UNBigData

MPD Task Team
Transport and Commuting Subgroup

OVERVIEW OF THE MOBILE PHONE DATA(MPD) HANDBOOK

Objectives of the MPD Handbook

UNBigData
MPD Task Team
Transport and Commuting Subgroup

- Provides practical guidance on how to collect and aggregate data from MPD so they can be used for producing transport statistics
- Uses non-domain-specific material developed from the other handbooks where relevant
- Detail number of use cases, which have either succeeded already in producing data to the standards of official statistics or are expected to in the future
- Analyze the approach taken in terms of data access, ensuring privacy maintenance, data quality control, and any relevant comparisons with existing datasets

Sources of Transport Data

UNBigData
MPD Task Team
Transport and Commuting Subgroup

Transport statisticians, planners, and operators rely on several sources of data

- To **understand how people and goods are moved around** and between countries using roads, railways, inland waterways, seas and the air
- To know **how people access jobs, services** and leisure, have access to transport services, average distances travelled by mode and the number of trips taken
- To **understand the quantities of goods shipped** on different modes in different countries

→ **Two Primary Data Sources**

✓ Transport Regulatory Data

Legacy Data Sources

- ✓ Travel Surveys & Census Data
- ✓ Transport Network Information
- ✓ Vehicle Traffic Data
- ✓ Public Transport & Shared Mobility Ticketing Data
- ✓ Safety & Accidents Data
- ✓ Logistics Data
- ✓ Vehicle & Driver Registration Data

Emerging Data Sources

- ✓ Probe Vehicle Data
- ✓ Connected and Autonomous Vehicles Data
- ✓ High-Definition Mapping
- ✓ **Mobile Phone Data**

What is Mobile Phone Data (MPD)?

UNBigData
MPD Task Team
Transport and Commuting Subgroup

Mobile positioning data (MPD) is a form of big data which results from the high data volumes of mobile positioning – tracking the location of mobile phones.

Mobile positioning data can be used for generating population and tourism statistics, for measuring human mobility, creating data-driven solutions in urban planning, establishing a response plan to disasters etc.

Data from Telecom / Mobile Network Operator (MNO)

Cell-tower associations of cellphones collected from telco providers. The advantage of telco data is that the **cell-tower switching happens every time a user moves,** however the accuracy of telco data depends on the density of cell-towers

There are three type of data probing done by telecom operators:

- **Call Detail Records –** This refers to **data produced** whenever the subscriber makes or receives a call, sends or receives a SMS, or accesses data using their phone
- **Passive Signaling Data –** This data **includes logs of every connection of a mobile device** to the network, which is typically more frequent than CDR events
- **Active Signaling Data –** These are **generated when MNOs decide to send** (or "ping") signals to mobile devices

Data from Smartphone Applications

Consisting of data collected by data suppliers by tapping into smartphone applications like **GPS sensors**. The advantage of this is the higher accuracy of location data, however due to privacy concerns, lot of users opt out

There are three types of location data collected by the smartphone applications

- **User Check-in Data –** Applications may **trigger users to check-in at locations** either as part of the social media postings, or as part of incentivizing the users to contribute to the community
- **Foreground Location Data –** Applications may **collect precise or approximate user-location** based on user setting when applications are running
- **Foreground Location Data –** Applications may also **collect data from users in the background**

Digital Survey Data

Consisting of **user responses** to generalized or targeted surveys pushed to applications by app developers. Typically, digital surveys are used as a value-add on top of the first two types of MPD to understand personalized experience

In the context of transport studies, digital survey data can be gathered in several ways

- **Online Surveys –** Respondents fill out **surveys** on their computers or smartphones
- **Mobile App Surveys –** Surveys conducted through dedicated mobile applications.
- **Digital Diaries –** Participants **record their travel activities in a digital format**, typically through a smartphone app, providing detailed, time-stamped data about their trips

Advantages and Drawback for MPD

UNBigData
MPD Task Team
Transport and Commuting Subgroup

MPD can support urban transport planning as a continuous exercise

- Greater spatio-temporal detail than corresponding traditional output
- Negligible incremental cost of generating forecasts
- Single source for understanding different aspects of mobility
- Time taken to collect this data is very less and the data is continuous and not a snapshot
- Data is automatically collected by mobile operators every time a call is made, a text is sent, or internet accessed on phone

