TRAFFIC MANAGEMENT USING MACHINE LEARNING

513221205019: SANTHOSH.D

Phase 2 Submission Document

Project: Traffic Management

Introduction:

Traffic prediction means forecasting the volume and density of traffic flow, usually for the purpose of managing vehicle movement, reducing congestion, and generating the optimal (least time- or energy-consuming) route.

Traffic prediction is mainly important for two groups of organizations (we're not talking about folks planning a weekend getaway, you know).

1. **National/local authorities**. In the last ten to twenty years, many cities adopted intelligent transportation systems (ITS) that support urban transportation network planning and traffic management. These systems use current traffic information as well as generated predictions to improve transport efficiency and safety by informing users of current road conditions and adjusting road infrastructure (e.g., street lights).
2. **Logistics companies**. Another area of implementation is the logistics industry. Transportation, delivery, field service, and other businesses have to accurately their operations and create the most efficient routes. Often, it's not only related to current trips, but also to activities in the future. Precise forecasts of road and traffic conditions to avoid congestion are crucial for such companies' planning and performance.

Data types and sources

Traffic is influenced by many factors, and you should consider all of them to make accurate predictions. So, there are several main groups of data that you'll have to obtain.

DATA NEEDED FOR TRAFFIC PREDICTION

Mapping data

Traffic data

Weather data

Other external data
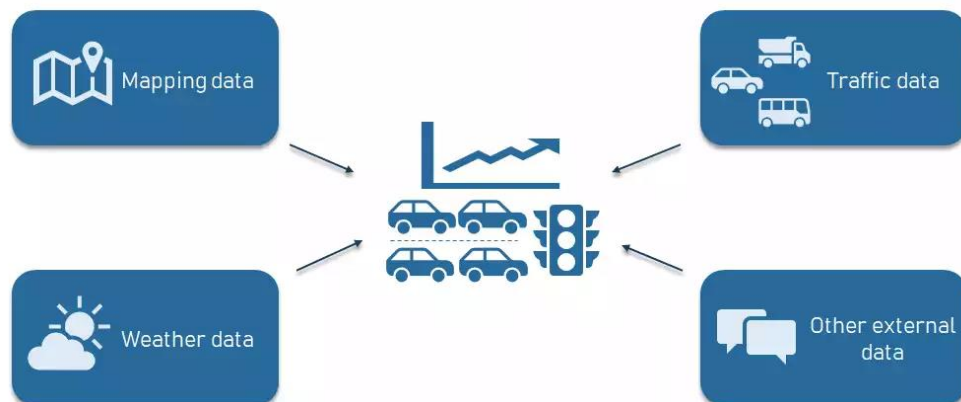
Mapping data. First of all, you need to have a detailed map with road networks and related attributes. Connecting to such global mapping data providers as [Google Maps](), [TomTom](), [HERE](), or [OSM]() is a great way to obtain complete and up-to-date information.

Traffic information. Then, you'll have to collect both historical and current traffic-related information such as the number of vehicles passing at a certain point, their speed, and type (trucks, light vehicles, etc.). Devices used to collect this data are

- loop detectors,

- cameras,

- weigh in motion sensors, and

- radars, or other sensor technologies.

Fortunately, you don't have to install these devices all over the place on your own. It's easier to get this information from the aforementioned providers that gather data from a system of sensors, diverse third-party sources, or make use of GPS probe data. (Just in case you're unfamiliar with how infrastructures for collecting, processing, and storing data are designed and work, you can visit our related post on [data engineering]() to get an idea.)

Other platforms such as [Otonomo]() use an innovative Vehicle to Everything (V2X) technology to collect so-called *connected car data* from embedded modems.

You can also get other important information on incidents (road closures or roadworks), places of interest, etc., from data providers.

Weather information. Weather data (historical, current, and forecasted) is also necessary as meteorological conditions impact the road situation and driving speed. There are lots of weather data providers you can connect to — such as OpenWeather or Tomorrow.io.

