

Analýza návštěvnosti ČR s využitím GSM dat

**Metodika měření a analýzy turistické
návštěvnosti České republiky**

VYTVOŘENO PRO SPOLEČNOST

Česká centrála cestovního ruchu – CzechTourism

Štěpánská 567/15,

Praha 2 – Nové Město 120 00

Česká republika

DATUM A VERZE DOKUMENTU

28.02.2025

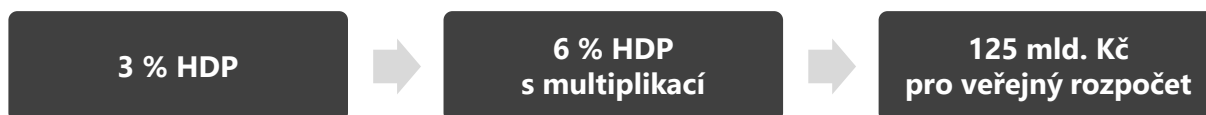
Verze 1.0

1 Obsah

1. Význam cestovního ruchu
2. Statistická data o cestovním ruchu
3. Využití geolokačních signalizačních dat pro analýzu cestovního ruchu
4. Zajištění bezpečnosti využití geolokačních signalizačních dat
5. Kalibrace geolokačních signalizačních dat
6. Metodika pro zpracování geolokačních signalizačních dat
7. Využití geolokačních signalizačních dat pro analýzu zahraničního cestovního ruchu
8. Přehled příkladů využití geolokačních signalizačních dat
9. Přínosy využití geolokačních signalizačních dat
10. Základní terminologie pro analýzy monitoringu návštěvnosti

2 Význam cestovního ruchu

Cestovní ruch je nedílnou součástí národních ekonomik i ekonomiky světové, je to průřezové a iniciační odvětví, jenž zahrnuje mimo jiné dopravu, ubytovací a stravovací zařízení, turistická zařízení, služby cestovních kancelářích a agentur, průvodcovské služby, informační systémy a informační centra aj. V České republice tvoří přibližně 3 % hrubého domácího produktu (HDP) a při zahrnutí multiplikace tvoří až 6 % HDP. Přínos turismu pro veřejný rozpočet dosahoval v roce 2024 dle předběžného odhadu společnosti Economic impactT úrovně cca 135 mld. Kč („Je odvětví cestovního ruchu budoucností Česka?“, únor 2025). United Nations World Tourism (UNWTO) definuje cestovní ruch jako činnost osob cestujících do míst a pobývajících v místech mimo své obvyklé prostředí po dobu kratší než jeden ucelený rok, za účelem trávení volného času a služebních cest (osoba nesmí být odměňována ze zdrojů navštíveného místa).



Turismus má pozitivní vliv na různá odvětví národní ekonomiky. Díky poptávce po službách cestovního ruchu dochází k multiplikačnímu efektu, který je zřejmý na tvorbě nových pracovních míst, snižování nezaměstnanosti, rozvoji regionů, zvyšování životní úrovně. Pro služby cestovního

ruchu je také budována potřebná infrastruktura. Turisté mají výdaje za potraviny, ubytovací a stravovací služby, služby průvodců, vstupné na atrakitivity a další, z čehož vyplývají příjmy, které mohou být v místě znovu reinvestovány. Cestovní ruch s sebou nese i negativní efekty jako například znečištění životního prostředí (uhlíková stopa), přetížení některých destinací v určité části sezóny nebo vznik turistických ghett (oblast se mění ve prospěch turistů, na úkor místních obyvatel) s čímž souvisí např. také turistická inflace, kdy začínají být ceny služeb a zboží pro místní obyvatele neakceptovatelné.

Pro identifikaci a využití turistického potenciálu destinace, pro měření a vyhodnocování pozitivních i negativních dopadů cestovního ruchu slouží statistiky v cestovním ruchu. Měřitelnost dopadů umožňuje vytvářet opatření, která pomáhají předcházet negativním důsledkům turismu. Naopak v oblastech, které nejsou zatím příliš turisticky využity je možné identifikovat a měřit jejich potenciál a podpořit tak regionální rozvoj udržitelným způsobem.