Inherent limitations mean MPD cannot replace traditional process entirely

- There is no information regarding the background of the users: their characteristics, socio-economic indicators, purpose of the trip, type of accommodation and means of transport used etc.
- MPD is spatially more precise in densely populated urban areas and near highways due to a larger number of cell towers, but in the countryside, for example, the accuracy is lower
- Limited to the coverage area of network cells, which can range from a few hundred meters to multiple kilometers

Measure of MPD useability

UNBigData
MPD Task Team
Transport and Commuting Subgroup

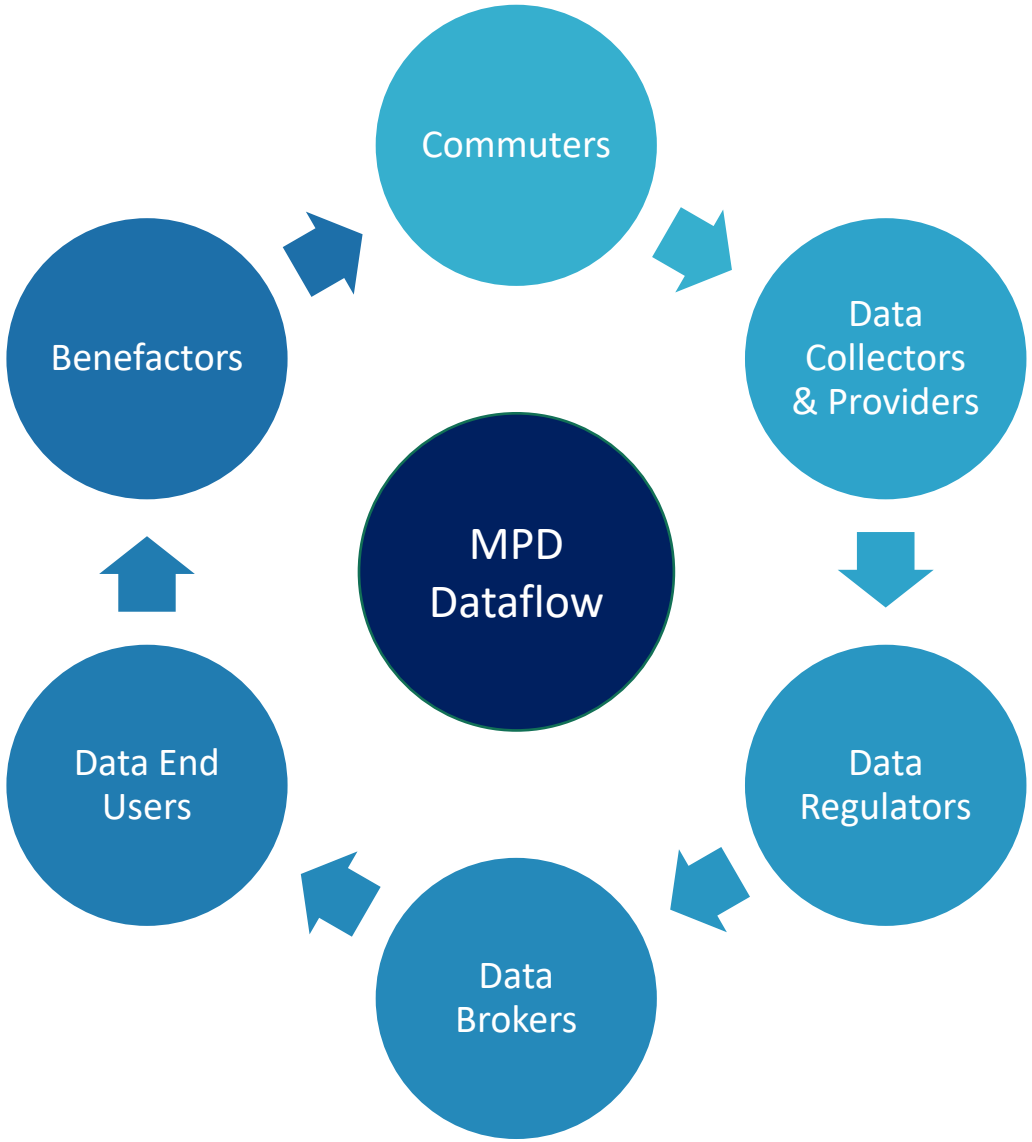
The useability of the Mobile phone data, regardless from which source it is collected, is generally attributed to following 5 factors

