https://doi.org/10.22034/aa.2025.732975
اندیشه آماری، سال ۱۴۰۳، دوره ۲۹، شماره ۲، ص ۲۶ – ۱

# آشنایی با داده‌های موقعیت‌یابی تلفن همراه در تولید آمار رسمی

**یاسین محمدآقایی[۱] و زهرا رضائی قهرودی[۲]***

[۱] دانشکده ریاضی، آمار و علوم کامپیوتر، دانشگاه تهران

[۲] دانشکده ریاضی، آمار و علوم کامپیوتر، دانشگاه تهران

تاریخ دریافت: ۱۴۰۴/۰۷/۰۴
تاریخ پذیرش: ۱۴۰۴/۱۰/۰۱.

## چکیده

با گسترش فناوری‌های ارتباطی و افزایش ضریب نفوذ تلفن همراه، حجم زیادی از داده‌های موقعیت‌یابی در جوامع تولید می‌شود. این داده‌ها، در صورت بهره‌برداری صحیح و قانونی، می‌توانند به‌عنون منبعی ارزشمند برای تولید آمار رسمی در حوزه‌هایی مانند مهاجرت، گردشگری، جابجایی در بحران‌ها و بلایا، حمل و نقل و ترافیک و مدیریت بحران تبدیل شوند. مقاله حاضر با هدف بررسی ظرفیت‌ها و محدودیت‌های به‌کارگیری داده‌های موقعیت‌یابی تلفن همراه و تبیین روش‌شناسی تولید آمار مبتنی بر این داده‌ها در نظام آماری کشور تدوین شده است.

با توجه به نبود منابع فارسی در این حوزه، این پژوهش تلاش می‌کند تا با مرور جامع ادبیات بین‌المللی و تحلیل یک مجموعه‌داده واقعی، چارچوبی کاربردی برای استفاده از داده‌های تلفن همراه در تولید آمار رسمی ارائه دهد. در این چارچوب، ضمن معرفی انواع داده‌های تلفن همراه و مرور کاربردهای آن‌ها در تولید آمار، الزامات قانونی، اخلاقی و فنی بهره‌گیری از این داده‌ها، مورد بررسی قرار گرفته است. همچنین، روش‌شناسی استخراج و پردازش داده‌های موقعیت‌یابی تلفن همراه تشریح شده و مطالعه‌ای موردی بر پایه مجموعه‌داده مرتبط با الگوهای تحرک شهری انجام شده است. دستاورد اصلی پژوهش، مستندسازی تجربیات و انتقال دانش فنی برای بومی‌سازی و بهره‌برداری از این نوع داده‌ها در نظام آماری کشور است. تحلیل داده‌های موردی نشان داد که بحران‌ها تأثیر قابل‌توجهی بر رفتار حرکتی کاربران داشته‌اند؛ به‌گونه‌ای که در شرایط عادی، بخش کوچکی از کاربران، دفعات زیادی مشاهده شده‌اند و تحرک بالایی (۵۰۰ مکان منحصربه‌فرد) از خود نشان داده‌اند، در حالی که این تحرک در وضعیت بحران به‌طور چشمگیری کاهش یافته و افراد بسیار کمتری با هر نوع سطح تحرکی (پایین، متوسط و بالا) مشاهده شده‌اند و ترددها عمدتاً در مناطق خاصی متمرکز شده است.

**واژه‌های کلیدی:** داده‌های موقعیت‌یابی تلفن همراه، آمار رسمی، مه‌داده‌ها، حریم خصوصی، چارچوب قانونی، تحلیل تحرک جمعیت.

## ۱ مقدمه

در عصر ارتباطات دیجیتال، شرکت‌های مخابراتی تلفن همراه یکی از منابع کلیدی در تولید و نگهداری داده‌های موقعیت‌یابی به‌شمار می‌آیند. این داده‌ها به دلیل گستردگی پوشش شبکه، فراوانی کاربران، و ثبت مستمر اطلاعات، از ارزش آماری بالایی برخوردارند. با توسعه زیرساخت‌های دیجیتال و فناوری‌های هوشمند، داده‌های موقعیت‌یابی جایگاه مهم‌تری در تحلیل‌های آماری، سیاست‌گذاری و طراحی خدمات

*نویسنده مسئول مقاله، z.rezaeigh@ut.ac.ir

آینده خواهند یافت. سازمان ملل متحد و برخی نهادهای آماری بین‌المللی، استفاده از این نوع داده‌ها را به‌عنوان یکی از مسیرهای نوین در تحول نظام آماری کشورهای در حال توسعه معرفی کرده‌اند. از سوی دیگر، دسترسی به این اطلاعات نیازمند وجود چارچوب‌های قانونی، توافق‌نامه‌های بین‌نهادی و پروتکل‌های دقیق برای حفظ حریم خصوصی و امنیت داده‌ها است. در این میان، داده‌های تلفن همراه یا داده‌های موقعیت‌یابی تلفن همراه (MPD) [۱]، به‌عنوان یکی از منابع نوظهور مه‌داده‌ها، ظرفیت بالایی برای بهبود فرآیندهای آماری به‌ویژه در کشورهایی که به‌دنبال ارتقای کیفیت و چابکی نظام آماری خود هستند، دارند. این داده‌ها، به‌عنوان یکی از نوآوری‌های عصر دیجیتال، تحولی چشم‌گیر در شیوه تولید و تحلیل آمارهای رسمی ایجاد کرده‌اند و با فراهم آوردن اطلاعات دقیق، بلادرنگ و در مقیاس گسترده، امکان رصد پویایی‌های جمعیتی، اقتصادی و اجتماعی را به شکلی بی‌سابقه فراهم می‌کنند. با همکاری بین **دفاتر آماری** [۲] (SO) و **اپراتورهای شبکه تلفن همراه** [۳] (MNO) امکان محاسبه آمار رسمی وجود دارد؛ به گونه‌ای که داده‌های MNO مکمل سایر منابع داده‌ای در دسترس دفاتر آماری (مانند داده‌های آمارگیری‌ها و ثبت‌های اداری) هستند و ترکیب آن‌ها می‌تواند به تولید نسل جدیدی از محصولات آماری با دقت مکانی و زمانی بالاتر و به‌روزرسانی سریع‌تر نسبت به آمار سنتی منجر شود [۱۶].

هم‌زمان با گسترش پوشش شبکه‌های مخابراتی و قابلیت‌های فنی گوشی‌های هوشمند، امکان گردآوری و تحلیل داده‌های مکانی مربوط به افراد و کسب اطلاعات درباره محل حضور یا حرکت کاربران در طول زمان فراهم شده است. به عنوان مثال، وقتی کاربری در یک محدوده جغرافیایی خاص از خدمات تلفن همراه استفاده می‌کند، داده‌هایی درباره موقعیت او، مدت‌زمان حضور در آن موقعیت و نحوه جابه‌جایی‌اش بین موقعیت‌های مختلف در سامانه‌های مخابراتی ثبت می‌شود که می‌توان از آن‌ها برای تحلیل الگوهای جمعیتی، رفتارهای انسانی، و تغییرات زمانی–مکانی استفاده کرد. این اطلاعات به‌طور مستقیم از طریق آنتن‌های مخابراتی، سامانه موقعیت‌یابی ماهواره‌ای، یا اتصال به شبکه‌های بی‌سیم شهری قابل استخراج هستند [۲۰].

از آنجا که داده‌های موقعیت‌یابی به‌صورت مداوم و بدون وقفه تولید می‌شوند و این ویژگی امکان تحلیل‌های به‌هنگام را فراهم می‌سازد؛ در نتیجه، نهادهای آماری می‌توانند به جای تکیه صرف بر داده‌های سنتی که زمان‌بر و پرهزینه هستند، از منابع پویاتری بهره بگیرند[۷]. این داده‌ها همچنین امکان ردیابی الگوهای رفت‌وآمد، شناسایی مناطق پرتردد، زمان‌های اوج ترافیک، و الگوهای مهاجرت داخلی را فراهم می‌کنند. در نتیجه، برای برنامه‌ریزی حمل‌ونقل، بهینه‌سازی خدمات عمومی، و مدیریت منابع شهری ابزار ارزشمندی محسوب می‌شوند. در شرایط بحران (مانند زلزله، سیل، یا همه‌گیری‌ها)، داده‌های موقعیت‌یابی می‌توانند اطلاعاتی درباره نحوه جابجایی جمعیت، مناطق متراکم، و نیازهای فوری ارائه دهند. استفاده از این داده‌ها به سازمان‌های امدادی و تصمیم‌گیران کمک می‌کند تا پاسخ‌های سریع‌تر و مؤثرتری داشته باشند.

در سال‌های اخیر، برخی کشورها به‌صورت رسمی از داده‌های موقعیت‌یابی تلفن همراه برای برآورد جمعیت در مناطق مختلف، تحلیل جابه‌جایی‌های فصلی و پایش اثرات حوادث غیرمترقبه بهره گرفته‌اند. به‌عنوان نمونه، در دوران همه‌گیری بیماری کووید ۱۹، برخی مراکز آماری با استفاده از داده‌های تلفن همراه، میزان رعایت محدودیت‌های ترددی را مورد سنجش قرار دادند[۴]. در اتحادیه اروپا، داده‌های تردد کاربران تلفن همراه به‌عنوان مبنایی برای بررسی الگوهای جابه‌جایی پناه‌جویان در زمان بحران‌های انسانی مورد استفاده قرار گرفت. این داده‌ها به سیاست‌گذاران در تدوین برنامه‌های حمایت اجتماعی کمک شایانی نمودند [۲]. تجربه کشورهای پیشرویی همچون استونی در استفاده از داده‌های ناشناس‌شده برای برنامه‌ریزی شهری و برآورد سفرهای ورودی و خروجی، اندونزی در برآورد تعداد بازدیدکنندگان بین‌المللی، فنلاند در مدیریت بحران با تحلیل تحرک جمعیت، کره جنوبی در بهینه‌سازی سیستم‌های ترافیکی، و ایتالیا در ردیابی الگوهای مهاجرت فصلی، گواهی بر این ظرفیت بی‌بدیل است.

در کشورهای استونی و اندونزی، ارائه‌دهندگان آمار، از داده‌های موقعیت‌یابی تلفن همراه به عنوان بخشی از تولید منظم آمار رسمی استفاده می‌کنند. از سال ۲۰۰۸، بانک مرکزی استونی [۵] با شرکت‌های خصوصی برای استفاده از مه‌داده‌ها در تولید آمار رسمی همکاری کرده است. این بانک از داده‌های تلفن همراه اپراتورهای شبکهٔ تلفن همراه استونی برای برآورد سفرهای ورودی و خروجی و از داده‌های پرداخت با کارت اعتباری برای تعدیل ارقام مربوط به هزینه‌ها استفاده می‌کند.

[1] Mobile Positioning Data
[2] Statistical Offices
[3] Mobile Network Operator
[۴] سازمان همکاری و توسعه اقتصادی، «نقش داده‌های تلفن همراه در پاسخ به بحران کووید–۱۹»، ۲۰۲۱.
[5] Eest Pank

ترکیب این دو مجموعه داده، همراه با چندین مجموعه داده مرجع دیگر، امکان تولید به‌موقع و کارآمد آمار تجارت خدمات را فراهم می‌آورد که در تدوین تراز پرداخت‌ها و حساب‌های ملی به کار می‌رود. آمار سفر مبتنی بر داده‌های تلفن همراه از سال ۲۰۱۲ به طور جداگانه و به‌عنوان آمار رسمی در وب‌سایت بانک مرکزی استونی به صورت فصلی منتشر شده است و سری زمانی آن به سال ۲۰۰۸ برمی‌گردد. این همکاری در گذر زمان به عنوان بخشی از فرآیند کسب‌وکار آماری در بانک مرکزی پایدار مانده است [۲۰]. برای کسب اطلاعات بیشتر در خصوص روش‌شناسی تدوین آمار سفرهای بین‌المللی در استونی به سایت بانک مرکزی استونی [۶] مراجعه شود.

برای افزایش پوشش داده‌های مربوط به تعداد ورود بازدیدکنندگان خارجی، به‌ویژه در مناطق مرزی که آمار آن‌ها در ایست‌های رسمی مهاجرت ثبت نمی‌شود، اداره آمار اندونزی (BPS) [۷]، و وزارت گردشگری از اکتبر ۲۰۱۶ در تلاش بوده‌اند روش‌شناسی محاسبه تعداد بازدیدکنندگان خارجی را با استفاده از داده‌های موقعیت‌یابی تلفن همراه بهبود دهند. داده‌های MPD در گذرگاه‌های مرزی شهرستان‌هایی به کار گرفته می‌شود که فاقد ایستگاه‌های رسمی مهاجرت هستند و اجرای پیمایش‌های پستی مرزی به دلیل شرایط جغرافیایی دشوار است.

مجموعه داده‌های سری زمانی مربوط به آمار ورود بازدیدکنندگان خارجی که شامل داده‌های موقعیت‌یابی تلفن همراه است، از اکتبر ۲۰۱۶ توسط اداره آمار اندونزی منتشر شده است. برای بررسی آمارهای گردشگری اداره آمار اندونزی و روش‌شناسی استفاده از داده‌های تلفن همراه به آدرس سایت اداره آمار اندونزی [۸] مراجعه کنید. آخرین آمار سالانه منتشر شده از ورود بازدیدکنندگان بین‌المللی توسط اداره آمار اندونزی در سال ۲۰۲۴ بوده است. منابع داده این انتشار از اداره کل مهاجرت وزارت مهاجرت و اصلاحات [۹]، وزارت گردشگری و همچنین داده‌های حاصل از موقعیت‌یابی تلفن همراه به دست آمده است. این انتشار شامل آمار مربوط به تعداد ورود بازدیدکنندگان بین‌المللی بر اساس بندر ورودی، ملیت و الگوی فصلی می‌باشد.

در هلند نیز، **اداره آمار هلند**[۱۰] از داده‌های تلفن همراه برای برآورد جمعیت روزانه و تحلیل تراکم جمعیت در ساعات مختلف روز استفاده کرده است، که در برنامه‌ریزی شهری و تخصیص منابع کاربرد داشته است [۲۰]. استین‌بروخن و همکاران [۱۷] مرور کلی بر مهم‌ترین مطالعات و پروژه‌هایی در کشور هلند داشتند که به استفاده از داده‌های استخراج‌شده از شبکه‌های تلفن همراه برای به‌دست‌آوردن اطلاعات مکانی و برآورد ترافیک افراد می‌پردازند.

گسترش کاربرد داده‌های موقعیت‌یابی تلفن همراه در نظام‌های آماری، نیازمند وجود چارچوب‌های معتبر جهانی و همکاری نهادهای بین‌المللی است. این داده‌ها به علت ماهیت پیچیده، وابستگی به فناوری‌های مخابراتی، و مسائل حقوقی و اخلاقی، نیازمند استانداردسازی، هدایت علمی و نظارت مستمر هستند. در این زمینه، نهادهای بین‌المللی از جمله بخش آمار سازمان ملل (UNSD) [۱۱]، اتحادیه بین‌المللی مخابرات، اداره آمار اتحادیهٔ اروپا، بانک جهانی و سازمان همکاری اقتصادی و توسعه، هر یک با مأموریت خاص خود، تلاش کرده‌اند تا استفاده از این داده‌ها را تسهیل و ساماندهی کنند. برای نمونه، اتحادیه بین‌المللی مخابرات با انتشار دستورالعمل‌های فنی–اخلاقی، مرکز آمار اتحادیه اروپا با تدوین پروتکل‌های یکپارچه‌سازی داده‌ها، و گروه تخصصی مه‌داده‌های سازمان ملل با ارائه چارچوب‌های حکمرانی داده، نقش بی‌بدیلی در تسهیل همکاری بین کشورها و کاهش چالش‌هایی مانند نقض حریم خصوصی یا ناهماهنگی حقوقی ایفا کرده‌اند. کمیته کارشناسان سازمان ملل در زمینه مه‌داده‌ها و علم داده‌ها برای آمارهای رسمی $(UN-CEBD)$ [۱۲] دارای یک گروه کاری کمیته کارشناسان سازمان ملل در زمینه داده‌های تلفن همراه هستند که با هدایت اتحادیه بین‌المللی مخابرات (ITU) به بررسی استفاده از مه‌داده‌های تلفن همراه در حوزه‌های مختلف آمار و توسعه روش‌شناسی‌ها می‌پردازند. این گروه کاری متشکل از حدود ۵۰ عضو حقیقی / ۳۰ نهاد شامل آژانس‌های بین‌المللی و منطقه‌ای، کشورها، دانشگاه‌ها و مؤسسات/شرکت‌های خصوصی است که ماهی یک بار (به صورت مجازی) تشکیل جلسه می‌دهد تا درباره خروجی‌ها، رویدادها و سایر موضوعات مرتبط بحث کند. زیرگروه‌های کاری این تیم در حوزه‌های زیر فعالیت می‌کنند.

