DataSpark
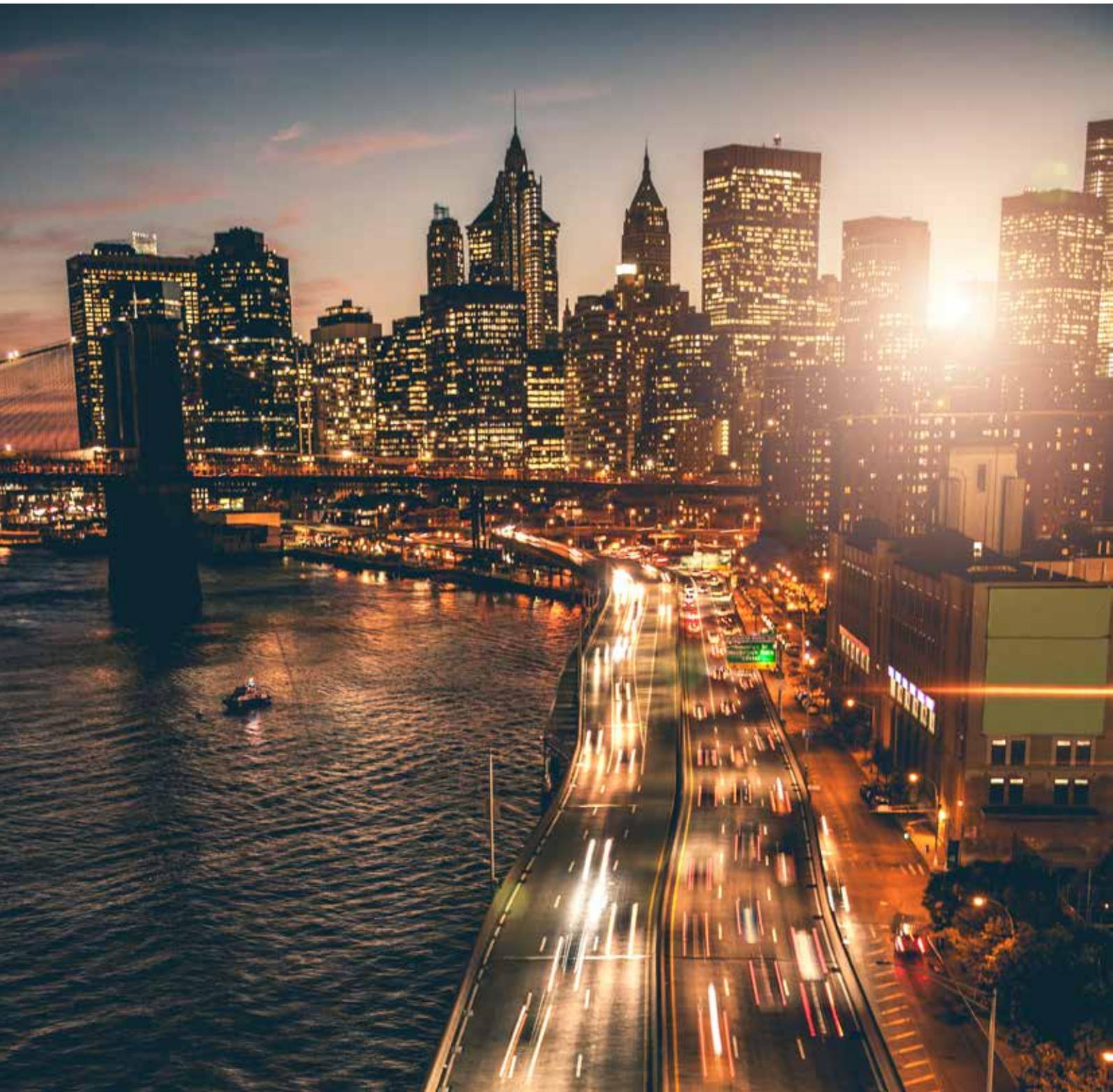
Empowering Transport Planning through Big Data

Introduction

Throughout history, transportation has been the key in enabling the mobility of people and goods, allowing for the transfer of resources, culture, and ideas. The need for robust transportation infrastructure and services is pertinent and particularly acute in urban centres where large populations of people live and work. Fifty-four percent of the world's population live in urban centres and this is expected to grow to sixty-six percent by 2050.[1] Urbanisation increases the need for mobility but also makes planning and delivering an efficient and effective transport system more challenging. Solving these challenges will have a significant impact on the economy, culture, and quality of life.

One of the key ways to plan, develop, and deliver transportation more effectively is by using data about how, when, and why people move. Today, with the ubiquity of mobile devices and the proliferation of the Internet of Things, there is an exponential increase in the amount of mobility data available, allowing transport planners to plan better transportation options. Having better quality data to understand where people live, what time they go to work, and which routes and modes of transport they use is vitally important to transport planners and modellers today who create models and anticipate needs for planning future infrastructure more effectively.