3 Statistická data o cestovním ruchu

Statistiky v cestovním ruchu fungují na základě doložitelných, reportů od subjektů se zpravodajskou povinností vůči Českému statistickému úřadu, dotazníkových šetření a dotazování v terénu. Zmiňované statistiky, především pak dotazování v terénu a dotazníková šetření, mohou vést ke zkreslení dat a také, vzhledem ke způsobu jejich sběru a zpracování, je jejich vyhodnocení poměrně časově náročné a mohou tak ztrácet aktuálnost.

V současné době roste význam dat ve všech odvětvích, a to jak pro potřeby různých analýz, tak i pro zlepšování chování firem vůči trhu a svým zákazníkům. Další z metod statistiky v cestovním ruchu je měření turistické návštěvnosti pomocí analýzy geolokačních signalizačních dat. Jedná se o novou moderní metodu užívající velký objem statistických dat od mobilních operátorů. Metoda je objektivní a zároveň spolehlivá, nezávislá na chování lidí a nezatížená zkresleními, která vznikají při terénních průzkumech a dotazníkových šetřeních. Příprava analýz turistické návštěvnosti je vzhledem k nutnosti získat a následně zpracovat geolokační signalizační data poměrně finančně a časově náročná, což je nutné zvážit při rozhodování o tom, kdy je vhodné tuto metodu aplikovat.

**Český statistický
úřad**

**Satelitní účet
cestovního ruchu**

**Statistiky MMR
a CzechTourismu**

4 Využití geolokačních signalizačních dat pro analýzu cestovního ruchu

Na základě geolokačních signalizačních dat od mobilních operátorů lze vytvářet analýzy návštěvnosti destinací, společenských, kulturních a sportovních akcí, analýzy mobility obyvatel a návazné analýzy rozvoje destinací či atraktivit pro destinační společnosti, veřejnou správu a soukromý sektor. Pro tyto analýzy jsou v současné době (únor 2025) využívána geolokační signalizační data od mobilních operátorů T-Mobile Czech Republic a O₂ Czech Republic v rámci projektu „Analýza návštěvnosti ČR s využitím GSM dat“.

Dle dat z Výroční zprávy Českého telekomunikačního úřadu (ČTÚ) za rok 2023 pokrývaly v polovině roku 2023 tyto dva největší čeští operátoři více než 70 % českého trhu mobilních služeb, kdy tržní podíl na základě celkového počtu aktivních mobilních SIM karet na maloobchodním trhu dosahoval úrovně 38,4 % v případě společnosti T-Mobile Czech Republic a 32,4 % v případě společnosti O₂ Czech Republic. Celkový počet aktivních SIM karet dle ČTÚ na konci roku 2023 přesáhl úroveň 15,5 mil. karet.

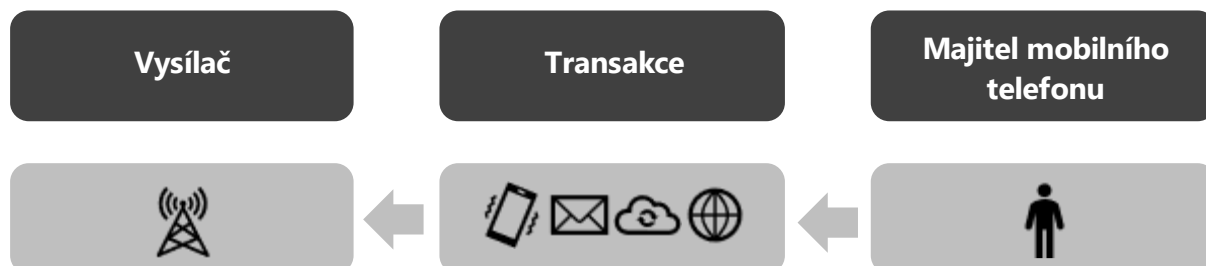
Podle údajů ČSÚ sesbíraných na jaře roku 2023 používá v ČR mobilní telefon 99 % osob starších 16 let, z toho 82 % osob používá chytrý telefon a 18 % klasický tlačítkový telefon.