۱. آمار گردشگری (با هدایت و رهبری اداره آمار اندونزی، $BPS$)

۲. آمار مهاجرت (با هدایت و رهبری اداره ملی آمار گرجستان، $GeoStat$)

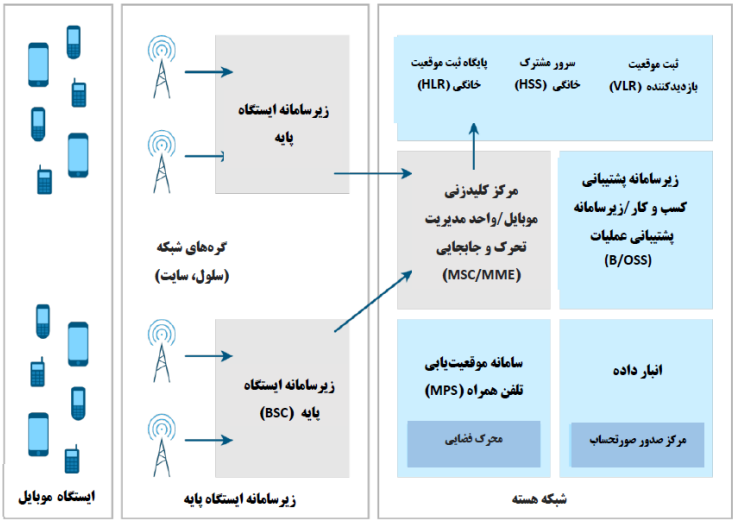
[6] https://statistika.eestipank.ee/failid/mbo/valisreisid_eng.html
[7] Statistics Indonesia (Indonesian: Badan Pusat Statistik, BPS)
[8] https://www.bps.go.id/en/statistics-table?subject=561
[9] Ministry of Immigration and Corrections
[10] CBS – Statistics Netherlands (Centraal Bureau voor de Statistiek)
[11] United Nations Statistics Division
[12] UN Committee of Experts on Big Data and Data Science for Official Statistics

۳. سرشماری و جمعیت پویا (با هدایت و رهبری شرکت خصوصی Positium در کشور استونی)
۴. جابجایی در بلایای طبیعی (با هدایت و رهبری دانشگاه توکیو)
۵. شاخص‌های جامعه اطلاعاتی [۱۳] (با هدایت و رهبری $ITU$)
۶. آمارهای حومه‌سفر و حمل و نقل [۱۴]
۷. آموزش و ظرفیت‌سازی

در این بخش، به اهمیت داده‌های موقعیت‌یابی تلفن همراه در آمارهای رسمی و مطالعات تطبیقی و تجربه کشورهای پیشرو در استفاده از داده‌های موقعیت‌یابی تلفن همراه پرداخته شده است. با توجه به نبود منابع فارسی در این حوزه، این پژوهش تلاش می‌کند تا با مرور جامع ادبیات بین‌المللی و تحلیل یک مجموعه‌داده واقعی، چارچوبی کاربردی برای استفاده از داده‌های تلفن همراه در تولید آمار رسمی ارائه دهد. در بخش‌های دوم و سوم، به‌ترتیب به منابع تولید داده‌های تلفن همراه و چارچوب‌های قانونی مرتبط با به‌کارگیری داده‌های موقعیت‌یابی در آمار رسمی پرداخته می‌شود. بخش چهارم، روش‌شناسی تولید آمار مبتنی بر داده‌های موقعیت‌یابی تلفن همراه را تبیین می‌کند و بخش پنجم به تحلیل مجموعه‌داده‌ی تحرک شهری $YJMob100K$ ژاپن اختصاص دارد. در پایان نیز جمع‌بندی و نتیجه‌گیری ارائه می‌شود.

## ۲ منابع تولید داده‌های تلفن همراه

منابع گردآوری داده‌های تلفن همراه از دو روش سامانه‌های اپراتورهای تلفن همراه به‌صورت خودکار و یا از طریق داده‌های حاصل از رویدادهای تلفن همراه و مستقیما از دستگاه کاربر که به داده‌های موقعیت‌یابی غیرفعال معروف هستند، گردآوری می‌شوند. در ادامه این دو نوع داده به همراه ویژگی‌های آن‌ها معرفی می‌شوند.




شکل ۱: منابع داده‌های موقعیت‌یابی تلفن همراه در سامانه‌های اپراتورهای تلفن همراه [۲۰]

### ۱.۲ داده‌های حاصل از سامانه‌های اپراتورهای تلفن همراه

برای بهره‌برداری مؤثر از داده‌های موقعیت‌یابی تلفن همراه در تولید آمارهای رسمی، شناخت دقیق از نحوه تولید، ذخیره‌سازی و استخراج این داده‌ها در زیرساخت‌های فنی اپراتورهای تلفن همراه ضروری است. داده‌های تلفن همراه، پیش از آنکه در قالب‌های تحلیلی مانند داده‌های داخلی یا رومینگ طبقه‌بندی شوند، در بستر شبکه‌های مخابراتی و سامانه‌های عملیاتی اپراتورها شکل می‌گیرند. این سامانه‌ها شامل اجزای متعددی هستند که هر یک نقش خاصی در ثبت، پردازش و انتقال داده‌های مکانی و زمانی ایفا می‌کنند.

استخراج داده‌های موقعیت‌یابی از سامانه‌های اپراتور تلفن همراه به راهکارهای فنی خاص هر اپراتور بستگی دارد. به‌طور معمول، هر

[13] Information society indicators
[14] Transport and commuting statistics

اپراتور دارای مرکز صدور صورتحساب [15] و انبار داده‌ای [16] است که اطلاعات به‌صورت دوره‌ای برای اهدافی چون صدور صورتحساب، برنامه‌ریزی و مدیریت شبکه، ذخیره می‌شوند (شکل ۱). علاوه بر سامانه‌های ذخیره‌سازی مرکزی که داده‌ها در آن‌ها به‌راحتی در دسترس هستند، امکان استخراج داده‌ها در فرایند فراکاوی و پایش [17] در سطح زیرسامانه‌ی ایستگاه پایه (BSS) [18] و زیرسامانه‌ی کلیدزنی شبکه (NSS) [19] نیز وجود دارد. در ادامه، در خصوص هر یک از اجزا، توضیحات بیشتری داده می‌شود.

### ۱.۱.۲ سامانه‌های ذخیره‌سازی مرکزی

استخراج داده از سامانه‌های ذخیره‌سازی مرکزی[20] ، آسان‌ترین روش برای به‌دست آوردن داده برای آمارهای ملی است. چندین منبع وجود دارد که می‌توان چنین داده‌هایی را که برای اهداف مختلف گردآوری شده‌اند، در آن‌ها یافت که در ادامه معرفی می‌شوند.

۱. **مرکز صدور صورتحساب.** این بخش مسئول ثبت و پردازش ثبت جزئیات تماس (CDR) [21] و ثبت جزئیات پروتکل اینترنت (IPDR) [22]، است و داده‌ها پس از تکمیل فرآیند صدور صورتحساب، در آن ذخیره می‌شوند. ثبت جزئیات تماس، شامل فراداده‌های مربوط به تراکنش‌های مخابراتی مانند تماس، پیامک است که معمولاً برای امور صورتحساب ثبت می‌شوند، در حالی که IPDR یا (DDR) [23]، فرادادهای مربوط به تبادل داده در شبکه تلفن را ثبت می‌کند.

۲. **پایگاه داده مشتریان** [24]، حاوی اطلاعاتی درباره کاربران مانند اطلاعات جمعیت‌شناختی و محل سکونت است که می‌تواند ارزش بیشتری به داده‌های CDR یا IPDR بدهد.

۳. **انبار داده**، پایگاهی برای تجمیع اطلاعات از منابع مختلف شبکه است. اما به دلیل حجم بالای داده‌های ذخیره‌شده، همیشه دسترسی به آن‌ها آسان نیست.

### ۲.۱.۲ فراکاوی و داده‌های سیگنال‌دهی

پایش و فراکاوی شبکه تلفن همراه یکی از روش‌های کلیدی برای استخراج داده‌های موقعیت‌یابی در سطح عملیاتی است. این فراکاوی به‌طورکلی به دو روش فعال و غیرفعال برای نظارت و استخراج داده از شبکه‌های مخابراتی انجام می‌شود. در روش **فراکاوی فعال**[25]، اپراتور به‌طور مستقیم ترافیک مصنوعی تولید می‌کند تا عملکرد شبکه را ارزیابی کند. اگرچه این روش امکان کنترل دقیق بر نوع داده‌ها را فراهم می‌کند، اما به‌دلیل ایجاد بار اضافی بر شبکه و هزینه‌های بالا، برای تولید آمارهای جمعیتی کمتر توصیه می‌شود. در این روش، بسته‌ها یا سیگنال‌های آزمایشی به طور عمدی در شبکه ارسال می‌شوند. سپس با اندازه‌گیری زمان رفت و برگشت (RTT) [26]، تأخیر، نرخ از دست رفتن بسته‌ها و پهنای باند، کیفیت شبکه یا کانال مخابراتی بررسی می‌شود.

در مقابل، **فراکاوی غیرفعال**[27] به رصد جریان واقعی داده‌ها بین اجزای شبکه می‌پردازد. این روش بدون ایجاد ترافیک مصنوعی، داده‌های واقعی مشترکین را در لحظه ثبت می‌کند و به‌دلیل دقت بالا و عدم اختلال در عملکرد شبکه، معمولاً ترجیح داده می‌شود. پیاده‌سازی این روش نیازمند نصب نرم‌افزارهای دارای مجوز و سخت‌افزارهای خاص در نقاط مختلف شبکه است. این سامانه‌ها می‌توانند به پایگاه‌هایی مانند ثبت موقعیت بازدیدکنندگان (VLR) [28]، نیز دسترسی داشته باشند و داده‌های مرتبط با موقعیت مشترکین را استخراج کنند. از مزایای فراکاوی غیرفعال می‌توان به دسترسی داده‌هایی که معمولاً در سامانه‌های صورتحساب یا انبار داده ذخیره نمی‌شوند و افزایش تعداد رکوردهای ثبت‌شده برای هر مشترک، اشاره کرد که به کاهش نوسانات روزانه و شبانه که ناشی از الگوهای تماس مشترکان است کمک می‌کند. با این حال، استفاده از داده‌های حاصل از فراکاوی

[15] Billing Domain
[16] Data Warehouse
[17] Probing
[18] Base Station Subsystem
[19] Network Switching Subsystem
[20] Central Storage Systems
[21] Call Detail Record
[22] Internet Protocol Detail Record
[23] Data Detail Record
[24] Customer databases
[25] Active Probing
[26] Round-Trip Time
[27] Passive Probing
[28] Visitor Location Register

و داده‌های سیگنال‌دهی با چالش‌هایی نیز همراه است. به‌عنوان مثال، سامانه‌های فراکاوی باید به‌صورت پیوسته فعال باشند تا داده‌های موقتی را ثبت و ذخیره کنند. همچنین همه اطلاعات برای اپراتور، الزامی به ذخیره شدن ندارند؛ برخی اطلاعات مانند **شناسه ناحیه موقعیت‌یابی** (LAI) [۲۹] الزامی هستند، اما شناسه دقیق **شناسه جهانی سلول** (CGI) [۳۰] ممکن است ذخیره نشود. این فعالیت باعث ایجاد بار اضافی بر شبکه اپراتور می‌شود. از طرف دیگر حجم بالای داده‌ها نیازمند منابع پردازشی و ذخیره‌سازی گسترده است. همچنین داده‌های سیگنال‌دهی ممکن است نویز آماری بالایی تولید کنند که باید در مرحله پردازش مدیریت شود. از طرف دیگر، نصب این سامانه‌ها هزینه‌بر است و معمولاً صرفاً برای تولید شاخص‌های آماری نصب نمی‌شوند. اما اگر اپراتورها از پیش نوعی سامانه پایش و زیرساخت ذخیره‌سازی داده را پیاده‌سازی کرده باشند، این داده‌ها می‌توانند منبع بسیار ارزشمندی برای شاخص‌های آماری باشند، زیرا حجم داده‌های بسیار بیشتری را نسبت به ثبت جزئیات تماس در سیستم‌های صورتحساب یا انبار داده‌ها در بر دارند. لازم به ذکر است که داده‌های پایش به داده‌های مربوط به رومینگ ورودی و داده‌های داخلی محدود می‌شوند. به‌طور معمول داده‌های سیگنال‌دهی میان اپراتورها مبادله نمی‌شوند، به همین دلیل داده‌های رومینگ خروجی در آن وجود ندارد.

| منبع داده‌ها | | چقدر تولید آن آسان است؟ | چه چیزی تولید می‌شود؟ | چه مقدار داده تولید می‌شود؟ | چقدر برای آمار مناسب است؟ |
|---|---|---|---|---|---|
| داده‌های غیرفعال | فعالیت سلولی | آسان و استاندارد | شدت استفاده از تلفن در سطح سلول | ● | ● |
| | سوابق جزئیات تماس (CDR+DDR) | نیاز به حفاظت از حریم خصوصی و روش‌های پردازش دارد. | فعالیت‌های سلولی | ●● | ●● |
| | پایش یا فراکاوی/ سیگنال‌دهی | نیازمند جمع‌آوری، حفاظت از حریم خصوصی و روش‌های پردازش است. | فعالیت‌های سلولی و ثبت‌های جابه‌جایی بین سلول‌ها؛ ویژگی‌های فردی استخراج‌شده از اپراتور | ●●● | ●●●● |
| داده‌های فعال | سیستم موقعیت‌یابی تلفن همراه | برای نمونه‌های کوچک آسان است، معمولاً نیاز به رضایت کاربر دارد که موافقت کند تا داده‌هایش جمع‌آوری یا استفاده شوند. | موقعیت‌یابی دقیق با بسامد نمونه‌برداری قابل تنظیم؛ امکان اجرای پرسش‌نامه با مشارکت پاسخ‌گو | ●●●●● | ●● |

شکل ۲: خلاصه‌ای از اصلی‌ترین منابع داده‌های موقعیت‌یابی تلفن همراه [۲۰]

### ۳.۱.۲ داده‌های موقعیت‌یابی فعال

داده‌های موقعیت‌یابی فعال به داده‌هایی اطلاق می‌شود که از طریق فرآیندهای هدفمند و مستقیم برای تعیین موقعیت مکانی مشترکین تلفن همراه تولید می‌شوند. این داده‌ها معمولاً از طریق روش‌های **مبتنی بر دستگاه**[۳۱] یا **مبتنی بر شبکه**[۳۲] ، و همچنین سامانه‌های ماهواره‌ای GPS یا **سامانه موقعیت‌یابی جهانی کمکی**[۳۳](A-GPS) به‌دست می‌آیند. این روش‌ها یا برای ارائه خدمات مبتنی بر مکان استفاده می‌شوند، یا در پاسخ به مقررات ملی که مستلزم گردآوری داده‌های موقعیت با دقت بالا (مثلاً برای اهداف نجات یا پلیس) هستند، به‌کار گرفته می‌شوند. در کشورهای در حال توسعه، توانایی و قابلیت‌های موقعیت‌یابی فعال اغلب به‌صورت موردی و در نمونه‌های کوچک برای مطالعات خاص یا ارزیابی‌های تطبیقی به کار گرفته می‌شوند. محققان از موقعیت‌یابی فعال تلفن همراه با نمونه‌های کوچک برای معیارسنجی و در مطالعاتی که نیاز به موقعیت‌یابی دقیق و مداوم

[29] Location Area Identity
[30] Cell Global Identity
[31] Device-centric
[32] Network-centric
[33] Assisted GPS
[34] Cellular Triangulation

دارند، استفاده کرده‌اند. روش‌های رایج در موقعیت‌یابی فعال شامل استفاده از **مثلث‌بندی سلولی**[۳۴] است که در آن فاصله نسبی مشترک با چند ایستگاه پایه برای برآورد موقعیت به‌کار می‌رود. همچنین، بهره‌گیری از سامانه موقعیت‌یابی جهانی، امکان موقعیت‌یابی دقیق با استفاده از ماهواره‌ها را فراهم می‌سازد، که گاهی با کمک شبکه برای تسریع فرآیند همراه است (شکل ۲). با وجود مزایای تحلیلی، داده‌های موقعیت‌یابی فعال معمولاً نیازمند رضایت مشترکین هستند و به‌دلیل محدودیت در حجم نمونه، برای تولید آمارهای ملی به‌تنهایی کافی نیستند. با این حال، در ترکیب با سایر منابع داده‌ای، می‌توانند نقش مکمل مؤثری ایفا کنند. در شکل ۲، منابع اصلی داده‌های موقعیت‌یابی تلفن همراه با یکدیگر مقایسه شده‌اند.

## ۲.۲ داده‌های رویدادهای تلفن همراه، داده‌های موقعیت‌یابی غیرفعال

نوع دیگری از داده‌های تلفن همراه، داده‌های رویدادهای تلفن همراه است که مربوط به داده‌های موقعیت‌یابی غیرفعال است. این داده‌ها به دو دسته اصلی داده‌های داخلی [۳۵] و داده‌های رومینگ یا فراگَردی [۳۶] تقسیم می‌شوند. هر یک از این دسته‌ها ویژگی‌ها و کاربردهای خاص خود را دارند که در ادامه به معرفی آن‌ها پرداخته می‌شود.