[1] United Nations, Department of Economic and Social Affairs, Population Division (2015). *World Urbanization Prospects: The 2014.*

Challenges

Conventional data collection methods (such as household travel surveys, traffic counters) are designed for yesterday's transportation system, they are expensive to collect, time consuming, and difficult to build. These data sources obtained through the conventional methods have several limitations:

Small Sample Size and Limited Coverage

Conducting surveys manually costs millions of dollars every year, yet they can only cover a fraction (typically less than one percent) of the population. Furthermore, surveys usually cover only a "typical day" for each household, and are not able to account for seasonal or daily variations. The process involves many stages that need very specific technical skills and implementation, a lot of quality control, and analysis to ensure accuracy of the surveys. On the other hand, traffic counters can only be deployed at a limited number of locations; usually the main highways and major roads. Even ticketing data that shows the movement of all public transport users cannot completely cover door-to-door journey, nor private transportation trips.

Low Reliability and Out of Date

The data from surveys are based on participants' recall of their typical day. Self-reported data may contain biases and is difficult to validate. Survey data is also often reported based on a multi-year rolling window, with obsolete data from as long as five years ago included in the final analysis.

Inflexible

Most data sources are published with fixed parameters, such as the dates, time, demographic profiles, and geographical attributes. Once published, the data cannot be queried in more flexible ways. For example, analysing how a major event affects traffic would require data around specific dates, it may not be possible with the existing data sources obtained through conventional methods.

Transport planners and modellers today face a far more complex transportation reality than their predecessors did, and their jobs are more challenging than ever. In addition, with billion-dollar investment decisions made around upgrading and implementing transport infrastructure and operations every year, it is critical to have the right data, at the right time.

We at DataSpark empower transport planners and modellers with big data to make better decisions by providing reliable, timely, useful, and easy-to-query data, using a combination of data from telecommunications networks and state-of-the-art algorithms. With that, they can spend less time collecting data and more time optimizing transportation plans.

Leverage Data for Better and More Timely Decisions

We use mobile network data to generate mobility intelligence of the population. Location signals from Singtel is analysed in an anonymous, aggregated way.

24/7/365 Coverage

Unlike surveys that capture a snapshot of a typical day, telecommunications data is captured 24 hours, 365 days a year. This allows transport planners and modellers to analyse mobility patterns on all days and times in a year, including seasonal trends, special events, and holidays.

Nationwide Coverage

The telecommunications data from Singtel covers more than half of the population, across all geographies nationwide. The data is available in high temporal and spatial resolution that provides new insights into the elements that influence transport planning. No other data set provides transport planners and modellers with a complete view on population movement.

Effortless Data Collection and Next Day Availability

Instead of waiting weeks or months for new data to be published, transport planners using DataSpark data can get the information they need the next day, allowing decisions to be made quicker.

Draw Valuable Insights for Transport Planning

Our proprietary Mobility Genome™ framework provides patented, state-of-the-art algorithms that extract meaningful insights from the telecommunications data that are immediately useful for transport modelling and planning. Here are some examples:

Origin-Destination Matrix

Origin-Destination Matrices are the most important and commonly-used data source for many types of transport modelling, including mode choice, route choice, and trip generation models. Our origin-destination matrices provide data about the number of trips and people travelling from place to place, countrywide. They cover all trips, including private and public transport, from door to door. They also provide information about the duration, distance, and routes for these trips.

Home and Work

Everyone has a unique mobility pattern. By analysing each person's daily movement patterns, we can infer the actual home and work locations, and the time that each person commutes to and from work. This helps transport planners to optimise the network for peak loads during the morning and evening commute hours.

Mode of Transport

Our algorithms infer the mode of transport for each trip. We infer this using various spatio-temporal features in the telecommunications data, such as speed, direction, and distance to roads or rail lines.

Trip Purpose

By analysing mobility behaviours and cross-referencing them to external data such as point of interests (POIs), our algorithms infer the purpose of each trip.

Flexible Querying to Meet Every Need

We deliver the data to transport planners and modellers through APIs that allow the data to be queried in a myriad of ways, with full control over the query parameters to get exactly the data needed for each project. The APIs allow users to query the data through industry-standard protocols which are accessible from most programming languages and data tools. This is the optimal choice for users with basic programming skills and yet need full flexibility in querying the data.

About DataSpark

DataSpark is a mobility intelligence company which aspires to deliver mobility intelligence in every application for every interaction. We are a Singtel Group Company. Singtel is a leading global multi-media and communications service provider that serves 640 million customers around the world.

Operating in 5 countries, DataSpark's team brings a rich background of deep and relevant technical and domain expertise. We combine rich carrier datasets, scalable platforms and innovative solutions to enable a deep understanding of mobility patterns of people. Our relevant domain expertise allows us to build data products and services to deliver high business impact.

Get Started Today

For more information, see our website https://apis.datasparkanalytics.com
Contact us at api@datasparkanalytics.com

DataSpark