Majitelé mobilních telefonů jsou vzhledem k jejich počtu a vysokému podílu na celkovém obyvatelstvu zdrojem užitečných dat, která se dají využít právě pro statistiku cestovního ruchu. Narozdíl od statistiky ČSÚ, která se zaměřuje primárně na údaje o počtu přenocování v hromadných ubytovacích zařízeních (HUZ), která data úřadu předávají, tato metoda dokáže zaznamenat i jednodenní návštěvníky, nebo návštěvníky, kteří nejsou ubytováni v jednom z některých HUZ, ale využijí např. sdílené ubytování, ubytování u známých nebo jiný typ pronájmu, u kterého není povinná evidence.

Geolokační signalizační data jsou souhrnem několika údajů o jednotlivých SIM kartách, jež jsou evidována mobilním operátorem na základě uskutečnění tzv. transakce mobilního telefonu k vysílači (hovor, odeslání SMS, připojení do datové sítě aj.). I když majitel neprovede transakci sám, mobilní telefon se nahlásí do sítě automaticky, a to vždy nejpozději každých 30 minut. Aktivita je tedy zaznamenána, i když majitel mobilní telefon přímo nepoužije. Obsahem balíku informací z každé telefonní transakce je:

- identifikační číslo SIM karty, které slouží pro přiřazení transakce jedné unikátní osobě,
- identifikační číslo přístroje
- čas události,
- identifikační číslo vysílače v níž transakce nastala,
- typ transakce.

Právě číslo vysílače je důležité pro určení lokace, a tak i pro umístění osoby v čase a prostoru. Vzhledem ke složitosti různých technologií mobilních sítí nelze přesně určit specifický algoritmus pravidelného ohlašování telefonu do mobilní sítě, jelikož má každý typ mobilní sítě svoje specifika.



5 Zajištění bezpečnosti využití geolokačních signalizačních dat

Dle zákona č. 151/2000 Sb. (telekomunikační tajemství) je na operátora uvalena povinnost data anonymizovat, tj. zbavit transakci osobních údajů konkrétního zákazníka, aby bylo zamezeno jeho dohledání. Operátor využívá před analýzou dat složité algoritmy anonymizace, které se pravidelně obměňují, aby bylo znemožněno jejich zneužití a spojení s konkrétním subjektem. Analýzy geolokačních signalizačních dat pro účely analýzy cestovního ruchu využívají pouze anonymizovaná vstupní data, která obsahují pseudoanonymizované číslo odpovídající právě 1 SIM kartě, údaje o umístění vysílače a čas transakce. Zpracované datové agregáty jsou dále podrobeny anonymizační metodě skrytí malých čísel, tak aby nebylo možné ztotožnit jedince ani v případě použití jiných datových základů. Koncový zákazník díky tomu získá statistické údaje o návštěvnosti, které může využívat pro své potřeby bez omezení za cenu zanedbatelného „umělého šumu“ v datech.

6 Kalibrace geolokačních signalizačních dat

Tato moderní metoda má ale svá úskalí. Například to, že jednotlivé vysílače mají vysílací přesah, jejich signál se navzájem překrývá, aby byla zajištěna určitá kvalita dostupnosti mobilní sítě, nebo existují jedinci, kteří mají v telefonu dvě SIM karty, naopak také ti, kdo nevlastní mobilní telefon (např. děti), dále je také nutné oddělit turisty a jednodenní návštěvníky od rezidentů.

Dále se může např. vyskytnout problém při přiřazování osob v jednotlivých územních celcích, kdy je připojení telefonu zaznamenáno právě v obou celcích. Pro tyto vymezující podmínky lze stanovit určitá kritéria, dle kterých lze data zpřesnit. Dále se v datech zohledňuje to, zda

se telefon pouze připojuje stylem „cikcak“ k jedné buňce a ke druhé buňce (tzv. „hop“) nebo zda se lidé skutečně přemísťují z místa na místo (mimo dosah prvního vysílače k druhému vysílači). Je tedy nezbytné, surová data od operátora přepočítávat a kalibrovat tak, aby byl výstup co nejbližší realitě.