### ۱.۲.۲ داده‌های داخلی

داده‌های داخلی به داده‌هایی اطلاق می‌شود که از دستگاه‌های تلفن همراه در داخل یک کشور و به‌طور معمول در شبکه‌های داخلی شرکت‌های مخابراتی گردآوری می‌شوند. این داده‌ها معمولاً شامل اطلاعات مکانی و ترافیکی کاربران در محدوده داخلی کشور است و برای تحلیل‌های جمعیتی و جغرافیایی به‌کار می‌روند. این داده‌ها انواع مختلفی دارند؛ یکی از این انواع، داده‌های موقعیت مکانی است که شامل اطلاعات مربوط به موقعیت جغرافیایی کاربران است که از طریق GPS ، **شبکه بی‌سیم محلی (Wi-Fi)**[۳۷]، یا سلول‌های مخابراتی گردآوری می‌شود. داده‌های مصرفی نیز نوع دیگری از داده‌های داخلی است که شامل اطلاعات مربوط به میزان مصرف اینترنت و خدمات تلفن همراه است که توسط کاربران در اپلیکیشن‌های مختلف وجود دارد و نهایتاً داده‌های مربوط به ترافیک و حرکت که شامل داده‌هایی است که از اپلیکیشن‌های نقشه و مسیریابی گردآوری می‌شوند و برای تحلیل ترافیک و بهبود حمل‌ونقل شهری استفاده می‌شوند. کاربرد داده‌های داخلی در تحلیل جابجایی و حرکت جمعیت است. داده‌های موقعیت مکانی می‌توانند برای تحلیل الگوهای جابجایی و تردد جمعیت در داخل کشور استفاده شوند. این اطلاعات برای برنامه‌ریزی شهری و طراحی سیستم‌های حمل‌ونقل عمومی مفید هستند. کاربرد دیگر این داده‌ها برای تحلیل رفتارهای مصرفی و پیش‌بینی الگوهای خرید و مصرف در کشور است. بهبود خدمات بهداشتی و بحران‌ها، زمینه دیگر استفاده از داده‌های داخلی است. داده‌های موقعیت مکانی همچنین در بحران‌ها (مثل بلایای طبیعی یا اپیدمی‌ها) برای ردیابی و کمک به برنامه‌ریزی بهتر برای ارائه خدمات استفاده می‌شود.

در فرانسه، از داده‌های داخلی شبکه‌های تلفن همراه برای تحلیل حرکت شهروندان در ایام تعطیلات و بحران‌های طبیعی استفاده می‌شود. این اطلاعات برای بهبود مدیریت ترافیک و خدمات اورژانس استفاده شده است.

### ۲.۲.۲ داده‌های رومینگ

داده‌های رومینگ به داده‌هایی اطلاق می‌شود که از طریق دستگاه‌های تلفن همراه گردآوری، اما در شبکه‌های مخابراتی خارجی (خارج از مرزهای کشور) استفاده می‌شوند. این داده‌ها شامل اطلاعات مکانی، ترافیکی و مصرفی کاربران زمانی است که در سفرهای بین‌المللی یا مناطق مختلف قرار دارند. این داده‌ها در دو بخش قابل بررسی هستند. اول **رومینگ ورودی** [۳۸] که به داده‌هایی گفته می‌شود که زمانی گردآوری می‌شوند که کاربران خارجی وارد شبکه مخابراتی یک کشور می‌شوند. به عبارت دیگر، این داده‌ها مربوط به افرادی است که به عنوان مسافر یا گردشگر به کشور شما وارد می‌شوند و از خدمات تلفن همراه شبکه داخلی استفاده می‌کنند. رومینگ ورودی در سه بخش دسته‌بندی می‌شود؛ دسته اول داده‌های موقعیت مکانی است که زمانی که یک کاربر خارجی به کشور وارد می‌شود، موقعیت مکانی او در شبکه‌های مخابراتی داخلی ثبت می‌شود. دسته دوم داده‌های مصرفی شامل میزان استفاده از اینترنت، تماس‌ها، پیامک‌ها، و دیگر خدمات مخابراتی توسط کاربران خارجی است و دسته سوم داده‌های ترافیکی است که مربوط به اطلاعات در مورد جابجایی کاربران خارجی بین مناطق مختلف کشور است. بطور کلی داده‌های رومینگ ورودی می‌توانند به‌طور مستقیم در تحلیل تعداد

[35] Domestic Data
[36] Roaming Data
[37] Wireless Fidelity
[38] Inbound Roaming

گردشگران و جابجایی‌های بین‌المللی در یک کشور کمک کنند و یا از این داده‌ها می‌توان برای پیش‌بینی نیازهای خدماتی در مناطق خاص (مثلاً مناطق گردشگری) استفاده کرد. همچنین در مواقع بحران یا بلایای طبیعی، این داده‌ها می‌توانند به شناسایی محل دقیق تجمع گردشگران و کمک به مدیریت اضطراری کمک کنند.

نوع دوم داده‌های رومینگ، **رومینگ خروجی** [۳۹] است و مربوط به داده‌هایی است که کاربران داخلی از کشور خارج می‌شوند و در شبکه‌های مخابراتی کشورهای دیگر از خدمات تلفن همراه استفاده می‌کنند. این داده‌ها شامل اطلاعات مربوط به موقعیت مکانی، مصرف اینترنت، و تماس‌های خارجی است. انواع داده‌های رومینگ خروجی عبارتند از داده‌های موقعیت مکانی کاربران داخلی که در سفر به کشورهای دیگر هستند. داده‌های مصرفی نوعی از رومینگ خروجی هستند که شامل اطلاعات در مورد میزان استفاده از خدمات مخابراتی در خارج از کشور، از جمله اینترنت و تماس‌های بین‌المللی و همچنین داده‌های ترافیک که اطلاعاتی در مورد جابجایی و موقعیت کاربران در کشورهای مختلف ارائه می‌دهند. کاربرد رومینگ خروجی در تحلیل مهاجرت و جابجایی‌های بین‌المللی، تحلیل رفتار مصرفی در سطح جهانی که به شرکت‌ها و کسب‌وکارها کمک می‌کند تا رفتار مصرفی کاربران خود را در بازارهای بین‌المللی تحلیل کنند و مدیریت هزینه‌ها و قیمت‌گذاری خدمات بین‌المللی است که اپراتورهای تلفن همراه می‌توانند از این داده‌ها برای بهینه‌سازی تعرفه‌های رومینگ و خدمات بین‌المللی خود استفاده کنند. به عنوان مثال، در کشور استونی، داده‌های تلفن همراه برای تولید آمار رسمی گردشگری مورد استفاده قرار گرفته‌اند. طبق گزارش گروه کاری سازمان ملل، بانک مرکزی استونی از سال ۲۰۰۸ با اپراتورهای تلفن همراه همکاری کرده تا داده‌های موقعیت‌یابی تلفن همراه را برای برآورد سفرهای ورودی و خروجی به‌کار گیرد. این داده‌ها نشان دادند که گردشگران خارجی عمدتاً در منطقه تاریخی تالین متمرکز هستند. این یافته موجب شد تا کمپین‌های تبلیغاتی برای جذب گردشگر به شهرهایی مانند پارنو و تارتو طراحی شود [۲۰].

در کشور آندورا، در مطالعه‌ای با استفاده از داده‌های تماس تلفن همراه (CDR) به تحلیل رفتار گردشگران در ۲۲ رویداد بزرگ شامل مسابقات ورزشی، اجراهای فرهنگی و جشنواره‌های موسیقی پرداخت. این داده‌ها امکان سنجش دقیق شاخص‌هایی مانند تراکم جمعیت، مدت اقامت، و تأثیرات اقتصادی رویدادها را فراهم کردند. نتایج نشان داد که داده‌های تلفن همراه می‌توانند به‌عنوان ابزار تجویزی و تشخیصی برای بهینه‌سازی برنامه‌ریزی گردشگری و ارزیابی اثربخشی کمپین‌های بازاریابی مورد استفاده قرار گیرند [۱۱]. بطورکلی این داده‌ها به‌طور گسترده در تحلیل‌های آماری، برنامه‌ریزی شهری، پیش‌بینی رفتار مصرفی، نظارت بر بحران‌ها و تحلیل تحرکات جمعیتی کاربرد دارند. استفاده بهینه از این داده‌ها می‌تواند به کشورهای مختلف در زمینه‌های مختلفی مانند مدیریت منابع، بهبود خدمات شهری، و پیش‌بینی روندهای اقتصادی و اجتماعی کمک کند تا به نتایج دقیق‌تری دست یابند.

در تحلیل‌های جمعیتی و مکانی مبتنی بر داده‌های تلفن همراه، دو دسته از ویژگی‌ها شامل ویژگی‌های زمانی و مکانی نقش بنیادی دارند. این ویژگی‌ها در فرآیند صدور صورتحساب و ثبت رویدادهای شبکه به‌صورت ساختاریافته ذخیره می‌شوند و طبق استانداردهای فنی $3GPP$ [۴۰]، الزامات خاصی برای ثبت آن‌ها تعریف شده است. در زمینه ویژگی‌های زمانی، در فرآیند تولید ثبت جزئیات تماس و ثبت جزئیات پروتکل اینترنت، ثبت زمان وقوع رویداد یکی از الزامات اصلی است. طبق استاندارد $3GPP$ (ETSI TS 132 250)، هر رکورد باید شامل حداقل چهار مؤلفه زمانی تاریخ، ساعت، دقیقه و ثانیه باشد. این دقت زمانی برای تحلیل‌های جمعیتی، الگوهای جابه‌جایی، و رفتارهای ارتباطی ضروری است. در فرآیند تماس، زمان‌های مختلفی توسط اجزای شبکه تولید و ثبت می‌شوند که هر یک نقش خاصی در تحلیل‌های ارتباطی دارند. استاندارد $3GPP$ سه نوع متفاوت از برچسب‌های زمانی برای مدیریت تماس تعریف می‌کند:

- **زمان اشغال یا زمان تصرف خط** [۴۱] زمانی است که منابع شبکه برای ارائه یک خدمت به مشترک اشغال/تصرف می‌شوند؛ این مؤلفه برای تماس‌های ناموفق الزامی است.
- **زمان پاسخ به تماس** [۴۲] که بیانگر موفقیت اتصال و پاسخ‌گویی گیرنده است و برای تماس‌های موفق الزامی تلقی می‌شود.
- **زمان آزادسازی منابع** [۴۳] زمانی است که منابع اشغال‌شده دوباره آزاد می‌شوند. به عبارت دیگر، لحظه‌ای است که بعد از پایان تماس یا قطع ارتباط، همان منابع دوباره آزاد می‌شوند تا برای مشترک دیگری مورد استفاده قرار گیرند. این مؤلفه اختیاری است و معمولاً در

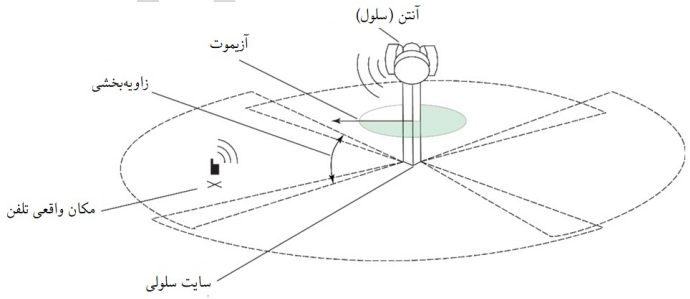
[39] Outbound Roaming
[40] 3rd Generation Partnership Project's
[41] Seizure Time
[42] Answer Time
[43] Release Time

تحلیل‌های تکمیلی مورد استفاده قرار می‌گیرد. این زمان‌ها توسط اجزای شبکه‌ای که خدمات ارتباطی را ارائه می‌دهند تولید می‌شوند و معمولاً در پایگاه‌هایی مانند پایگاه ثبت موقعیت خانگی (HLR) [۴۴]، یا پایگاه ثبت موقعیت بازدیدکننده (VLR) ذخیره نمی‌شوند، بلکه در فرآیند صدور صورتحساب و هنگام وقوع رویداد قابل شارژ ثبت می‌گردند.

در زمینه ویژگی‌های مکانی، ارجاع جغرافیایی دقیق به رویدادها از طریق **شناسه سلول (CID)**[۴۵] یا محدوده پوشش سلول انجام می‌شود. در شبکه‌های تلفن همراه، CID یک شماره یا کد یکتاست که هر سلول (آنتن یا دکل مخابراتی) را در شبکه مشخص می‌کند. این شناسه به اپراتور کمک می‌کند بداند یک مشترک دقیقاً به کدام سلول (دکل یا بخش شبکه) متصل است. طبق استانداردهای $3GPP$ در داده‌های داخلی و رومینگ ورودی، شناسه سلول یا موقعیت آنتن به‌عنوان مرجع مکانی ثبت می‌شود. اما در داده‌های رومینگ خروجی، الزام به ثبت موقعیت وجود ندارد و معمولاً فقط کد کشور ثبت می‌شود.

برای سامانه‌های سلولی، کوچک‌ترین عنصر ساختاری، یک سلول شبکه است. هر سلول شبکه دارای مجموعه‌ای از ویژگی‌های فنی و مکانی است که در تحلیل‌های آماری باید مورد توجه قرار گیرد. این ویژگی‌ها شامل **آزیموت**[۴۶] (زاویه‌ای که آنتن سلول نسبت به شمال جغرافیایی دارد)، **زاویه‌بخشی** [۴۷] (محدوده زاویه‌ای است که آنتن پوشش می‌دهد)، شکل و اندازه ناحیه پوشش، نوع آنتن (تک‌جهته، چندجهته یا هوشمند)، و موقعیت مکانی یا مختصات دقیق محل نصب آنتن یا سلول هستند. شبکه مخابرات تلفن همراه می‌تواند شامل هزاران سلول مجزا باشد. از آنجا که هر سلول تنها می‌تواند تعداد محدودی از کاربران را پشتیبانی کند، سلول‌ها در سطح منطقه تحت پوشش اپراتور تلفن همراه به‌طور یکنواخت توزیع نشده‌اند. به‌گونه‌ای که در مناطق شهری با تراکم بیشتری قرار دارند، در حالی که در مناطق روستایی، سلول‌ها به‌صورت پراکنده مستقر شده‌اند. این تفاوت در تراکم سلولی باید در مدل‌سازی مکانی و تحلیل‌های جمعیتی لحاظ شود تا نتایج آماری دچار سوگیری نشوند.

یک دکل یا برج سلولی، مجموعه‌ای از فرستنده‌ها در یک مکان مشخص است که می‌توان از آن با عنوان سایت سلولی [۴۸] یاد کرد (شکل ۳). این بدان معناست که یک مکان می‌تواند چندین سلول مختلف را زمانی که آن‌ها را بر اساس مختصات جغرافیایی گروه‌بندی کنیم، در خود جای دهد. سلول‌های مجاور، معمولاً ۳۰ یا ۴۰ سلول، در یک ناحیه مکانی (LA) [۴۹] گروه‌بندی می‌شوند. این کار برای متعادل‌سازی بار سیگنال‌دهی انجام می‌شود، به‌طوری که دستگاه تنها زمانی نیاز به ارسال اطلاعات مربوط به موقعیت کاربر (به‌روزرسانی ناحیه مکانی) دارد که وارد ناحیه مکانی دیگری شود، و به‌صورت دوره‌ای در بازه‌های زمانی از پیش تعیین‌شده قرار داشته باشد.




شکل ۳: سلول، سایت سلولی و منطقه مکانی، و روابط آن‌ها با یکدیگر [۲۰]

[44] Home Location Register
[45] Cell Identity
[46] Azimuth
[47] Sector Angle
[48] Cell Site
[49] در شبکه‌های LTE ، که به شبکه‌های $4G$ اشاره دارد، به جای اصطلاح «ناحیه مکانی» (LA) از اصطلاح «ناحیه ردیابی» (TA) استفاده می‌شود.

# ۳ چارچوب‌های قانونی استفاده از داده‌های موقعیت‌یابی تلفن همراه در آمار رسمی

یکی از مهم‌ترین پیش‌نیازهای بهره‌گیری از داده‌های موقعیت‌یابی تلفن همراه در تولید آمار رسمی، وجود چارچوب‌های حقوقی و اخلاقی شفاف و قابل استناد است. اطلاعات مکانی در دسته‌ی داده‌های حساس قرار می‌گیرند و چنانچه بدون سازوکار قانونی روشن، مدیریت یا منتشر شوند، می‌توانند منجر به نقض حریم خصوصی، بی‌اعتمادی عمومی و حتی بحران‌های اجتماعی شوند.

استفاده از داده‌های تلفن همراه در تولید آمار رسمی و تحلیل‌های اجتماعی و اقتصادی، در کنار فرصت‌های بی‌شماری که ارائه می‌دهد، با چالش‌های قانونی و مقرراتی نیز مواجه است. این چالش‌ها شامل مسائلی مانند حریم خصوصی افراد، حفاظت از داده‌ها، و حقوق مالکیت معنوی می‌شود.