1. **Latency**: The **time delay** between collection of mobile phone data and its availability for the end-user
2. **Precision**: The **accuracy of the mobile phone data** in regarding to collecting user's location
3. **Ping Rate**: The rate of location pings by MNO to determine user location
4. **User Penetration**: The **percent of commuters** who actively contribute to the MPD data provisioning pool
5. **Completeness**: This represents the **number of attributes** present in the mobile phone data

	MNO Data			Application Data		
	Call Detail Records	*Passive Signaling Data*	*Active Signaling Data*	*User Check-in Data*	*Foreground Location Data*	*Background Location Data*
Latency						
Precision						
Ping Rate						
User Penetration						
Completeness						

Typical flow of data for MPD

UNBigData
MPD Task Team
Transport and Commuting Subgroup



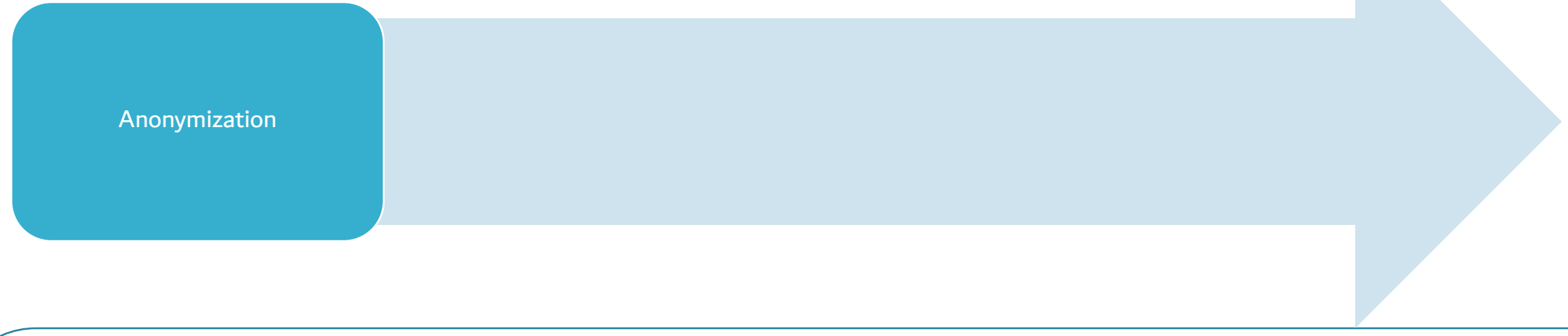
Typically, the dataflow of MPD occurs between different stakeholders listed below (non-exhaustive)

- **Commuters** represents the **commuters whose mobile phones are used as the raw data source.** Their role is to understand what data is collected, by whom and what privacy controls are available at their disposal
- **Data Collectors & Providers** represents the **stakeholder who collect data from the commuters and provide it to the end-users in a clean**, usable way that not only protects their business, but also their commuter privacy
- **Data Regulators**, represents the **governmental or non-governmental organizations that actively or passively regulate** such activity and ensure protection of user rights, data protection laws and citizen privacy
- **Data End-Users**, represents the **users of MPD data**, and could be statistics departments, transport operators, metropolitan planning organizations (MPOs) or private companies who use MPD to target and improve their business
- **Benefactors**, represents the **stakeholder groups** that are benefitted from the end-users' use of MPD data

Processing of MPD

UNBigData

MPD Task Team
Transport and Commuting Subgroup



Understanding the Limitations of MPD

UNBigData

MPD Task Team
Transport and Commuting Subgroup

Applications of Mobile Phone Data in Transportation

UNBigData

MPD Task Team
Transport and Commuting Subgroup

Transportation agencies and statistics commissions across the world are moving towards either solely using high-resolution people movement data from mobile devices in their reporting, planning and operational needs, or at least use it in supplementing other available data (such as travel surveys, ticketing data etc.)