Z dodaných dat nelze zjistit, jak přesně se lidé pohybují právě proto, jelikož neexistují kontinuální záznamy připojování se k určitým buňkám, jelikož může být zaznamenána transakce v místě A, o nějakou dobu později v místě B, ale v tomto intervalu, mezi těmito dvěma časy žádné údaje neexistují.

Zvolená metodika a způsob zpracování dat nicméně reagují na uvedené negativní jevy, díky čemuž je možné dodat na výstupu očekávané metriky.

7 Metodika pro zpracování geolokačních signalizačních dat

Veškerá data jsou zpracovávána v několika iteračních krocích, které vedou ke zpřesnění dat tak, aby měla vypovídající hodnotu, která co nejvíce odpovídá skutečnosti.

- 1** Očištění dat o zbytné události, technické a neaktivní SIM karty atd.
- 2** Převedení mapy buněk na územní jednotky.
- 3** Kalibrace dat – vynásobení naměřené hodnoty koeficientem, který je proměnlivý podle místa a času.
- 4** Zpracování dat – již vznikají validní data, segmentace osob na kategorie dle vztahu k místu.
- 5** Transformace dat na informace.

8 Využití geolokačních signalizačních dat pro analýzu zahraničního cestovního ruchu

U zahraničních uživatelů je jejich původ určován prostřednictvím tzv. mobile country code číselníku, který určuje obsluhujícího operátora a zemi původu SIM karty. Zahraniční uživatele tak lze rozdělit dle země původu SIM karty, se kterou se v rámci roamingu připojují do tuzemské mobilní sítě.

U zahraničních uživatelů, tedy zákazníků, kteří využívají mobilní síť v roamingu, je extrapolace přepočtu dat na celkovou populaci, která navštívila sledované území, založena na počtu připojení mobilních zařízení uživatelů zahraničních SIM k obsluhujícímu vysílači ve sledovaném území, informacích o tržním zastoupení mobilních operátorů daných SIM karet v zemi jejich původu, a koeficientu, vycházející z mezinárodních dohod mezi tuzemskými a zahraničními operátory, o využívání kapacity tuzemské mobilní sítě pro účely roamingu. Vzhledem k tomu, že nelze sledovat trajektorii pohybu zahraničních SIM, využívajících roamingové služby v tuzemské mobilní síti, jsou veškeré statistiky počítány pomocí matematického modelu, který odhaduje chování zahraničních uživatelů na základě množství zaznamenaných připojení, tzv. eventů, mobilních telefonů se zahraničními SIM kartami k vysílačům, obsluhující sledovanou oblast.

Před samotným výpočtem statistik proběhne agregace všech zaznamenaných eventů z vysílačů v analyzované oblasti (např. kraj ČR) a následně je pomocí matematického modelu odhadován počet uživatelů a jejich časoprostorové chování, kteří návštěvu sledované oblasti vykonali. Hlavními vstupy pro tzv. trénování matematického modelu jsou informace o počtu eventů českých uživatelů, kteří zároveň udělili souhlas se zpracováním osobních dat a u kterých je tak známá trajektorie pohybu SIM v mobilní síti. Na základě toho je model schopen odhadnout z počtu eventů zahraničních uživatelů i jejich časoprostorové chování, které pak vstupuje do procesu zařazení SIM karty do uživatelských skupin.

Finální extrapolace na celkovou populaci zahraničních návštěvníků ve sledovaném území je realizována algoritmem, zohledňujícím kombinaci tržního podílu zahraničních operátorů, jejichž SIM karty byly připojeny na vysílače, obsluhující sledované území a koeficientu, vycházejícího z využívání kapacity tuzemské mobilní sítě zahraničními operátory pro účely roamingu. Pro co největší spolehlivost výsledků, jsou data také dále kalibrována pomocí dalších datových zdrojů, jako např. data o tržním zastoupení jednotlivých roamingových partnerů, jejichž zákazníci přicestovali do sledovaných území v ČR, dále koeficienty, vycházející z mezinárodních dohod tuzemských a zahraničních operátorů o využívání kapacity tuzemské mobilní sítě pro účely roamingu, data o počtu ubytovaných v HUZ, experimentální datové vzorky z online ubytovacích platforem či na základě typově podobných analýz.