با گسترش استفاده از داده‌های شخصی و حساس در دنیای دیجیتال، قوانین مختلفی در سطح بین‌المللی برای حفاظت از حریم خصوصی افراد و مدیریت داده‌های شخصی ایجاد شده است. یکی از معروف‌ترین و جامع‌ترین قوانین در این زمینه، مقررات عمومی حفاظت از داده‌های اتحادیه اروپا است که در سال ۲۰۱۸ اجرایی شد. این مقررات به‌ویژه بر گردآوری، ذخیره‌سازی و پردازش داده‌های شخصی تاکید دارد و از جمله مهم‌ترین اصول آن می‌توان به شایستگی و شفافیت در استفاده از داده‌ها، حق دسترسی و اصلاح داده‌ها و حق حذف داده‌ها اشاره کرد.

طبق این مقررات، هرگونه استفاده از داده‌های موقعیت‌یابی باید با اطلاع و رضایت کاربر باشد و علاوه بر این، فرد باید از حق دسترسی به داده‌های خود و همچنین امکان حذف آنها برخوردار باشد. همچنین، این قانون بر اهمیت امنیت داده‌ها و رعایت ضوابط محافظت از داده‌ها تأکید دارد و سازمان‌ها را ملزم می‌کند که اقدامات امنیتی مناسبی برای جلوگیری از نقض داده‌ها انجام دهند [۴].

برخی نهادهای بین‌المللی، راهنماهای مشخصی برای استفاده از داده‌های تلفن همراه ارائه کرده‌اند تا ضمن حفظ امنیت ملی، از حریم خصوصی نیز صیانت شود. نمونه موفق این رویکرد در کشور استونی دیده می‌شود. در این کشور، داده‌های مکانی تلفن همراه، پس از پردازش و حذف اطلاعات شناسایی فردی برای تولید آمار گردشگری و جمعیت استفاده می‌شود. داده‌ها به‌صورت تجمیعی و ناشناس در اختیار نهادهای آماری قرار می‌گیرند و شهروندان می‌توانند گزارش‌های استفاده از داده‌های خود را مشاهده کنند. این شفافیت و کنترل، اعتماد عمومی را افزایش داده و استفاده از داده‌های تلفن همراه را در چارچوب قانونی ممکن ساخته است [۲۰].

در سطح ملی، بسیاری از کشورها قوانین خاصی برای حفاظت از حریم خصوصی و داده‌های شخصی دارند. به عنوان مثال، در ایالات متحده، قانون حفاظت از حریم خصوصی اینترنت و قانون حفاظت از اطلاعات شخصی و داده‌های الکترونیکی، از جمله قوانین مهمی هستند که بر نحوه گردآوری، استفاده و نگهداری اطلاعات شخصی تأثیرگذارند. این قوانین در راستای محافظت از اطلاعات حساس افراد و کنترل دسترسی به آن‌ها طراحی شده‌اند.

## ۱.۳ چالش‌های قانونی در استفاده از داده‌های تلفن همراه

در استفاده از داده‌های موقعیت‌یابی تلفن همراه، چندین چالش قانونی وجود دارد که به‌ویژه در حریم خصوصی و مالکیت داده‌ها مطرح است. یکی از این چالش‌ها، گردآوری داده‌های موقعیت جغرافیایی بدون رضایت کاربران است. اگر کاربران از گردآوری داده‌های موقعیت خود آگاه نباشند یا این گردآوری بدون رضایت آنها انجام شود، این امر می‌تواند نقض قوانین حفاظت از داده‌ها محسوب شود.

چالش دیگر مربوط به داده‌های ناشناس است. در حالی که بسیاری از داده‌های گردآوری‌شده از تلفن‌های همراه به نظر ناشناس هستند، اما با تجزیه‌وتحلیل این داده‌ها می‌توان به اطلاعات شخصی حساس افراد دست پیدا کرد. برای مثال، می‌توان از داده‌های موقعیت‌یابی برای شناسایی الگوهای سفر روزانه افراد، مکان‌های خاصی که آنها به طور منظم به آنجا می‌روند، و حتی ارتباطات اجتماعی آنها بهره برد. این امکان باعث می‌شود که داده‌های به ظاهر ناشناس، شخصی‌سازی شده و قابل شناسایی شوند، که این امر چالشی جدی برای قوانین حریم خصوصی است.

از دیگر چالش‌های قانونی می‌توان به دسترسی به داده‌ها توسط مقامات دولتی اشاره کرد. در برخی کشورها، قوانین اجازه می‌دهند که مقامات دولتی برای اهداف امنیتی یا قضائی به داده‌های موقعیت‌یابی دسترسی داشته باشند. این امر می‌تواند با حقوق کاربران در تضاد باشد و نگرانی‌های زیادی را در زمینه نظارت دولتی و آزادی‌های فردی ایجاد کند[۲۲].

## ۲.۳ حقوق مالکیت معنوی در داده‌های تلفن همراه

یکی از مسائل دیگر در چارچوب‌های قانونی، مربوط به مالکیت داده‌ها است. داده‌های گردآوری‌شده از تلفن‌های همراه می‌توانند شامل اطلاعات حساس و ارزشمندی باشند که متعلق به شرکت‌های مخابراتی، شرکت‌های مخابراتی تلفن همراه یا حتی خود کاربران است. در این راستا، سوالات بسیاری در خصوص حقوق مالکیت معنوی این داده‌ها وجود دارد.

برخی از سوالات کلیدی در این زمینه عبارتند از: آیا کاربران باید حق مالکیت بر داده‌های خود را داشته باشند؟ آیا شرکت‌های مخابراتی می‌توانند به‌طور قانونی این داده‌ها را برای اهداف تجاری بفروشند؟ و چه ضمانت‌هایی برای حفاظت از داده‌ها در برابر سوءاستفاده‌های احتمالی وجود دارد؟

در ادبیات حقوقی، مالکیت داده‌ها به‌عنوان یک مفهوم نوظهور مطرح شده که هنوز در بسیاری از نظام‌های حقوقی به‌طور کامل تعریف نشده است. در ایالات متحده، برخی دادگاه‌ها، داده‌های شخصی را به‌عنوان دارایی قابل حفاظت در نظر گرفته‌اند، اما هنوز قانون فدرالی مشخصی برای مالکیت داده‌ها وجود ندارد. در اتحادیه اروپا، مقررات عمومی حفاظت از داده‌ها به‌جای تعریف مالکیت، بر «کنترل» داده‌ها توسط افراد تأکید دارد و حق دسترسی، اصلاح، حذف و انتقال داده‌ها را برای کاربران تضمین می‌کند[۳].

در اسناد بین‌المللی مرتبط با داده‌های تلفن همراه، از جمله فصل سوم راهنمای استفاده از داده‌های تلفن همراه در آمار رسمی منتشرشده توسط گروه کاری جهانی سازمان ملل، تأکید شده که دسترسی به داده‌های تلفن همراه باید در چارچوب‌های قانونی روشن، با رعایت اصول محرمانگی، و از طریق مشارکت چندجانبه بین اپراتورها، دولت‌ها و نهادهای آماری انجام شود [۲۰]. این اسناد به‌جای تمرکز بر مالکیت فردی، بر حکمرانی داده‌ها و مسئولیت‌پذیری نهادی تأکید دارند.

# ۴ روش‌شناسی تولید آمار از طریق داده‌های موقعیت‌یابی تلفن همراه

از داده‌های تلفن همراه می‌توان آمارهایی در حوزه‌های مختلف شامل آمار گردشگری، حساب‌های اقماری گردشگری، حمل‌ونقل مسافری، جمعیت، مهاجرت و آمار رفت‌وآمد استخراج کرد. داده‌های موقعیت‌یابی تلفن همراه ابزاری برای اندازه‌گیری مدت زمان سفر [۵۰]، تعداد بازدیدها، مدت زمان اقامت برحسب مجموع ساعات، روزهای حضور، و شب‌های سپری‌شده فراهم می‌کنند. این شاخص‌ها می‌توانند در تحلیل رفتارهای اجتماعی، بررسی الگوهای جابه‌جایی جمعیت، و تحلیل سایر پارامترهای اقتصادی و اجتماعی کمک کنند. آمارهای قابل دست‌یابی از داده‌های موقعیت‌یابی تلفن همراه به دو دسته اصلی شاخص‌های مکانی و زمانی تقسیم می‌شوند. شاخص‌های مکانی به تحلیل مکان‌های جغرافیایی مختلف و تغییرات آنها در طول زمان می‌پردازند، در حالی که شاخص‌های زمانی به بررسی تغییرات در زمان‌های مختلف و نحوه جابه‌جایی افراد در این بازه‌ها توجه دارند. این شاخص‌ها اغلب مکمل داده‌های سنتی مانند سرشماری یا حساب‌های ملی هستند و می‌توانند به‌صورت لحظه‌ای یا کوتاه‌مدت به‌روزرسانی شوند. در ژاپن، داده‌های موقعیت‌یابی GPS کاربران تلفن همراه اپراتور au برای تحلیل تغییرات حرکتی مردم در برخی مناطق کلیدی شهرها و بررسی ارتباط آن با شاخص‌های کلان اقتصادی مانند تولید ناخالص داخلی و مصرف به‌کار گرفته شد [۱۸].

به عنوان نمونه‌ای دیگر، **مؤسسه ملی آمار اسپانیا (INE)**[۵۱] با همکاری شرکت‌های مخابراتی تلفن همراه، از داده‌های شبکه برای شناسایی ورود گردشگران بین‌المللی از طریق مرزها استفاده می‌کند. این داده‌ها بر اساس موقعیت ثبت‌شده اولین برج مخابراتی در بدو ورود تحلیل می‌شود و به تفکیک ملیت، زمان ورود و مقصد اولیه تهیه شده است [۹]. همچنین برآورد مدت‌زمانی که یک گردشگر در یک منطقه جغرافیایی حضور دارد، از زمان ورود تا خروج، بر پایه تحلیل مداومت سیگنال‌های متصل به برج‌های سلولی در ناحیه خاص به‌دست می‌آید.

برای پیش‌بینی درست تقاضای حمل‌ونقل و سازمان‌دهی عرضه آن، تهیه برآوردهای دقیق ماتریس‌های مبدأ–مقصد حائز اهمیت است. با این حال، به‌دست آوردن این ماتریس‌ها از طریق روش‌های متعارف پیمایشی بسیار دشوار و پرهزینه است. بنابراین، اپراتورهای شبکه تلفن همراه (MNOs) می‌توانند ماتریس‌هایی با دقت بالاتر و هزینه کمتر از داده‌های موقعیت‌یابی تلفن همراه فراهم کنند. البته، این داده‌ها نوع وسیله نقلیه یا هدف سفر را نشان نخواهند داد.

داده‌های موقعیت‌یابی تلفن همراه با وجود دقت کمتر نسبت به سرشماری‌ها، به دلیل به‌روز بودن، می‌توانند اطلاعات مفیدی درباره

[50] Duration of the Trip

[51] Instituto Nacional de Estadística

رفت‌وآمد، مهاجرت و جابه‌جایی جمعیت ارائه دهند. ایجاد نظام آماری بر پایه این داده‌ها پرهزینه است، اما در صورت استفاده هم‌زمان در حوزه‌هایی مانند گردشگری، حمل‌ونقل و جمعیت، هزینه‌های اضافی کاهش می‌یابد. برای بهره‌گیری از این داده‌ها، تعریف دقیق متغیرها ضروری است، زیرا برخی شاخص‌ها قابل محاسبه دقیق طبق تعاریف رسمی نیستند.

برای تولید آمار با استفاده از داده‌های موقعیت‌یابی تلفن همراه، لازم است پس از گردآوری داده‌ها، فرآیندهایی مانند پیش‌پردازش، محاسبه شاخص‌ها و نهایتا ارزیابی داده‌ها صورت پذیرد که در شکل ۴ نمایش داده شده است. براساس این فرآیند، هر روز فایل‌های CDR طبق فرمت مشخص توسط اپراتورهای تلفن همراه آماده می‌شود. سپس، شناسه‌های حساس توسط اپراتور ناشناس و داده‌های ناشناس‌شده از طریق کانال امن به تیم تحلیل منتقل می‌شوند. سپس فرایند پیش‌پردازش و پاک‌سازی داده‌ها شامل بررسی مقادیر گم‌شده یا اشتباه، تبدیل فرمت‌ها، تطبیق شناسه سلول با جدول مختصات، و محاسبه آمار پایه مانند تعداد رکوردها و مشترکین روزانه، انجام می‌شود. سپس از داده‌های پیش‌پردازش شده برای برآورد جمعیت در نواحی مختلف استفاده می‌شود. تعیین مکان غالب کاربران و گردآوری تعداد مشترکین یکتا در هر ناحیه از جمله اقداماتی است که معمولا در مرحله بعد از پیش‌پردازش داده‌ها انجام می‌شود. برای بررسی جابجایی و تحرکات کاربران بین نواحی مختلف، توالی زمانی رکوردهای هر کاربر بررسی می‌شود و ماتریس‌های مبدأ–مقصد برای بازه‌های زمانی موردنظر تولید می‌شود. از خروجی مراحل قبل، شاخص‌هایی مانند تعداد مشترکین یکتا در ناحیه، ماتریس‌های مبدأ–مقصد و شاخص‌های جریان یا تحرکات جمعیت تولید می‌شود. در مرحله آخر، برای ارزیابی نتایج، شاخص‌های محاسبه شده براساس داده‌های موقعیت‌یابی تلفن‌های همراه با نتایج حاصل از آمارگیری‌های نمونه‌ای و سرشماری‌ها یا سایر منابع ثبتی مقایسه می‌شود و در صورت مغایرت، خطاها و محدودیت‌ها شناسایی می‌شود نهایتا با کم‌پوششی و بیش‌پوششی مقابله می‌شود. در ادامه برخی از این مراحل با جزئیات بیشتری معرفی می‌شود.

شکل ۴: فرآیند گردآوری، پیش‌پردازش، محاسبه و ارزیابی شاخص‌ها

## ۱.۴ روش‌شناسی پردازش داده‌ها

فرآیند پردازش داده‌های موقعیت‌یابی تلفن همراه برای تولید آمار گردشگری و سایر حوزه‌ها شامل مراحلی مانند پاک‌سازی داده‌های ناقص یا نادرست، آماده‌سازی تکمیلی داده‌های رویداد، تبدیل و فیلتر کردن داده‌ها، و محاسبه آمارهای پایه است. داده‌های اولیه توسط اپراتورهای شبکه تلفن همراه بر اساس رویدادهای شبکه استخراج

می‌شوند که حضور هر مشترک را در زمان و مکان مشخص ثبت می‌کنند. در مرحله آماده‌سازی تکمیلی، عملیات مختلفی مانند ارجاع جغرافیایی داده‌ها [۵۲] ، حذف دستگاه‌های غیرانسانی، بررسی پوشش زمانی و مکانی، و مدیریت داده‌های گم‌شده انجام می‌شود. پس از انجام این مراحل توسط اپراتورهای تلفن همراه، فرآیندهای پردازشی بعدی بر روی داده‌های آماده‌شده اجرا می‌گردد.

## ۲.۴ برآورد جمعیت مبدأ ـ مقصد

برای تهیه آمارهای جابجایی و سفر، پس از پاک‌سازی و پیش‌پردازش داده‌ها، با به‌کارگیری الگوریتم‌های شناسایی سفر، به شناسایی هر سفر مشترک به کشور مورد نظر همراه با زمان شروع و پایان هر سفر پرداخته می‌شود. سپس با تمایز قایل شدن بین فعالیت‌های گردشگری و غیرگردشگری، به تعیین تعاریف و مفاهیم مرتبط با شاخص‌های مورد نظر پرداخته می‌شود. به‌عنوان مثال، به مشخص کردن تعریف مشترکان رومینگ که در واقع از مرز عبور نکرده و وارد کشور نمی‌شوند (ورودی، خروجی)، تعریف ساکنان (ورودی، خروجی)، تعریف محل اقامت و محیط معمول زندگی (داخلی)، شناسایی سفرهای عبوری در سطح کشور (ورودی) و شناسایی کشورهای مقصد و ترانزیت (خروجی) پرداخته می‌شود. سپس به تعریف متغیرها (تعداد بازدیدها، مدت سفر، رده‌بندی‌ها و غیره) و نهایتاً برآورد جمعیت و برآورد مبدا – مقصد از طریق تجمیع داده‌ها در بُعد زمان و مکان (روز، هفته، ماه، فصل، شبکه‌بندی مکانی مانند یک کیلومتر مربع، نقاط روستایی، نقاط شهری، کشور) و ترکیب داده‌ها از اپراتورهای مختلف و نهایتاً محاسبه شاخص‌های نهایی آماری پرداخته می‌شود.