Transport Statistics

These statistics describe the mobility of people and goods across various modes of transport (e.g., cars, trains, planes, etc.).

Mobile phone data can provide information such as
- Origins and destinations of trips
- Modes of transport
- Population density
- Total Personal Miles Traveled (PMT)

Traffic Statistics

Traffic Statistics: Traffic statistics focus on road usage, specifically the movements of vehicles.

By monitoring location data from mobile phones, it's possible to
- Estimate the volume of traffic
- Identify congested areas
- Assess the average speeds
- Vehicle Miles Traveled (VMT)

Transport and Urban Planning

Mobile phone data can contribute to understanding the patterns of movement within a city or region.

This data can be used to
- Identify commuting patterns
- Popular routes
- Origin-Destination Matrix including Modal Split
- Travel Time Index

Transit & Public Transport Planning

Mobile phone data can be used to determine how, when, and where people are using public transportation.

This includes identifying
- Busy stations,
- Peak usage times,
- Popular routes
- Planning Transit and Public Transport

Traffic Analysis

Mobile phone data can be used to conduct detailed analyses of traffic conditions.

This can include
- Identifying patterns of congestion
- Studying the impact of road works accidents
- Analyzing the effects of weather on traffic

Operational Decision Support

In transport operations, mobile phone data can help predict demand and optimize schedules and routes.

For example, it can support decisions on
- Where to deploy additional buses during peak times
- How to reroute traffic in response to an incident

Events and Crisis Management

During events or crises, mobile phone data can provide real-time insights into population movements.

This could be useful for
- Managing crowds at a major event
- Understanding evacuation patterns during a crisis

Traffic & Transit Demand Prediction

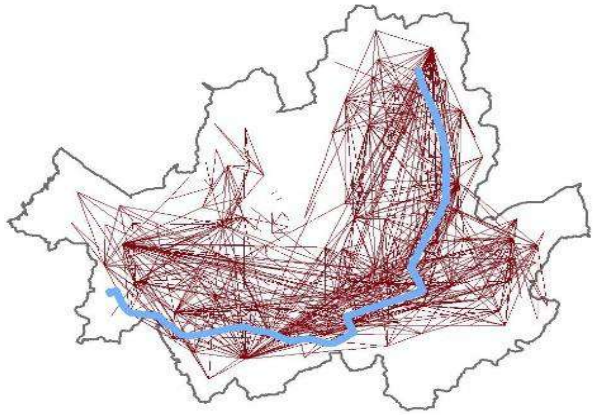
With advanced analytics, mobile phone data can be used to predict future traffic and transit demand.

This could involve
- Forecasting daily traffic volumes
- Predicting the impact of a new transport service
- Modeling the effects of population growth on transport demand

Case Studies

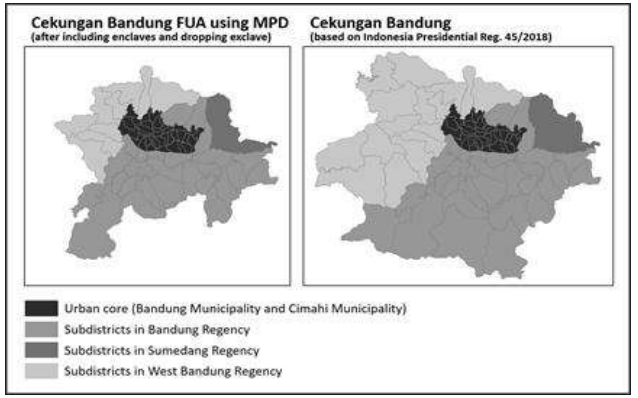
UNBigData

MPD Task Team
Transport and Commuting Subgroup



Seoul's Use of MPD in Planning Night Bus Service

The Data and Statistics division of Seoul's metropolitan government analyzed late night call records provided by the Korean telecom KT Company to plan bus routes in a more inclusive manner.



Replacing Indonesia's Household Travel Surveys

In 2019, BPS-Statistics Indonesia worked with the Indonesian Ministry of National Development Planning to conduct a pilot project which explored the use of MPD to delineate metropolitan areas.