Uvedená metoda byla navržena s cílem získání co možná nejvíce přesných dat o zahraničním cestovním ruchu při současném respektování limitů souvisejících se zachováním legislativně ukotvenou ochranou osobních údajů v oblasti roamingu.

9 Přehled příkladů využití geolokačních signalizačních dat

1) Statistiky cestovního ruchu

Jedná se o unikátní statistiky o domácím i zahraničním cestovním ruchu. Analýza probíhá na základě geolokačních signalizačních dat od mobilních operátorů, kdy jsou evidována data za zvolené období a místo. Analytici data zpracovávají na základě interních algoritmů a kalibrace. Výsledný report by měl také obsahovat informace od konkrétních destinačních společností, měst, obcí, kulturních a sportovních zařízení, jejich webových stránek, strategických dokumentů a dalších, za účelem vytvoření komplexního výstup na míru zadavateli. Následně je možné report na základě analytických zjištění doplnit i o návrhovou část či část se strategickým plánem pro danou destinaci.

2) Analýza zdrojů cest

Jedná se o unikátní data mobility obyvatel, kdy se vytváří matice zdrojů cest pro jakoukoli zvolenou oblast. Vzhledem k faktu, že matice obsahují data o pohybu všech návštěvníků v území, slouží jako vhodný podklad pro oblast státní či územní správy.

3) Retail Geo-Intelligence

Signalizační data poskytují jedinečný pohled na rozložení návštěvy zákazníků. Odpovídají na otázky, díky kterým lze pochopit jejich časoprostorové rozložení a chování. Výstupem mohou být podklady pro analytiku, tištěné i interaktivní reporty s pokročilými vizualizacemi, grafy či hloubkové analýzy poskytující unikátní pohled na zákazníky či společnost.

4) Site-Rank

Služba Site-Rank slouží k výběru optimálních míst pomocí analýzy vztahů mezi geodemografií a potenciálem lokalit. Výstupem mohou být data pro informační systémy, reporty či analýzy.

10 Přínosy využití geolokačních signalizačních dat

Metoda měření návštěvnosti za pomoci geolokačních signalizačních dat skýtá mnoho výhod oproti klasickým metodám používaným v cestovním ruchu, mezi něž patří například rychlost zpracování výstupu, objektivní a reálný podklad pro statistiku nebo také přidaná hodnota analýzy daná konzultanty, kteří jsou odborníky ve svém oboru.

Pro oblast cestovního ruchu, rozvoje území, objektů v území má analýza návštěvnosti mnoho přínosů:

- může sloužit jako podklad pro efektivní řízení (kompetence destinačních společností, spolupráce mezi nimi a zřizovateli),

- spolehlivý podklad k nastavení kvalitních strategických cílů a návazné tvorbě dlouhodobých strategických dokumentů,
- umožňuje definici počtu návštěvníků a jejich rozprostření po regionu / oblasti,
- poskytuje informace o oblíbených, rentabilních a hojně navštěvovaných akcích v místě / ve městě / regionu / oblasti, které vedou k získání podpůrných argumentů pro jejich realizaci či rozšíření v budoucnu,
- informace o rozložení počtu návštěvníků v průběhu roku, měsíce, dnů v týdnu může poskytnout východisko pro boj se sezónností v podobě regulace návštěvnosti v období sezóny pomocí různých nástrojů a činností,
- analýzu je možné provádět kontinuálně anebo opakovaně s identicky nastavenými procesy zpracování dat, díky čemuž je možné výstupy mezi sebou porovnávat,
- díky masivnímu souboru vstupních dat má analýza vysokou vypovídající úroveň,
- analýza ulehčuje tvorbu komunikační a marketingové strategie na základě informací o zdrojových trzích, délce pobytu, sezónnosti a obvyklém chování návštěvníků ve vztahu k jejich pobytu v destinaci, tyto strategie ulehčují tvorbu vhodných a přesně zacílených produktů cestovního ruchu dle profilu návštěvníků destinace,
- prioritizace opatření na podporu cestovního ruchu na základě reálných podkladových dat o destinaci.