## ۳.۴ چارچوب تضمین کیفیت برای تولید آمار با داده‌های شبکه تلفن همراه

کیفیت نقشی بسیار مهم در ایجاد اعتماد به محصولاتی که توسط یک نهاد تولید می‌شوند، دارد. تضمین کیفیت بخشی جدایی‌ناپذیر از فرآیند تولید آمار در هر دفتر آماری است. در پروژه‌های داده‌های موقعیت‌یابی تلفن همراه، تضمین کیفیت به مسئله‌ای با اهمیت تبدیل شده است و چارچوب تضمین کیفیت (QAF) [۵۳]، بر اطمینان از کیفیت در هر مرحله از پردازش شامل

- تضمین کیفیت داده‌های ورودی
- تضمین کیفیت فرآیندهای آماری
- تضمین کیفیت محصول (خروجی)

تأکید دارد. از طریق این سه مرحلهٔ تضمین کیفیت، کیفیت داده‌های آماری تولیدشده به‌خوبی حفظ خواهد شد، چرا که در هر مرحله از فرآیند تولید آمار، بررسی و ارزیابی صورت می‌گیرد.

یک چارچوب تضمین کیفیت خوب، مفاهیم کیفیت مرتبط با تولید آمارهای رسمی را بازتاب می‌دهد (برای نمونه: ESS ۲۰۱۲ [۵۴] ، UN ۲۰۱۵ [۵۵] ، UNECE ۲۰۱۴ ) [۵۶].

چارچوب تضمین کیفیت باید به موارد زیر بپردازد:

۱. **محیط نهادی** قراردادها و توافق‌های سطح نهادی بین ذی‌نفعان، که به جزئیات ارتباطات، جنبه‌های مربوط به محرمانگی، مقررات، نقش‌ها و سایر پرسش‌های کلی می‌پردازد.

۲. **کیفیت ورودی** هدایت و راهنمایی کلی در این مرحله تضمین می‌کند که داده‌های خام ورودی برای پردازش آماده باشند. در اولین دسترسی یا هنگام وقوع تغییرات مهم در داده‌های خام، فراداده‌ها و شاخص‌های کیفیت شامل معیارهایی از دقت و خودانتخابی گردآوری می‌شوند. این مرحله بر فراداده‌های همراه با انتقال‌های منظم داده نیز تمرکز دارد.

۳. **کیفیت پردازش** جزئیات پردازش آماری عمدتاً به کیفیت روش‌شناسی و الگوریتم‌های پردازش شامل تأیید الگوریتم‌ها، اعتبار داده‌های مدل‌سازی‌شده، فراداده‌ها، ثبت و گزارش‌دهی در طول پردازش می‌پردازد.

۴. **کیفیت خروجی** جنبه‌های کیفیتی مربوط به نتایج پردازش داده‌ها، اعتبار فایل‌های داده منتقل‌شده، و پایبندی خروجی به اصول کیفیت که معمولاً از طریق سطوح سازگاری (قابلیت مقایسه در طول زمان، در مناطق جغرافیایی مختلف و با داده‌های مرجع) اندازه‌گیری می‌شوند.

[۵۲] به این معنا که داده‌ها یا رویدادهایی که ثبت شده‌اند (مثل اتصال تلفن همراه به دکل مخابراتی یا ثبت یک رویداد شبکه) به یک موقعیت جغرافیایی مشخص روی نقشه مرتبط شوند.

[53] Quality Assurance Framework

[54] EU, (2012), European Statistical System "Quality Assurance Framework of the European Statistical System"

$http://epp.eurostat.ec.europa.eu/cache/ITY_PUBLIC/QAF_2012/EN/QAF_2012-EN.PDF$

[55] UNECE (2014), A Suggested Framework for the Quality of Big Data

[56] United Nations (2015), Fundamental Principles of Official Statistics: Implementation guidelines

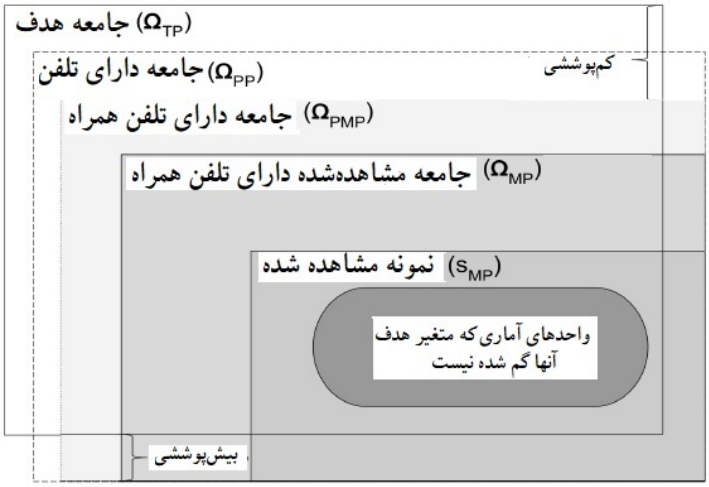
مرحله‌ی آخر – یعنی اندازه‌گیری کیفیت خروجی – آشناترین بخش برای ادارات آماری است. با این حال، اگر کیفیت داده‌ها در مراحل قبلی به شکلی تخصصی و منسجم تضمین شود، اندازه‌گیری‌های کیفیت خروجی نباید غافلگیرکننده باشند. آنچه اهمیت دارد، تضمین کیفیت در مراحل اولیه پردازش داده‌های موقعیت‌یابی تلفن همراه است؛ زیرا کیفیت خروجی بازتاب همین مراحل است.

هیچ استاندارد رسمی‌ای برای ایجاد چارچوب تضمین کیفیت داده‌های تلفن همراه وجود ندارد، اما در UNECE تلاش‌هایی در جریان است تا چارچوبی برای مه‌داده‌ها طراحی شود که با کمی تغییر برای داده‌های موقعیت‌یابی تلفن همراه نیز قابل استفاده است. در اندونزی، اداره آمار اندونزی [۵۷] یک چارچوب جامع تضمین کیفیت برای روش‌شناسی برآورد تعداد گردشگران ورودی فرامرزی از داده‌های موقعیت‌یابی تلفن همراه دارد.

داده‌های موقعیت‌یابی تلفن همراه با وجود ارزش بالای تحلیلی، مستعد بروز خطاهایی هستند که می‌توانند بر دقت آمار رسمی تأثیرگذار باشند. به‌عنوان نمونه، می‌توان به خطاهای موقعیت‌یابی ناشی از دقت محدود GPS یا وابستگی به دکل‌های مخابراتی موجب انحراف مکانی داده‌ها اشاره کرد. نرخ نمونه‌گیری نابرابر به‌دلیل تفاوت در تناوب ثبت داده میان کاربران یا مناطق، منجر به سوگیری زمانی و مکانی در تحلیل تحرک می‌گردد. همچنین، تأثیر پوشش شبکه در مناطق با سیگنال ضعیف یا ازدحام بالا باعث کاهش نمایندگی داده‌ها و برآورد نادرست از جمعیت یا الگوهای جابه‌جایی می‌شود. در نتیجه، کنترل و تعدیل این خطاها برای بهره‌برداری معتبر از داده‌های تلفن همراه در تولید آمار رسمی ضروری است. در ادامه با تمرکز بر خطاهای پوشش شبکه، برخی راه‌های مقابله با کم‌پوششی و بیش‌پوششی ارایه می‌شود.

## ۴.۴ مقابله با کم‌پوششی و بیش‌پوششی

در مورد داده‌های موقعیت‌یابی تلفن همراه، یک مغایرت از مقایسه تعریف جامعه هدف و تعریف جامعه‌ی چارچوب آشکار می‌شود. در جامعه‌ی چارچوب، تمام مشترکانی که از تلفن همراه استفاده می‌کنند، حضور دارند، در حالی که جامعه هدف شامل همه افرادی است که در کشور ساکن هستند.




شکل ۵: گذار جامعه هدف به جامعه مشاهده‌شده در داده‌های شبکه تلفن همراه [۲۰]

شکل ۵ رابطه بین جامعه هدف و نمونه نهایی به‌دست‌آمده را نشان می‌دهد. در اینجا بین جمعیت دارای تلفن (از هر نوع) و جمعیت دارای تلفن همراه تمایز قائل شده است. ممکن است اتفاق بیفتد که برای تمام واحدهای نمونه مشاهده‌شده، اطلاعاتی از متغیر هدف (مانند محل سکونت) در دسترس نباشد، که در این صورت باید جانهی صورت گیرد. در این حالت می‌توان مسئله‌ی متغیرهای پنهان را در نظر گرفت؛ متغیرهایی که مستقیماً قابل مشاهده نیستند اما به‌طور غیرمستقیم مثلاً از طریق تحلیل موقعیت با استفاده از ایستگاه‌های فرستنده–گیرنده پایه

[57] BPS Statistics Indonesia
[58] Base Transceiver Stations
[59] Selectivity

(BTS) [۵۸]، قابل استنباط هستند.

چرا خودانتخابی [۵۹]، در قالب زیرمجموعه‌سازی جامعه هدف مطرح می‌شود؟ فارغ از تعریف جامعه هدف (مثلاً سفرها، افراد)، برخی واحدها همیشه کنار گذاشته خواهند شد. برای نمونه، زیرساخت شبکه تلفن همراه امکان شناسایی دقیق همه سفرها را محدود می‌کند. با این حال، این موضوع باید به‌عنوان یک خطای پوشش در نظر گرفته شود. سفرهای مشاهده‌شده یک زیرمجموعه از کل جامعه سفرها هستند (مثلاً محدود به یک اپراتور یا فقط کسانی که تلفن همراه دارند) و شناسایی سفر نیز به پوشش زیرساخت بستگی دارد.

دو دلیل اصلی برای مشکلات پوشش و خودانتخابی عبارت‌اند از:

۱. پوشش جامعه هدف (صرف‌نظر از تعریف آن). باید توجه داشت که جمعیت مشاهده‌شده در میان مشترکان اپراتورهای تلفن همراه شامل اشخاص حقیقی و حقوقی است و این دو می‌توانند هم‌پوشانی داشته باشند (مثلاً یک فرد هم تلفن شخصی و هم تلفن کاری داشته باشد).

۲. مسئله دوم به موضوع خودانتخابی ایستگاه‌ها و سطحی از تجمیع داده‌ها که قابلیت استفاده دارند، مربوط است. ایستگاه‌های BTS به‌طور طبیعی متناسب با تراکم جمعیت نصب می‌شوند و در ادبیات علمی این موضوع عمدتاً با استفاده از نمودارهای ورونوی [۶۰]، نمایش داده می‌شود.

یکی از مهم‌ترین فعالیت‌ها در آغاز هر پروژه داده‌های موقعیت‌یابی تلفن همراه، تضمین کیفیت داده‌های ورودی است، تا توصیف بسیار دقیقی از مجموعه داده و جنبه‌های پوشش آن، از دیدگاه پروفایل مشترک، داده‌های رویداد و پوشش سلولی داشته باشیم.

جدول ۱: فهرست برخی مشکلات احتمالی پوشش داده‌ها از اپراتورهای موبایل (MNO)

| مسئله | کم‌پوششی/ بیش‌پوششی | راه‌حل ممکن |
|---|---|---|
| ۱. داده‌ها فقط از اپراتورهای منتخب شبکه تلفن همراه (MNO) در دسترس است؛ به عبارت دیگر، همه‌ی داده‌های اپراتورها موجود نیست. | کم‌پوششی | سهم سایر اپراتورها را می‌توان با استفاده از سایر منابع و اطلاعات درباره‌ی نرخ نفوذ و مشخصات مشتریان برآورد کرد. |
| ۲. افرادی که از تلفن همراه استفاده نمی‌کنند: | کم‌پوششی | در مرحله‌ی برآورد، از اطلاعات منابع اضافی برای ساخت و بررسی فرضیات مربوط به رفتار گروه کم‌پوشش، استفاده کنید و سپس از مدل‌های مناسب برای تعدیل برآوردها استفاده کنید. می‌توان از برآوردگرهای مدل‌مبنا یا برآوردگرهای مدل‌یار استفاده کرد. |
| ۳. استفاده از بیش از یک دستگاه تلفن همراه:<br>۱. همه‌ی دستگاه‌ها از همان شبکه استفاده می‌کنند و باعث تکرار در داده‌های یکی از اپراتورها (MNO) می‌شوند.<br>۲. دستگاه‌ها از شبکه‌های مختلف استفاده می‌کنند و باعث تکرار در داده‌های ارائه‌شده توسط اپراتورهای مختلف می‌شوند. | بیش‌پوششی | در مرحله‌ی برآورد، از اطلاعات منابع اضافی – در صورت موجود بودن – استفاده می‌شود تا فرضیات مربوط به تعداد افرادی که چند تلفن همراه دارند ساخته و بررسی شود. سپس از مدل‌های مناسب برای تعدیل برآوردها استفاده می‌شود. میزان تعدیل به فراوانی وقوع این پدیده بستگی دارد. |
| ۴. نرخ‌های نفوذ متفاوت اپراتورها (MNO) در میان مشترکان خارجی کشورهای خاص | کم‌پوششی یا بیش‌پوششی | با فرض در دسترس بودن اطلاعاتی که امکان ایجاد مدل‌های جداگانه بر پایه اپراتورهای تلفن همراه (MNO) با پارامترهای ویژه هر کشور را فراهم کند، می‌توان برآوردهای وزن‌دهی‌شده را محاسبه کرد. |
| ۵. ضریب‌های نفوذ منطقه‌ای و جمعیت‌شناختی متفاوت برای اپراتورها (MNO) | کم‌پوششی یا بیش‌پوششی | در صورت وجود، می‌توان از آمارگیری‌هایی که ضریب‌های نفوذ اپراتورها را بر اساس منطقه و ویژگی‌های جمعیت‌شناختی پوشش می‌دهند، در برآوردها استفاده کرد. |

[60] Voronoi Diagrams

به طور خلاصه، مشکلات زیر در خصوص استنباط مبتنی بر داده‌های تلفن همراه مشاهده می‌شوند:

- مشکلات مربوط به تعریف و استخراج جمعیت هدف مشاهده‌شده در داده‌های شبکه تلفن همراه
- دسترسی محدود به اطلاعات کاربران دستگاه‌های تلفن همراه (محل سکونت، جنسیت، سن، وضعیت تأهل)، و ضرورت جانهی یا پروفایل‌سازی
- شناسایی پوشش بیش از حد در چارچوب‌ها
- شناسایی واحدهای آماری
- شناسایی موقعیت‌های دقیق
- لحاظ کردن عدم قطعیت ناشی از جانهی در برآوردها

از برخی مشکلات پوشش، می‌توان در مرحله تشکیل چارچوب یا گردآوری داده‌ها (قبل از انجام برآورد) با شناسایی مشاهداتی که جزو جامعه هدف نیستند و حذف آن‌ها از چارچوب، اجتناب کرد. در جدول ۱ فهرست برخی مشکلات احتمالی پوشش داده‌ها از اپراتورهای موبایل (MNO) و راه‌حل‌های ممکن ارایه شده اشت.

# ۵ تحلیل مجموعه داده تحرک شهری K۱۰۰YJMob ژاپن

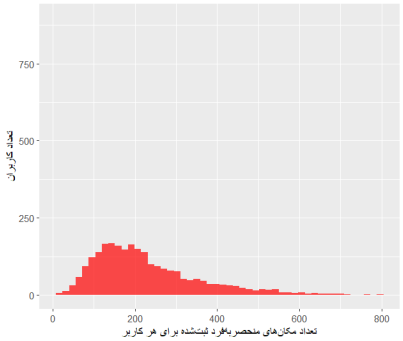
در این بخش به تحلیل داده‌های موقعیت‌یابی تلفن همراه در زمینه تولید آمار رسمی می‌پردازیم. هدف اصلی، تحلیل مجموعه داده تحرک شهری K۱۰۰YJMob در ژاپن با بهره‌گیری از شاخص‌ها و آمارهایی است که در ادامه معرفی می‌شوند. این داده‌ها برای مدل‌سازی حمل‌ونقل، مدیریت بحران، برنامه‌ریزی شهری و سایر حوزه‌ها مفید هستند.