Optimizing Dubai's Bus Laybys based on First & Last Mile

Dubai RTA and Locatium used mobile phone data to optimize bus station layout based on passenger usage patterns.

Case Study: Seoul's Use of MPD in Planning Night Bus Service

UNBigData
MPD Task Team
Transport and Commuting Subgroup

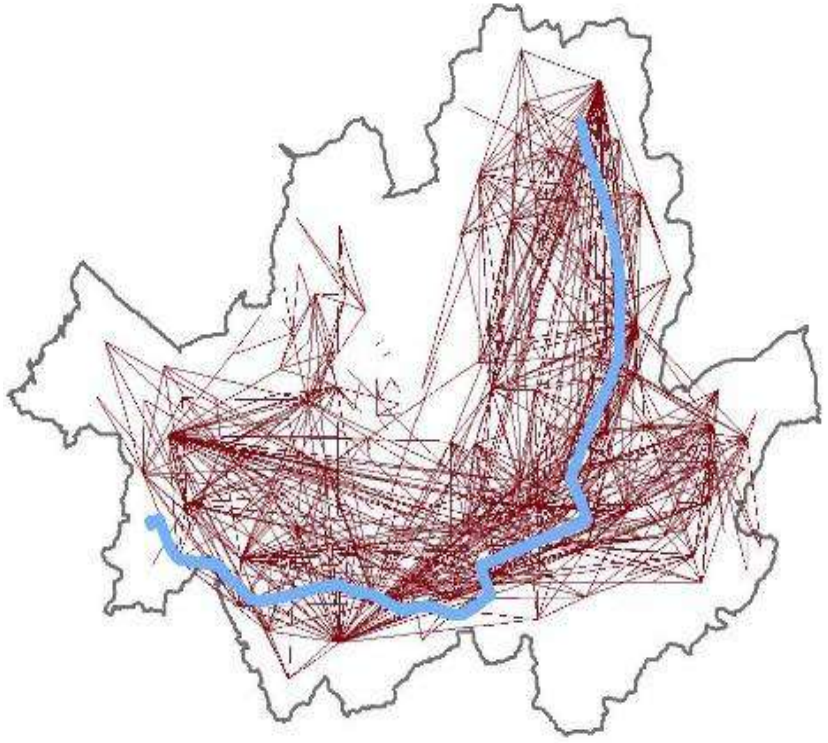


Figure showing travel patterns traced in Seoul for owl bus service

Case Study

The **Data and Statistics division of Seoul's metropolitan government** analyzed late night call records provided by the Korean telecom KT Company to plan bus routes in a more inclusive manner.

Objective

Designing routes for the night bus service proved difficult. Daytime traffic data could be misleading since commuter behavior might be different at night. **Relying on the intuition and guesswork of experts, which was the convention for route design**, would fall short of providing optimal routes and service frequency.

Methodology

The Seoul government reached out to a major telecommunications provider to gain access to a huge set of anonymous mobile communication data and the location of mobile users was used as a proxy for the movement pattern of commuters. **Over 3 billion mobile call logs were gathered** over a period of 1 month to map the distribution of late-night travelers across the city. **Journey data from over 5 million taxi rides**, collected through the T-Money card system, were also used. Seoul was then divided into 1,250 cell units of 1-kilometer radius each. The geographic information system-enabled Night Bus Route Design Support System then **overlaid the data onto the cells to produce a visualized pattern of the late-night floating population**

Impact

The result was **the creation of the "owl bus," which operates late into the night until five o'clock in the morning**. Since 2012, Seoul's "Owl Bus" service has saved the city's lower-income groups approximately $1.2 million by providing more affordable transportation using cell phone data to plan late night bus routes.