Pro oblast státní a územní správy jsou výstupy analýz mobility obyvatel využitelné v následující oblastech:

- možnost zefektivnění chodu správy území díky definování denního pohybu osob:
 - počet obyvatel v průběhu dne a v noci,
 - dojíždka a vyjíždka z okolí (na úrovni krajů, správních obvodů obcí s rozšířenou působností, správních obvodů obcí s pověřeným obecním úřadem, obcí apod.),
 - krátkodobé cesty (služební cesty, nákupy, volný čas...),
- podklady slouží pro tvorbu kvalitních strategických dokumentů,
- mohou vést k optimalizaci umístění úřadů veřejné správy s ohledem na pohyb obyvatel,
- úpravy hranic správních obvodů,
- napomáhají efektivnímu nastavení toků dopravy (železnice vs. autobusové spoje, optimalizace městské hromadné dopravy) a s tím spojené obslužné vybavenosti jednotlivých obcí,
- poskytnutí zpětné vazby o reálné návštěvnosti podpořených akcí na území obce / kraje / státu a na to navazující adekvátnost finanční podpory,

- taktéž je velmi důležitým benefitem pomoc při plánování strategického rozvoje území na základě dat o pohybu a chování obyvatel.

11 Základní terminologie pro analýzy monitoringu návštěvnosti

Pro analýzy monitoringu návštěvnosti destinace je využívána následující terminologie, která rozděluje účastníky cestovního ruchu do následujících skupin podle jejich charakteristických rysů chování, resp. podle počtu strávených hodin v destinaci. Základní rozdělení je na vícedenního návštěvníka neboli turistu a jednodenního návštěvníka, resp. výletníka. Tabulka znázorňuje referenčního turistu, který přijel do destinace dvakrát a výletníka, který přijel do destinace třikrát. Červené body znázorňují zachycení účastníka v konkrétních dnech.

	Den 1	Den 2	Den 3	Den 4	Den 5	Den 6	Den 7	CELKEM
TURISTA	●							1
TURISTONÁVŠTĚVA		●				●		2
PŘENOCOVÁNÍ		●	●	●		●	●	4
24 h POBYT		●	●	●		●	●	2
TURISTODEN		●	●	●		●	●	5
VÝLETNÍK	●							1
VÝLETODEN		●				●	●	3

Vícedenní návštěvník

- TURISTA = vícedenní návštěvník.** Turista přijel do destinace 2x. Při první návštěvě strávil v destinaci 3 noci, při druhé 1 noc.
- TURISTONÁVŠTĚVA = počet opakovaných návštěv.** Turista přijel do destinace 2x. Při první návštěvě strávil v destinaci 3 noci, při druhé návštěvě 1 noc.
- PŘENOCOVÁNÍ = počet přenocování v destinaci.** Turista přijel do destinace 2x. Při první návštěvě přenocoval v destinaci 3x, při druhé návštěvě 1x. Dohromady tedy v destinaci přenocoval 4x.
- 24 h POBYT = počet celých dnů strávených v destinaci.** Turista přijel do destinace 2x. Při první návštěvě strávil v destinaci 2 celé dny, při druhé návštěvě v destinaci celý den nestrávil.
- TURISTODEN = počet dnů strávených v destinaci.** Turista přijel do destinace 2x. Při první návštěvě strávil v destinaci 3 turistodny, při druhé návštěvě 2 turistodny.

Jednodenní návštěvník

- **VÝLETNÍK = jednodenní návštěvník.** Výletník přijel do destinace 3x, žádnou noc zde nestrávil.
- **VÝLETODEN = počet jednodenních návštěv do destinace.** Výletník do destinace vykonal 3 jednodenní návštěvy.

Uvedené metriky jsou vztaženy relativně k domovu návštěvníka, který je definován jako obec, kde jedinec stráví většinu času v noci (1:00 – 6:00) za poslední 3 měsíce. Pokud se jedinec nachází v DMO, respektive kraji, kde má domovskou obec, není zde započítáván jako návštěvník.

Ze zpracování jsou dále vyloučeny pravidelné cesty (typicky za prací a do škol), které by ovlivňovaly metriky cestovního ruchu ve smyslu jejich nadhodnocování.