این مجموعه داده از رکوردهای GPS گوشی‌های هوشمندی گردآوری شده‌اند که کاربران آنها اپلیکیشن‌های شرکت $YahooJapan$ که اکنون به $(LYCorporation)$ تغییر نام داده را روی گوشی خود نصب کرده بودند. این شرکت از طریق اپلیکیشن‌های خود، داده‌های مکانی را به صورت خودکار و با رضایت کاربران دریافت کرده است. کاربران هنگام نصب برنامه، با سیاست حفظ حریم خصوصی موافقت کرده‌اند که در آن ذکر شده داده‌های مکان به صورت ناشناس‌شده برای اهداف پژوهشی و تحلیلی در همکاری با شرکای تحقیقاتی استفاده می‌شود. هر رکورد مکانی، شامل شناسه یکتای کاربر، زمان مشاهده، طول و عرض جغرافیایی است. نرخ نمونه‌گیری داده‌ها تقریباً ۵ درصد از کل جمعیت ساکن در محدوده سلول شبکه براساس داده‌های سرشماری بوده است. فراوانی گردآوری داده‌های GPS بسته به سرعت حرکت کاربر تغییر می‌کند تا بار روی باتری گوشی هوشمند کاربر به حداقل برسد. اگر تشخیص داده شود که کاربر مدت طولانی در یک مکان ثابت حضور دارد، داده‌ها با فرکانس نسبتاً پایین گردآوری می‌شوند و اگر کاربر در حال حرکت باشد، داده‌ها با فرکانس بیشتری گردآوری می‌شوند. پس از گردآوری، داده‌ها برای حفظ حریم خصوصی، به صورت مکانی و زمانی پردازش و ناشناس‌سازی شده‌اند. برای این منظور، داده‌های خام با برش مکانی و زمانی محدود شده‌اند. ابتدا یک منطقه شهری ناشناس در ژاپن به‌عنوان محدوده انتخاب شده و تنها کاربرانی در مجموعه نهایی قرار گرفته‌اند که بیش از ۱۰ بار در طی یک بازه ۱۰ روزه در آن محدوده مشاهده شده باشند. برای حفظ محرمانگی، مختصات مکانی به شبکه‌ای از سلول‌های با ابعاد ۵۰۰ متر در ۵۰۰ متر و زمان به بازه‌های ۳۰ دقیقه‌ای تقسیم شده است و تاریخ واقعی مشاهدات نیز حذف شده است. داده‌های کاربران خارج از محدوده نیز حذف و در نهایت، برای مجموعه داده شماره ۱، تعداد ۱۰۰۰۰۰ کاربر و برای مجموعه داده شماره ۲، تعداد ۲۵۰۰۰ کاربر انتخاب شده‌اند. داده‌ها شامل اطلاعات مسیر حرکتی این کاربران در بازه‌ی ۷۵ روز است که در دو بازه‌ی زمانی شامل شرایط عادی و اضطراری گردآوری شده‌اند. در مجموعه داده اول، مسیرهای حرکتی در شرایط عادی به مدت ۷۵ روز ثبت شده است، و مجموعه داده دوم شامل ۶۰ روز داده‌های عادی و ۱۵ روز داده‌های مربوط به شرایط اضطراری است. با توجه به سیاست‌های حفظ حریم خصوصی، ماهیت دقیق شرایط اضطراری مشخص نیست.

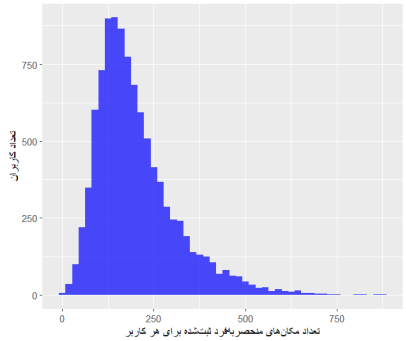
در این مجموعه داده، اطلاعات مسیرهای حرکتی افراد از طریق چندین متغیر کلیدی ذخیره شده است. متغیر `user_id` شناسه‌ای ناشناس برای هر فرد است که در طول مجموعه داده ثابت می‌ماند. متغیر day نشان‌دهنده‌ی شماره‌ی روز از بازه‌ی ۷۵ روزه، بدون نمایش تاریخ واقعی برای حفظ حریم خصوصی کاربران است. متغیر –times lot بازه‌های زمانی ۳۰ دقیقه‌ای را مشخص می‌کند و مقدار آن عددی بین ۰ تا ۴۷ است که ساعات مختلف روز را نشان می‌دهد. علاوه بر این، ثبت موقعیت یا سیگنال‌های مکانی کاربران در شبکه‌ای از سلول‌های ۵۰۰ متر در ۵۰۰ متر ثبت شده است که از طریق متغیرهای x و y نمایش داده می‌شود. مقدار این متغیرها بین ۱ تا ۲۰۰ متغیر است، که نشان‌دهنده‌ی موقعیت افقی و عمودی در سیستم شبکه‌ای است. ترکیب این دو متغیر امکان بررسی دقیق مسیرهای حرکتی افراد را، بدون اینکه مختصات جغرافیایی واقعی فاش شود، فراهم می‌کند.

در دوره‌ی ۱۵ روزه بحران، رفتار کاربران به طور قابل توجهی تغییر می‌کند. برخی کاربران کمتر حرکت می‌کنند، بیشتر در موقعیت‌های

خاص باقی می‌مانند یا استفاده‌ی کمتری از دستگاه‌های متصل به شبکه دارند. این موضوع باعث کاهش تعداد کل کاربران فعال در مجموعه داده شده است. مطالعات بین‌المللی نشان می‌دهند که یافته‌های این پژوهش با نتایج گزارش شده در سایر کشورها هم راستا است. برای نمونه، در اسپانیا، موسسه ملی آمار (INE) [۶۱]، با استفاده از داده‌های موقعیت‌یابی تلفن همراه، تحرک جمعیت و تغییرات مکانی آن را در مراحل مختلف همه‌گیری کووید–۱۹ مورد بررسی قرار دادند [۱۵]. نتایج این مطالعه نشان داد که در دوره‌های قرنطینه و محدودیت تردد، تعداد تلفن‌های همراه فعال در جریان سفرها به طور چشمگیری کاهش یافته و تنها در کلان‌شهرها تحرک محدودی مشاهده می‌شود. پس از لغو محدودیت‌ها، جابه‌جایی‌ها به‌ویژه در مناطق ساحلی و گردشگری افزایش یافته و الگوی بازگشت تدریجی به رفتار عادی سفر در داده‌ها آشکار گردیده است. در مطالعه‌ای دیگر در ایتالیا، از داده‌های تماس تلفن همراه برای سنجش میزان جابه‌جایی جمعیت پس از زمین لرزه‌های مرکز ایتالیا استفاده شد [۸]. نتایج نشان داد که تغییرات تعداد تلفن‌های فعال در مناطق آسیب‌دیده با داده‌های رسمی سرشماری همبستگی بالایی دارد و می‌تواند به صورت کوتاه‌مدت جابه‌جایی واقعی جمعیت را پیش‌بینی کند. این پژوهش رابطه‌ی معنی‌داری میان شدت خسارت، فاصله از مرکز زلزله و میزان کاهش حضور جمعیت نشان داد. مقایسه این یافته‌ها با نتایج پژوهش حاضر نشان می‌دهد که در هر سه مطالعه، داده‌های تلفن همراه الگوهای معناداری از تغییرات رفتاری و فضایی جمعیت در شرایط بحرانی منعکس می‌کنند. در این پژوهش، همانند تجربه‌ی اسپانیا در دوران همه‌گیری و ایتالیا در بحران زلزله، کاهش محسوسی در میزان جابه‌جایی و تنوع مکانی مشاهده شده است. این شباهت‌ها بیانگر آن است که داده‌های موقعیت‌یابی تلفن همراه نه تنها در شرایط خاص یک کشور، بلکه در مقیاسی جهانی به‌عنوان شاخصی قابل اعتماد برای سنجش تحرک و پویایی جمعیت عمل می‌کنند. لازم به ذکر است در تمامی نمودارهای ارایه شده در ادامه، آزمون کولموگروف–اسمیرنوف که یک آزمون ناپارامتری آماری بر پایه تابع توزیع تجمعی تجربی (ECDF) [۶۲] است، برای مقایسه دو توزیع داده در شرایط عادی و بحران استفاده شده است که با مقدار $p$ –مقدار بسیار نزدیک صفر، فرض برابری توزیع دو داده رد شده است.




شکل ۶: نمودار بافت‌نگار تعداد مکان‌های منحصربه‌فرد ثبت‌شده برای هر کاربر، در طی ۷۵ روز در شرایط بحران




شکل ۷: نمودار بافت‌نگار تعداد مکان‌های منحصربه‌فرد ثبت‌شده برای هر کاربر، در طی ۷۵ روز در شرایط عادی

## تحرک کاربران

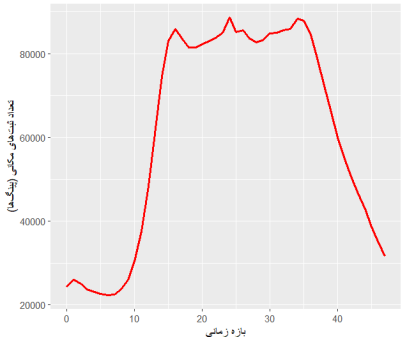
در شکل‌های ۶ و ۷، بافت‌نگار تعداد مکان‌های منحصربه‌فرد ثبت‌شده برای هر کاربر طی ۷۵ روز، در دو شرایط عادی و بحران مقایسه شده است. در این نمودار برای هر کاربر، تمام نقاط منحصربه‌فردی که در آن‌ها حضور داشته بر اساس طول و عرض جغرافیایی شمرده می‌شود. نمودار ۷ مربوط به شرایط عادی، نشان‌دهنده توزیع چوله به راست است، به‌طوری‌که بیشترین تراکم کاربران در بازه ۵۰ تا ۲۰۰ مکان منحصربه‌فرد ثبت‌شده قرار دارد و مُد داده‌ها مربوط به حدود ۱۵۰ مکان ثبت‌شده برای کاربران است. در این حالت، دامنه جابجایی کاربران گسترده‌تر بوده و مواردی با بیش از ۷۵۰ مکان منحصربه‌فرد نیز ثبت شده‌اند. این موضوع بیانگر تحرک بالای کاربران در شرایط معمولی و نبود محدودیت خاصی بر رفتار جابجایی آن‌هاست. در مقابل، نمودار مربوط به شرایط بحران (نمودار ۶)دارای توزیع چوله به راست اما با نقطه بیشینه‌ای کشیده‌تر و دامنه‌ای که گرچه مشابه شرایط عادی است، تراکم داده‌ها در آن بسیار

[61] The Spanish Institute of Statistics
[62] Empirical Cumulative Distribution Function

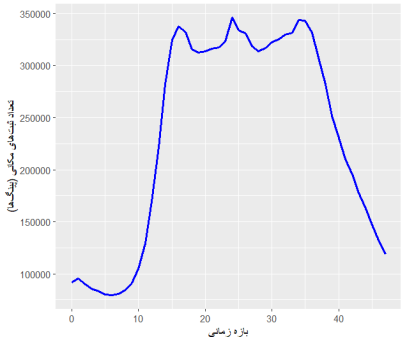
پایین‌تر بوده و کاربران در مجموع تحرک کمتری داشته‌اند. بیشترین فراوانی کاربران در این نمودار در حدود ۱۷۰ مکان منحصربه‌فرد ثبت‌شده است، اما ارتفاع این نقطه بیشینه به مراتب کمتر از شرایط عادی است. این کاهش فراوانی در تمام سطوح دامنه نشان‌دهنده کاهش کلی فعالیت مکانی کاربران در دوره بحران است. در مجموع، مقایسه این دو توزیع نشان می‌دهد که بحران‌ها تأثیر قابل‌توجهی بر رفتار حرکتی کاربران داشته‌اند. بخش کوچکی از کاربران، دفعات زیادی مشاهده شده‌اند و تحرک بالایی (۵۰۰ مکان منحصربه‌فرد) از خود نشان داده‌اند؛ این درحالی است که در وضعیت بحران، تعداد کاربران با تحرک‌های مختلف، از وضعیت عادی بسیار کمتر هستند.

## تغییرات زمانی تعداد ثبت‌ها یا سیگنال‌های مکانی در طول زمان

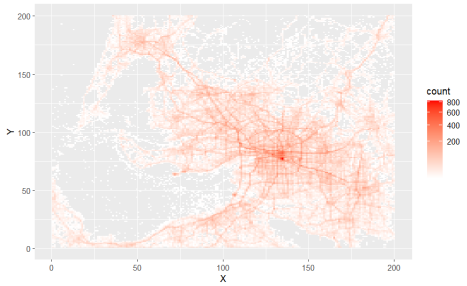
هر بار که یک گوشی تلفن همراه با دکل مخابراتی، آنتن BTS یا شبکه ارتباط برقرار می‌کند (مثل برقراری تماس، ارسال پیام، استفاده از اینترنت یا حتی تغییر مکان)، یک ثبت موقعیت یا سیگنال‌های مکانی در سیستم اپراتور ذخیره می‌شود. به این ثبت‌های لحظه‌ای پینگ می‌گویند. در شکل ۹، میزان ثبت‌های مکانی یا سیگنال‌های مکانی در طول زمان برای شرایط عادی که دارای نوسانات طبیعی است و در برخی نقاط به اوج‌های بالایی می‌رسد، نشان داده شده است. این نشان‌دهنده یک شهر پویا با دسترسی آزاد به حمل‌ونقل و تحرک گسترده کاربران است. در مقابل، شکل ۸، میزان ثبت‌های مکانی در شرایط بحران را نمایش می‌دهد. در نگاه کلی، روند کلی نمودار قرمز مشابه نمودار آبی است، اما در نمودار قرمز، برخی بخش‌ها دارای تغییرات یکنواخت‌تری هستند. این یافته‌ها نشان می‌دهند که بحران منجر به افت قابل توجهی در تحرک کاربران شده است. در نمودار قرمز، نقاط اوج و سایر نقاط نسبت به نمودار آبی پایین‌تر هستند، که نشان می‌دهد در شرایط بحرانی، تعداد کاربران فعال در بازه‌های زمانی کاهش یافته است. بیش‌ترین ثبت‌ها یا سیگنال‌های مکانی در شرایط بحران حدود ۹۰ هزار و در حالت عادی حدود ۳۵۰ هزار ثبت بوده است.



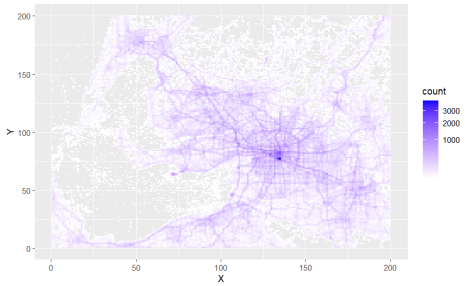

شکل ۸: تغییرات زمانی تعداد ثبت‌های مکانی در هر بازهٔ نیم‌ساعته (از بازهٔ ۰ تا ۴۷، معادل ۲۴ ساعت) در شرایط بحرانی




شکل ۹: تغییرات زمانی تعداد ثبت‌های مکانی در هر بازهٔ نیم‌ساعته (از بازهٔ ۰ تا ۴۷، معادل ۲۴ ساعت) در شرایط عادی




شکل ۱۰: بافت‌نگار دوبُعدی تعداد کاربران منحصربه‌فرد مشاهده‌شده در طول ۷۵ روز در شرایط بحران

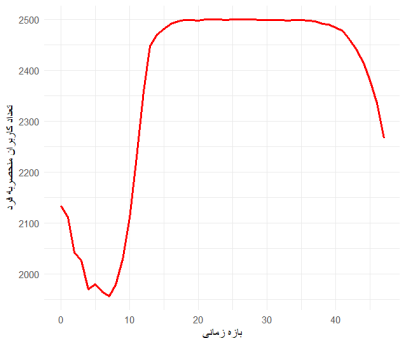



شکل ۱۱: بافت‌نگار دوبُعدی تعداد کاربران منحصربه‌فرد مشاهده‌شده در طول ۷۵ روز در شرایط عادی

## تراکم کاربران

نمودارهای ۱۰ و ۱۱ به ترتیب، بافت نگار دوبُعدی تعداد کاربران منحصربه‌فرد مشاهده‌شده در طول ۷۵ روز در شرایط بحران و شرایط عادی را نشان می‌دهد. در شرایط عادی، بیشترین تعداد کاربران حدود ۴۰۰۰ نفر بوده است، در حالی که این رقم در شرایط بحران به حدود ۸۰۰ نفر کاهش یافته است. نقاط پرتراکم‌تر با رنگ‌های تیره‌تر مشخص شده‌اند که نشان‌دهنده‌ی سطح بالای تعامل و فعالیت کاربران است. در شرایط بحران، تعداد کاربران به‌طور قابل‌توجهی کاهش یافته است. رنگ‌های قرمز نشان‌دهنده‌ی نقاط پرتراکم‌تر هستند. با وجود کاهش تراکم، نقاط تجمع کاربران در شرایط بحرانی تقریباً مشابه شرایط عادی است، اما تعداد کاربران در هر نقطه کاهش یافته است. این موضوع نشان می‌دهد که در زمان بحران، الگوی پراکندگی کاربران تغییر زیادی نمی‌کند، بلکه میزان فعالیت و حضور آن‌ها کمتر می‌شود.

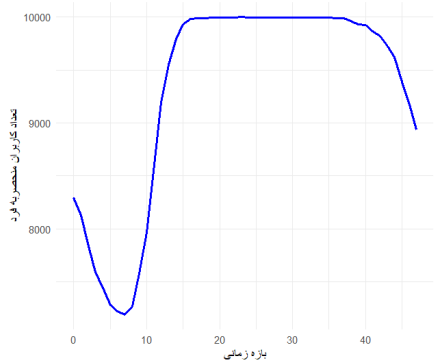
وجود نقاط جدید در شرایط بحرانی نشان می‌دهد که برخی کاربران در موقعیت‌هایی ظاهر شده‌اند که قبلاً حضور نداشته‌اند. این می‌تواند دلایل مختلفی داشته باشد؛ مثلا شاید شرایط بحرانی باعث شده برخی افراد به مکان‌های جدیدی منتقل شوند، یا نوع تعامل آن‌ها تغییر کرده باشد، به‌طوری که در این نقاط خاص فعالیت بیشتری نشان داده‌اند. همچنین ممکن است در شرایط بحران، نیازهای خاصی مثلاً برای دسترسی به منابع، ارتباطات اضطراری یا تغییر رفتار در پاسخ به بحران به وجود آمده باشد که باعث شده کاربران به مکان‌های جدیدی بروند.




