monitoring and visualizing large-scale* collective behavior in LBSNs. *Journal of Network and Computer Applications* [online]. **55**(55), 170-180 [cit. 2019-04-28]. DOI: 10.1016/j.jnca.2015.05.010. ISSN 10848045. Dostupné z: https://linkinghub.elsevier.com/retrieve/pii/S1084804515001101

YANG, Xin-She, 2010. Nature-inspired metaheuristic algorithms. 2nd ed. Frome,: Luniver Press. ISBN 978-1-905986-28-6.

YANG, Xin-She*, 2013. Chaos-Enhanced Firefly Algorithm with Automatic* Parameter Tuning. SHI, Yuhui, ed., Yuhui SHI. Recent Algorithms and Applications in Swarm Intelligence Research [online]. IGI Global, s. 125-136 [cit. 2019-04-26]. DOI: 10.4018/978-1-

4666-2479-5.ch007. ISBN 9781466624795. Dostupné z: http://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/978-1-4666-2479-5.ch007

YAN, Xiao-Yong, Xiao-Pu HAN, Tao ZHOU a Bing-Hong WANG, 2011. *Exact Solution of* the Gyration Radius of an Individual's Trajectory for a Simplified Human Regular Mobility Model. Chinese Physics Letters [online]. 28(12), 1–4 [cit. 2019-05-02]. DOI: 10.1088/0256-307X/28/12/120506. ISSN 0256-307X. Dostupné z: http://stacks.iop.org/0256-307X/28/i=12/a=120506?key=crossref.4620b64985fa71145505832311e3cc9b

ZIPF, George, 1946. The P 1 P 2 D Hypothesis: *On the Intercity Movement* of Persons. American Sociological Review [online]. 11(6) [cit. 2019-04-26]. DOI: 10.2307/2087063. ISSN 00031224. Dostupné z: http://www.jstor.org/stable/2087063?origin=crossref

8 Seznam obrázků

Obrázek 1: Kontinuum mobility. Zdroj: (Pooley, 2005) 13
Obrázek 2: Pohyb albatrosů, který popisuje Lévyho model. Zdroj: (Viswanathan, 1996) 27
Obrázek 3: Porovnání gravitačního a radiačního modelu. Zdroj: (Stefanouli, 2017).....29
Obrázek 4: Hrubá vizualizace tras *check-inů* uživatelů v Praze. Dobře rozpoznatelná je nejnavštěvovanější *venue* – Letiště. Zdroj: vlastní 33
Obrázek 5: Vizualizace *venues* v Praze. Vlevo *venues* jen s více *checkiny*. Zdroj: vlastní.....34
Obrázek 6: Vizualizace oblastí o ploše 0,25km^2, které měly alespoň 5 *venues*. Zdroj: vlastní38
Obrázek 7: Mapa zón v Praze bikesharingového projektu Rekola. Zdroj: (Rekola, 2019).....48
Obrázek 8: Ukázka 1000 náhodných jízd z datasetu Rekola. Zdroj: vlastní 49

9 Seznam grafů

Graf 1: Histogram délky tras z Foursquare datasetu. Zdroj: vlastní 35
Graf 3: Graf závislosti hustoty *venues* a průměrné délky trasy. V dolní části oranžově Praha. Zdroj: vlastní a (Noulas, 2012) 39
Graf 4: Graf závislostí plochy a průměrné délky trasy. V horní části oranžově Praha. Zdroj: vlastní a (Noulas, 2012) 39
Graf 2: Histogram ranku *venues* z Foursquare datasetu. Zdroj: vlastní 41
Graf 5: Graf vztahu Rank a funkce hustoty pravděpodobnosti dat z aplikace Foursquare. Zdroj: vlastní 42
Graf 6: Histogram vzdáleností jízd projektu Rekola. Zdroj: vlastní 50
Graf 7: Graf vztahu Rank a funkce hustoty pravděpodobnosti dat z projektu Rekola. Zdroj: vlastní 51

10 Seznam tabulek

Tabulka 1: Přehled výzkumů vzorců pohybu. Zdroj: vlastní.....25
Tabulka 2: Ukázka z datasetu Foursquare. Zdroj: vlastní a (Yang, 2015) 32
Tabulka 3: Tabulka s přehledem pohybu, míst, hustoty míst, plochy a průměrné délky trasy v porovnávaných městech. Zdroj: (Noulas, 2012).....36
Tabulka 4: Ukázka shlukové analýzy matice kategorií *venues* a uživatelů a poměry jejich označení v daných kategoriích. Zdroj: vlastní.....44
Tabulka 5: Ukázka datasetu Rekola. Zdroj: (Rekola, 2019).....49

11 Použité zkratky

API – *Application Programming Interface*, rozhraní pro programování a propojování s aplikací

PDF – *Probability Density Function,* funkce hustoty pravděpodobnosti