Case Study: Replacing Indonesia's Household Travel Surveys

UNBigData

MPD Task Team
Transport and Commuting Subgroup

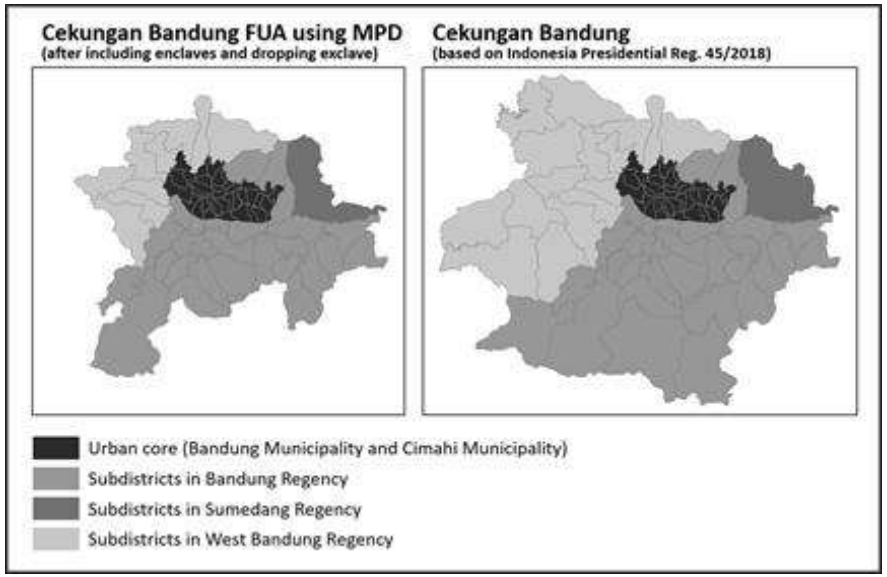


Figure above shows the visualization of MPD delineation and Indonesian Government delineation comparison on Cekungan Bandung metropolitan area

Case Study

In 2019, **BPS-Statistics Indonesia** worked with the Indonesian **Ministry of National Development Planning (BAPPENAS)** to conduct a pilot project which explored the use of MPD to delineate metropolitan areas.

Objective

Cekungan Bandung one of the metropolitan areas that have been delineated using commuting survey. The objective of this case study was **to use MPD to delineate the same metropolitan area** Cekungan Bandung and then compare with the last result

Methodology

The data itself was obtained from Telkomsel, one of the largest MNO in Indonesia, which is a state-owned private enterprise. **There were 50,907 Telkomsel subscribers** used as the sample within November 2019. To analyze commuter patterns using MPD, BPS-Statistics Indonesia **developed algorithms to identify commuters and estimate commuting flows at sub-district level** (the third level of local area unit in Indonesia). The **commuting flows were used to measure the integration** between the urban core and surrounding hinterlands in Cekungan Bandung

Impact

The delineation from **MPD results were then compared to the delineation** already determined by the Government of Indonesia, to identify which sub-districts were included in the metropolitan area by law but have a low rate of commuting flow. The **results was a recommendation for the Government** of Indonesia to help **determine a more appropriate delineation** by using MPD, especially in Cekungan Bandung

Case Study: Optimizing Dubai's Bus Laybys based on First & Last Mile

UNBigData

MPD Task Team
Transport and Commuting Subgroup

Figure shows the GPS pings of users of buses to Dubai's Al Ghubaiba station, along with their final destination.

Case Study

In 2021, **Dubai RTA** and **Locatium** conducted studies using Mobile Phone Data to understand how passengers use bus stations prior and after to their public transport trip to optimize bus station layout.

Objective

Public Transport Agencies have a good understanding of their users travel pattern within their system via smart ticketing data. However, Dubai wanted to understand how this pattern continued before and after the PT system usage with the aim of optimizing bus laybys. Al Ghubaiba station forms a major transit center with all modes of transport converging in, including metro, taxi, marine and bus.

Methodology

For this pilot, Locatium collected anonymized smartphone-based GPS pings from a set of users. For the users whose journeys end or start at the Dubai's Al Ghubaiba Station, the dwell-time data was utilized to understand the origin or final destination of the trips. Based on rigorous analysis, it was found that a simple reconfiguration of laybys can help the users reach their next mode of transport with little walk from the station.

Impact

Without MPD data, such an analysis wouldn't have been possible. Use of MPD data supported in raising the customer satisfaction levels at the station by 7%. The station, now, consistently ranks very high in Google Reviews as well.

THANK YOU

Join us in unleashing the power of Mobile Phone Data

Latifa.Alshehhi@rta.ae
bigdata@un.org