شکل ۱۲: تغییرات زمانی تعداد کاربران منحصربه‌فرد در هر بازهٔ نیم‌ساعته (از بازهٔ ۰ تا ۴۷، معادل ۲۴ ساعت) در شرایط بحران

## تغییرات زمانی تعداد کاربران منحصربه‌فرد در طول زمان

نمودارهای ۱۲ و ۱۳ به‌ترتیب تغییرات زمانی تعداد کاربران منحصربه‌فرد در هر بازهٔ نیم‌ساعته (از بازهٔ ۰ تا ۴۷، معادل ۲۴ ساعت) در شرایط بحران و شرایط عادی را نشان می‌دهد. هر دو نمودار نشان‌دهنده یک روند مشابه هستند که ابتدا با کاهش تعداد کاربران و سپس افزایش تدریجی در میانه روز، و در نهایت افت مجدد روبرو است. اما یک تفاوت مهم در دو نمودار وجود دارد. در شرایط بحرانی کاهش اولیه هموار و یکنواخت نیست، بلکه نوساناتی در همان ابتدای روز دیده می‌شود. این نوسانات می‌تواند نشان‌دهنده رفتار متفاوت کاربران در شرایط بحرانی باشد. در مقابل، نمودار شرایط عادی روند کاهشی یکنواخت‌تری دارد، که بیانگر یک الگوی رفتاری پایدارتر است. این الگوی کلی نشان‌دهنده یک چرخه رفتاری است که معمولاً در فعالیت‌های کاربران در طول روز دیده می‌شود. در ساعات ابتدایی روز، به‌دلیل غیرفعال بودن کاربران، تعداد کاربران کاهش می‌یابد. معمولاً در این ساعات، فعالیت کمتر است، زیرا بسیاری از افراد هنوز در حال آماده‌سازی برای روز کاری یا استراحت هستند. با نزدیک شدن به میانه روز، افزایش تعداد کاربران نشان‌دهنده اوج تعامل است. این زمان می‌تواند مربوط به ساعات کاری یا بازه‌هایی باشد که افراد بیشترین دسترسی را به خدمات دارند. معمولاً در این بخش، فعالیت و مشارکت به بیشترین مقدار خود می‌رسد، زیرا کاربران در حال استفاده از سرویس‌ها یا برقراری ارتباط هستند. در پایان روز، نمودار دوباره کاهش می‌یابد که می‌تواند ناشی از کاهش تعامل کاربران، ورود به ساعات پایانی روز، یا کاهش دسترسی باشد.



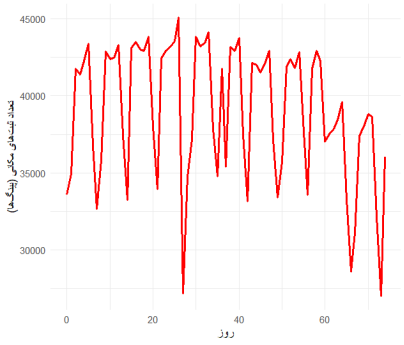

شکل ۱۳: تغییرات زمانی تعداد کاربران منحصربه‌فرد در هر بازهٔ نیم‌ساعته (از بازهٔ ۰ تا ۴۷، معادل ۲۴ ساعت) در شرایط عادی

## تغییرات زمانی تعداد ثبت‌های مکانی (پینگ‌ها) GPS در طول روز

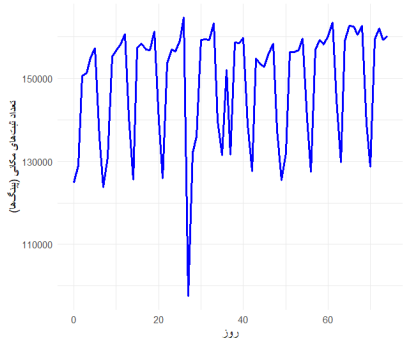
نمودارهای ۱۴ و ۱۵ به‌ترتیب تغییرات زمانی تعداد ثبت‌های مکانی در هر بازهٔ نیم‌ساعته (از بازهٔ ۰ تا ۴۷، معادل ۲۴ ساعت) در شرایط بحران و شرایط عادی را نشان می‌دهد. در شرایط عادی، تعداد ثبت‌های مکانی بین ۱۱۰۰۰۰ تا ۱۵۰۰۰۰ متغیر است و روندی نسبتاً پایدار دارد، در حالی که در شرایط بحران، این تعداد به شدت کاهش یافته و بین ۳۰۰۰۰ تا ۴۵۰۰۰ است. این تفاوت چشمگیر احتمالاً ناشی از عوامل مختلفی مانند محدودیت‌های ارتباطی، کاهش فعالیت کاربران، یا اختلالات زیرساختی در بحران است. روز ۶۰ آغاز شرایط بحرانی است و پس از آن شاهد کاهش چشمگیر تعداد ثبت‌های مکانی هستیم. این تغییر نشان می‌دهد که بحران از این روز به بعد تأثیر قابل‌توجهی بر ارتباطات گذاشته است. روند نزولی پس از روز ۶۰ می‌تواند ناشی از عواملی مانند تشدید اختلالات زیرساختی، محدودیت‌های اعمال‌شده، یا تغییر رفتار کاربران باشد. در مقابل، در روزهای قبل از ۶۰، نمودار شرایط عادی را نشان می‌دهد که پایدار و بدون افت قابل‌توجه است. این موضوع تأکید می‌کند که بحران تأثیر تدریجی و عمیقی داشته که از یک نقطه مشخص شروع شده و سپس روند نزولی تشدید شده است.

## تغییرات تعداد کاربران منحصربه فرد در طول روز

نمودارهای ۱۶ و ۱۷ به‌ترتیب تغییرات زمانی تعداد کاربران منحصربه‌فرد در هر روز (از روز ۰ تا ۷۴) در شرایط بحران و شرایط عادی را نشان می‌دهد. در شرایط عادی، داده‌ها نوسان دارند و تغییرات روزانه نسبتاً پایدار است. روند کلی نشان می‌دهد که تعداد کاربران دارای الگوی مشخصی از نوسان است، اما تغییرات شدید در آن مشاهده نمی‌شود. در شرایط بحرانی، تقریبا تا روز شصتم نمودار نوسان دارد اما پایدار است و مثل شرایط عادی هست اما از روز شصتم به بعد تغییرات شدیدتری مشاهده می‌شود. مثلاً در برخی روزها افت ناگهانی وجود دارد، در حالی که در روزهای دیگر شاهد افزایش هستیم. این الگو نشان‌دهنده تأثیر بحران بر رفتار کاربران است، که ممکن است ناشی از عواملی مانند مشکلات دسترسی، تغییر نیازها، یا افزایش استفاده در شرایط خاص باشد. اما در مجموع از آغاز بحران، روند تعداد کاربران افزایشی است که این می‌تواند به دلیل نیاز مردم برای رفع نیازهای روزمره، ارتباطات، و دسترسی به اطلاعات به فضای آنلاین وابسته شدند.




شکل ۱۴: تغییرات زمانی تعداد ثبت‌های مکانی در هر روز (از روز ۰ تا ۷۴) در شرایط بحران




شکل ۱۵: تغییرات زمانی تعداد ثبت‌های مکانی در هر روز (از روز ۰ تا ۷۴) در شرایط عادی

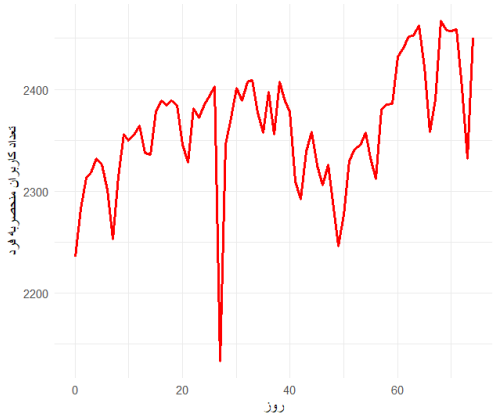
## تحلیل داده‌های زمانی

نمودارهای ۱۸ و ۱۹ به‌ترتیب تغییرات زمانی میانگین جابجایی کاربران (میانگین مسافت طی‌شده توسط هر کاربر در آن روز با استفاده از فاصله اقلیدسی) در هر روز (از روز ۰ تا ۷۴) در شرایط بحران و شرایط عادی را نشان می‌دهد. میانگین جابجایی روزانه کاربران (ADUD)[63] در روز $d$ ام به صورت

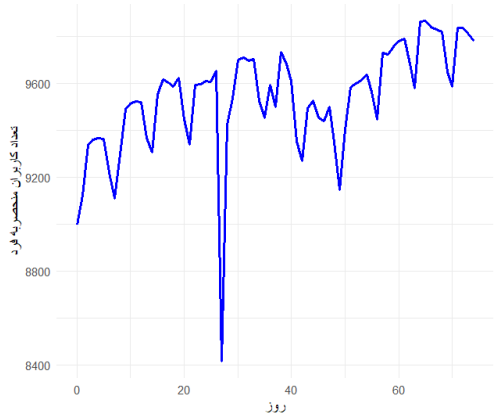
$$ADUD_d = \sum_{i=1}^{n} \sum_{t=1}^{47} \frac{\sqrt{(y_{i,t-1} - y_{i,t})^2 + (x_{i,t-1} - x_{i,t})^2}}{N_d}$$

محاسبه می‌شود که در آن $x_{i,t}$ و $y_{i,t}$ به ترتیب طول و جغرافیایی کاربر $i$ در زمان $t$، $N_d$ تعداد کل جابه‌جایی‌های ثبت‌شده در روز $d$ است [۱۳]. نوسان واضح و منظمی در هر دو نمودار دیده می‌شود، این نوسان

[63] Average Daily User Displacement

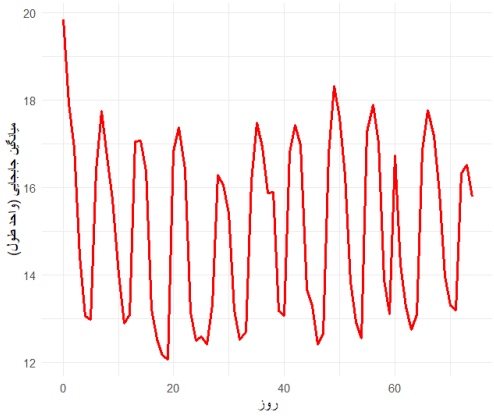



شکل ۱۶: تغییرات زمانی تعداد کاربران منحصربه‌فرد در هر روز (از روز ۰ تا ۷۴) در شرایط بحران

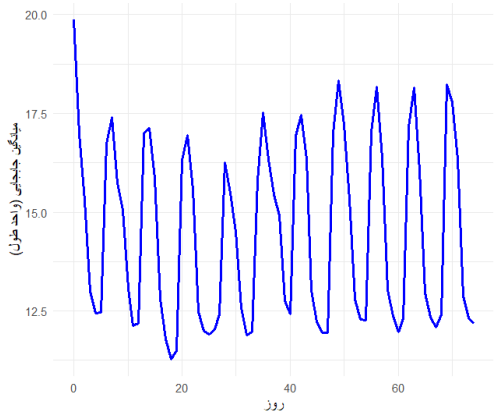



شکل ۱۷: تغییرات زمانی تعداد کاربران منحصربه‌فرد در هر روز (از روز ۰ تا ۷۴) در شرایط عادی

تقریباً هفتگی (۷ روزه) به نظر می‌رسد، یعنی کاربران هر هفته الگوی رفتاری مشخصی دارند. در شرایط عادی، نوسانات هفتگی تقریباً منظم و متقارن هستند؛ یعنی کاربران در طول هفته روال مشخصی داشتند، و تغییرات شدید در رفتار آن‌ها مشاهده نمی‌شود. الگوی تحرک کاربران تقریبا ثابت است و کاربران مطابق با نیازهای معمول خود رفتار می‌کنند. در شرایط بحرانی، بعضی هفته‌ها اوج و فرود شدیدتر است که نشان‌دهنده تغییرات بیشتر تحرک کاربران در شرایط بحران است. پس بحران باعث افزایش نوسانات و کاهش ثبات تحرک کاربران شده است. با آغاز بحران کرونا، جابجایی کاربران کمتر پایدار و با نوسانات هفتگی نامنظم‌تر بود. این نوسانات هفتگی نامنظم‌تر ناشی از عدم پیش‌بینی‌پذیری و محدودیت‌های حرکت در دوران بحران کرونا است. ولی همچنان یک الگوی هفتگی قابل شناسایی وجود دارد، اما نسبت به شرایط عادی کمتر منظم و متقارن است. این نشان می‌دهد که عادات رفتاری هفتگی کاربران دستخوش تغییر شده است. در شرایط عادی تحرک کاربران پایدار (انحراف معیار برابر ۱۰۹٫۱۱)، قابل پیش‌بینی و منظم است ولی در شرایط بحرانی تحرک کاربران بی‌ثبات‌تر (انحراف معیار برابر ۱۲۷٫۰۳)، پراکنده‌تر و واکنشی است، همچنین تغییرات هفتگی نامنظم‌تر است اما همچنان الگوی کلی تحرک کاربران تا حدی حفظ می‌شود.



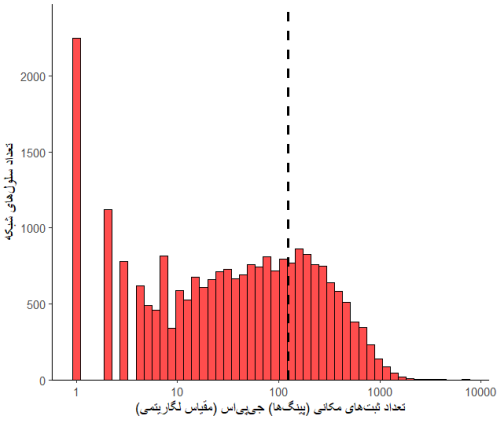

شکل ۱۸: تغییرات زمانی میانگین جابجایی کاربران (میانگین مسافت طی‌شده توسط هر کاربر در آن روز با استفاده از فاصله اقلیدسی) در طول دوره ۷۵ روزه، در شرایط بحرانی



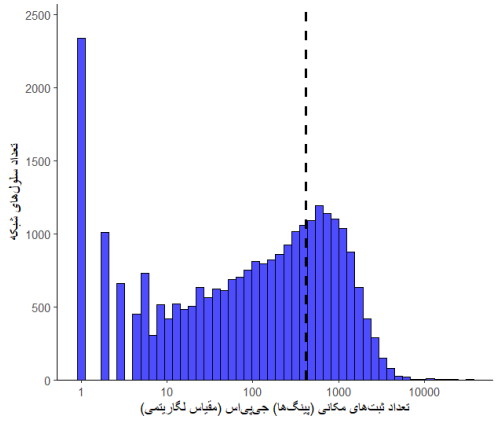

شکل ۱۹: تغییرات زمانی میانگین جابجایی کاربران (میانگین مسافت طی‌شده توسط هر کاربر در آن روز با استفاده از فاصله اقلیدسی) در طول دوره ۷۵ روزه، در شرایط عادی

## تعداد ثبت‌های مکانی (پینگ‌ها) داده‌های موقعیت‌یاب GPS

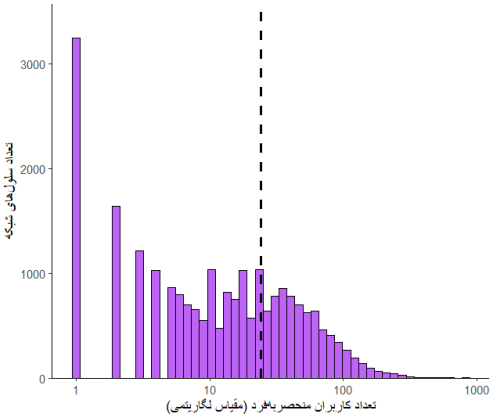
نمودارهای ۲۰ و ۲۱ بافت‌نگار تعداد ثبت‌های مکانی داده‌های موقعیت‌یاب GPS به هر سلول شبکه (تقسیم‌بندی فضای جغرافیایی به سلول‌های مربعی یا چندضلعی منظم )، در طول دوره‌ی ۷۵ روزه موجود در شرایط بحرانی و شرایط عادی را نشان می‌دهد. محور افقی در هر دو نمودار به صورت لگاریتمی مقیاس‌بندی شده است. هر دو نمودار یک توزیع دومُدی را نشان می‌دهند، به‌طوری‌که بخش بزرگی از سلول‌ها (حدود ۲۲۰۰ سلول) تنها یک بار پینگ و یا حدود ۱۰۰۰ سلول تنها ۳ ثبت مکانی (پینگ) بازدید شده‌اند. این در حالی است که در یک مُد دیگر در حدود ۱۰۰۰ ثبت مکانی تعداد بسیار کمی سلول را مشاهده کرده‌اند. همانطور که در نمودار مشخص است، الگوی بافت‌نگار تعداد ثبت‌های مکانی در شرایط عادی و بجرانی متفاوت است.



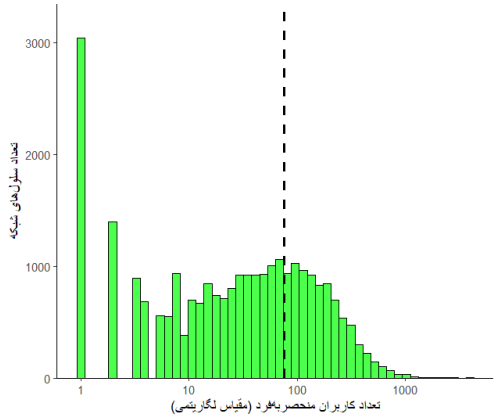

شکل ۲۰: بافت‌نگار تعداد ثبت‌های مکانی داده‌های موقعیت‌یاب GPS به هر سلول شبکه (تقسیم‌بندی فضای جغرافیایی به سلول‌های مربعی یا چندضلعی منظم)، در طول دوره‌ی ۷۵ روزه موجود در شرایط بحرانی




شکل ۲۱: بافت‌نگار تعداد ثبت‌های مکانی داده‌های موقعیت‌یاب GPS به هر سلول شبکه (تقسیم‌بندی فضای جغرافیایی به سلول‌های مربعی یا چندضلعی منظم)، در طول دوره‌ی ۷۵ روزه موجود در شرایط عادی




شکل ۲۲: بافت‌نگار تعداد کاربران منحصربه‌فردی که در هر سلول شبکه بازدید کرده‌اند (تقسیم‌بندی فضای جغرافیایی به سلول‌های مربعی یا چندضلعی منظم)، در طول دوره‌ی ۷۵ روزه موجود در شرایط بحرانی




شکل ۲۳: بافت‌نگار تعداد کاربران منحصربه‌فردی که در هر سلول شبکه بازدید کرده‌اند (تقسیم‌بندی فضای جغرافیایی به سلول‌های مربعی یا چندضلعی منظم)، در طول دوره‌ی ۷۵ روزه موجود در شرایط عادی

## تعداد کاربران منحصربه‌فردی که در هر سلول شبکه بازدید کرده‌اند

نمودارهای ۲۲ و ۲۳ بافت‌نگار تعداد کاربران منحصربه‌فردی که هر سلول شبکه (تقسیم‌بندی فضای جغرافیایی به سلول‌های مربعی یا چندضلعی منظم) را در طول دوره‌ی ۷۵ روزه موجود در شرایط بحرانی و شرایط عادی، بازدید کرده‌اند را نشان می‌دهد. توجه داشته باشید که محور افقی در هر دو نمودار به صورت لگاریتمی مقیاس‌بندی شده است. هر دو نمودار یک توزیع دومُدی را نشان می‌دهند، به‌طوری‌که بخش بزرگی از سلول‌ها (حدود ۳۰۰۰ سلول) تنها یک بار (یک کاربر منحصربه‌فرد) و یا حدود ۱۵۰۰ سلول توسط ۳ کاربر بازدید شده‌اند. این در حالی است که در یک مُد دیگر در حدود ۱۰۰۰ کاربر منحصربه‌فرد تعداد بسیار کمی سلول را مشاهده کرده‌اند. در شرایط عادی، نمودار نشان‌دهنده تمرکز بالای کاربران در تعداد کمی سلول بوده است، در حالی که تعداد کمی از کاربران در بخش بزرگی از سلول‌ها دیده شده‌اند. این الگو معمولاً در سیستم‌هایی دیده می‌شود که یک بخش عمده از کاربران به‌طور محدود با پلتفرم تعامل دارند، در حالی که تنها تعداد کمی به‌شدت فعال‌اند. در شرایط بحرانی نیز نمودار دومُدی است، اگر چه شکل آن با حالت عادی کمی متفاوت است.

## ۶ نتیجه‌گیری

در دنیای امروز که داده‌های دیجیتال به‌طور فزاینده‌ای در فرآیندهای آماری وارد می‌شوند، داده‌های موقعیت‌یابی تلفن همراه به‌عنوان یک منبع نوین و حیاتی در تولید آمار دقیق و به‌روز در حوزه‌های مختلف مانند تحلیل تحرک جمعیت، پیش‌بینی مهاجرت‌ها، برنامه‌ریزی شهری و نظارت بر بحران‌ها مطرح شده است.

تحلیل داده‌های K۱۰۰YJMob به‌روشنی نشان می‌دهند که داده‌های موقعیت‌یابی تلفن همراه، ابزاری کارآمد برای فهم رفتارهای افراد در زمان و مکان‌های مختلف است. در شرایط بحرانی، مانند بلایای طبیعی یا همه‌گیری بیماری‌ها، تحلیل داده‌ها نشان‌دهنده کاهش معنادار جابه‌جایی‌ها، تنوع مکانی و ثبات زمانی است؛ در حالی‌که در شرایط عادی، رفتار کاربران الگوهای منظم‌تری از فعالیت، تردد، اشتغال، و استفاده از زیرساخت‌های شهری را منعکس می‌کند.

یکی از مهم‌ترین یافته‌های این پژوهش، قابلیت بالای داده‌های مکانی در تجزیه‌وتحلیل‌های آماری است. این داده‌ها می‌توانند اطلاعات دقیق‌تری در خصوص الگوهای رفتاری و جغرافیایی افراد ارائه دهند که در روش‌های سنتی گردآوری داده‌ها کمتر به‌دست می‌آید. با این وجود، استفاده از این داده‌ها نیازمند توجه ویژه به چندین محدودیت و چالش است. اولاً، نگرانی‌های حقوقی و اخلاقی مرتبط با حریم خصوصی افراد و حفاظت از داده‌ها باید در اولویت قرار گیرد. دولت‌ها و نهادهای آماری باید چارچوب‌های قانونی و اخلاقی مناسبی برای گردآوری، پردازش و اشتراک‌گذاری این داده‌ها تدوین کنند تا از استفاده غیرمجاز و نقض حریم خصوصی جلوگیری شود. در این راستا، مثال‌هایی از کشورهای مختلف نشان داده‌اند که استانداردسازی‌ها و چارچوب‌های قانونی می‌توانند به عنوان یک راهکار عملی در جلوگیری از بروز مشکلات حقوقی عمل کنند. ثانیاً، تحلیل داده‌های تلفن همراه نیازمند زیرساخت‌های مناسب فنی و منابع انسانی متخصص است. در این پژوهش به وضوح نشان داده شد که برخی کشورهای پیشرفته با همکاری‌های بین‌المللی و طراحی سیستم‌های تحلیل داده‌های پیشرفته توانسته‌اند از پتانسیل این داده‌ها به‌طور مؤثر بهره‌برداری کنند. در مقابل، برخی کشورها به دلیل عدم دسترسی به فناوری‌های مناسب و یا کمبود تخصص در این زمینه، با چالش‌هایی در استفاده از داده‌های تلفن همراه روبه‌رو بوده‌اند. از محدودیت‌های این پژوهش، می‌توان به محدودیت دسترسی به داده‌های تلفن همراه و اندازه نمونه کوچک برای تحلیل اشاره کرد. پیشنهادات این پژوهش شامل تقویت همکاری‌های بین‌المللی در زمینه تبادل دانش و استانداردهای بین‌المللی برای استفاده از داده‌های تلفن همراه، توسعه ظرفیت‌های فنی و آموزشی در کشورهای در حال توسعه و ایجاد چارچوب‌های قانونی و اخلاقی شفاف است. علاوه بر این، با توجه به گسترش تلفن‌های همراه هوشمند، اینترنت تلفن همراه و سامانه تعیین موقعیت جهانی، پژوهشگران و متخصصان آمار باید به طور مستمر از روش‌های آماری ساده به روش‌های پیچیده یادگیری ماشین و یادگیری عمیق برای تحلیل و پردازش داده‌های تلفن همراه و به‌ویژه مه‌داده‌ها پیش بروند و تحقیقات آینده در کشور در این مسیر ادامه یابد تا از تمامی پتانسیل‌های موجود به‌طور بهینه استفاده کنند.

در نهایت، داده‌های موقعیت‌یابی تلفن همراه به‌عنوان یک منبع نوین می‌توانند تحولی در تولید آمار رسمی به‌وجود آورد و تحلیل‌های دقیق و به‌روز از رفتار جمعیت و الگوهای تحرک شهری ارائه کنند که در سیاست‌گذاری شهری، حمل‌ونقل و مدیریت بحران کاربرد مستقیم دارد. همچنین ترکیب این داده‌ها با روش‌های آماری پیشرفته و اعتبارسنجی با داده‌های سنتی، امکان تولید نتایج کمّی قابل اعتماد و جامع را برای پژوهشگران و نهادهای آماری فراهم می‌کند. برای بهره‌برداری مؤثر و مسئولانه، توسعه چارچوب‌های قانونی و استانداردهای بین‌المللی،

ظرفیت‌سازی فنی و همکاری‌های علمی ضروری است، تا از تمامی پتانسیل‌های داده‌های مکانی در آمار رسمی استفاده شود.

اجرای این پیشنهادات مستلزم اقدامات زیر است:

۱. **ایجاد زیرساخت‌های فنی و محاسباتی پیشرفته** برای ذخیره‌سازی، پردازش و تحلیل حجم عظیم داده‌های موقعیت‌یابی با رعایت الزامات امنیت و حریم خصوصی.

۲. **برگزاری کارگاه‌های آموزشی و برنامه‌های مهارت‌افزایی** برای متخصصان آمار، علوم داده و برنامه‌ریزان شهری با تمرکز بر تحلیل داده‌های مکانی و الگوریتم‌های یادگیری ماشین.

۳. **ایجاد چارچوب‌های همکاری بین نهادهای دولتی، اپراتورهای تلفن همراه و مراکز پژوهشی** برای تسهیل دسترسی ایمن به داده‌ها و تعریف استانداردهای مشترک برای فرمت و سطح تجمیع داده‌ها.

۴. **تدوین دستورالعمل‌های اخلاقی و حقوقی** در زمینه استفاده از داده‌های مکانی، با تأکید بر ناشناس‌سازی، رضایت آگاهانه کاربران و حفظ حریم خصوصی.

۵. **اجرای نمونه‌های آزمایشی** در مقیاس شهری یا استانی برای آزمودن قابلیت کاربردی روش‌های تحلیلی پیشنهادی و ارزیابی اثرات آن‌ها در تصمیم‌گیری‌های سیاستی.

۶. **توسعه مدل‌های ترکیبی داده‌ای** که داده‌های تلفن همراه را با منابع سنتی مانند سرشماری و داده‌های تردد ترکیب می‌کنند تا دقت برآوردها افزایش یابد.

این پژوهش امید دارد که مسیرهای جدیدی را برای استفاده بهینه از داده‌های دیجیتال در آمار رسمی هموار کرده و پژوهش‌های آینده را در این حوزه هدایت کند.

پذیرفته شده

# مراجع


[1] Blondel, V. D., Esch, M., Chan, C., Clérot, F., & Krings, G. (2015). *Data for Development: The D4D Challenge on Mobile Phone Data*.arXiv preprint arXiv:1210.0137.

[2] European Commission (2023). *Mobility Data in Migration Crisis*.

[3] European Parliament and Council. (2016). *General Data Protection Regulation (EU) 2016/679*. Official Journal of the European Union.

[4] European Parliament and Council of the European Union. (2016). *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation)*.

[5] Eurostat (2014). Feasibility Study on the Use of Mobile Positioning Data for Tourism Statistics. Report 3a − Feasibility of use, methodological issues.

[6] Eurostat (2020). *Mobile Phone Data for Official Statistics: A Eurostat Handbook*.

[7] Eurostat Big Data Task Force (2014). *ESS big data action plan and roadmap 1.0 – work programme objective 11.1*.

[8] Giardini, F., Hadjidimitriou, N. S., Mamei, M., Bastardi, G., Codeluppi, N., & Pancotto, F. (2023). Using mobile phone data to map evacuation and displacement: a case study of the central Italy earthquake. Scientific reports, 13(1), 22228.

[9] Instituto Nacional de Estadística (INE), Spain. (2022). *Annual Statistical Report*.

[10] ITU. (2019). *Measuring the Information Society Report*. International Telecommunication Union.

[11] Leng, Y., Noriega, A., & Pentland, A. (2021). *Tourism event analytics with mobile phone data*. ACM/IMS Transactions on Data Science, **2**(3), Article 25.

[12] Lu, X., Bengtsson, L., & Holme, P. (2012). Predictability of population displacement after the 2010 Haiti earthquake. Proceedings of the National Academy of Sciences, 109(29), 11576–11581.

[13] Maria, E., Budiman, E., & Taruk, M. (2020, February). Measure distance locating nearest public facilities using Haversine and Euclidean Methods. In Journal of Physics: Conference Series (1450 (1), 012080). IOP Publishing.

[14] Novak, J., Ahas, R., Aasa, A., & Silm, S. (2013). Application of mobile phone location data in mapping of commuting patterns and functional regionalization: A pilot study of Estonia. Journal of Maps, 9(1), 10-15.

[15] Osorio Arjona, J., & de las Obras-Loscertales Sampériz, J. (2023). Estimation of mobility and population in Spain during different phases of the COVID-19 pandemic from mobile phone data. Scientific Reports, 13(1), 8962.

[16] Ricciato, F., De Meersman, F., Wirthmann, A., Seynaeve, G., & Skaliotis, M. (2018, October). Processing of mobile network operator data for official statistics: the case for public-private partnerships. In 104th DGINS conference.

[17] Steenbruggen, J., Borzacchiello, M. T., Nijkamp, P., & Scholten, H. (2013). Mobile phone data from GSM networks for traffic parameter and urban spatial pattern assessment: a review of applications and opportunities. GeoJournal, 78(2), 223-243.

[18] Suimon, Y. (2021). Economic analysis based on the mobile phone GPS data and monitoring consumer behavior during the COVID-19 pandemic. Information Engineering Express, 7(2), 49–58.

[19] UN Council of Experts in Big Data. *Use of Mobile Phone Data in Transportation* (2023). Working Party on Transport Statistics, Seventy-fourth session United Nations, Geneva, 15-17.

[20] United Nations Global Working Group (GWG) on Big Data for Official Statistics. (2020). *Handbook on Mobile Phone Data for Official Statistics*.

[21] UN Global Pulse. (2018). *Annual Report*.

[22] United Nations High Commissioner for Human Rights. (2021). *The right to privacy in the digital age*. Report A/HRC/48/31.

*Andisheye Amari, 2024, Vol. 29, No. 2, pp. 1-26*
https://doi.org/10.22034/aa.2025.732975

# Familiarity with Mobile Positioning Data in the Production of Official Statistics

**Yasin Mohammad Aghaei**[1], **Zahra Rezaei Ghahroodi**[2] *

[1] *School of Mathematics, Statistics and Computer Science, University of Tehran, Tehran, Iran*

[2] *School of Mathematics, Statistics and Computer Science, University of Tehran, Tehran, Iran*

*Received: 2025/09/26*

*Accepted: 2025/12/22*

## Abstract

The rapid expansion of mobile communication technologies has led to the continuous generation of large volumes of location data by mobile network users. When used appropriately and within legal frameworks, these data can become a valuable source for producing official statistics in areas such as migration, tourism, Displacement in crises and disaster, Transport and traffic monitoring, informal employment, and crisis management.

The present paper is designed to explore the potentials and limitations of employing mobile positioning data and to outline a methodological framework for producing statistics based on these data within the national statistical system. Given the lack of Persian references in this field, this study seeks to provide a practical framework for the use of mobile phone data in the production of official statistics through a comprehensive review of international literature and the analysis of a real dataset. Within this framework, various types of mobile phone data and their applications in statistical production are introduced, and the legal, ethical, and technical requirements for their use are examined. Furthermore, the methodology for extracting and processing mobile positioning data is described, and a case study based on a dataset related to urban mobility patterns is presented. The main contribution of this research lies in documenting practical experiences and transferring technical knowledge to support the localization and application of such data within the national statistical system.

The analysis of the case study data revealed that crises have a significant impact on users' mobility behavior. Under normal conditions, a small subset of users appeared very frequently and exhibited high mobility (around 500 unique locations). However, during crisis situations, this mobility decreased sharply, with far fewer individuals displaying any level of movement (low, medium, or high), and travel activities became concentrated in specific areas.

**Keywords:** Mobile positioning data, Official statistics, Big data, Privacy, Legal framework, Population mobility analysis.

*Corresponding author, z.rezaeigh@ut.ac.ir