Additional data on road conditions. There are external data sources that can provide important information that impacts traffic. Think social media posts about sports events in the area, local news about civil protests, or even police scanners about crime scenes, accidents, or road blockages.
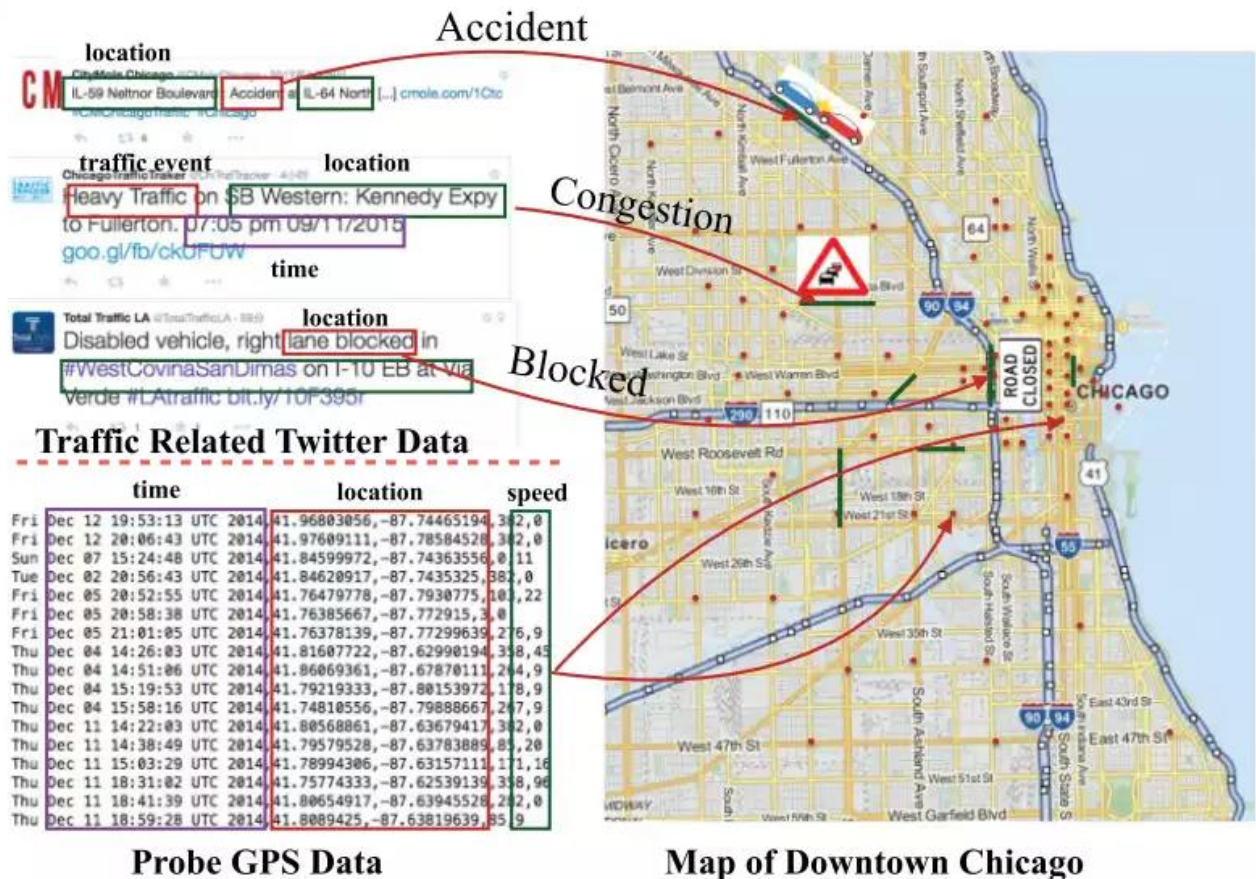
Algorithms for generating traffic predictions

Traffic prediction involves forecasting drivable speed on particular road segments, as well as jam occurrence and evolution. Let's take a look at different approaches to this task.
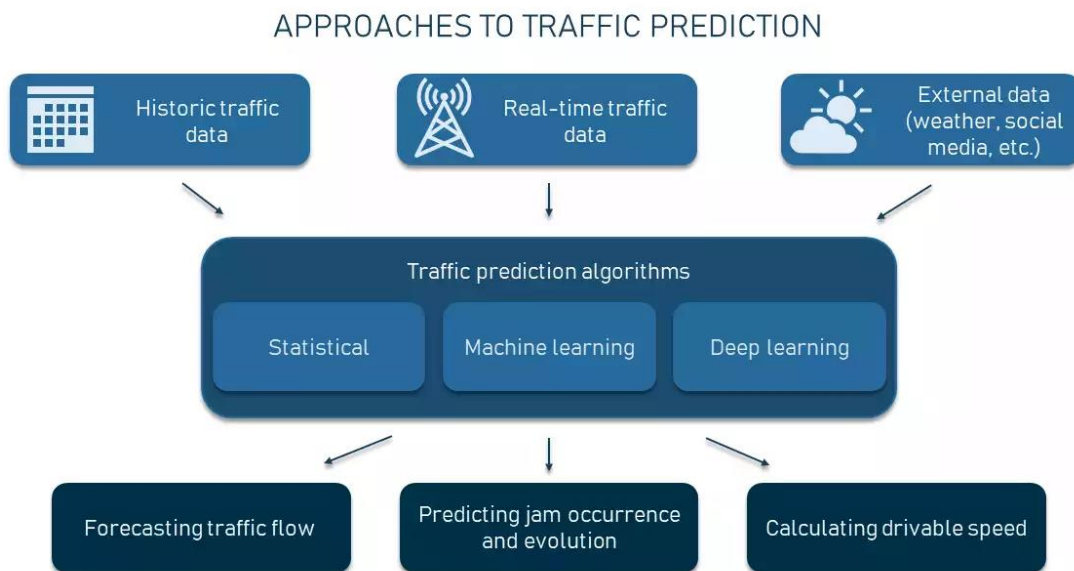
Statistical approach

Statistical methods allow you to identify traffic patterns at a different scale: during the day, on different days of the week, seasonal, etc. They are usually easier, faster, and cheaper to implement than machine learning ones. However, they are less accurate since they can't process as much multivariate data.

Specifically, **auto-regressive integrated moving average** (ARIMA) models have been actively used to predict traffic since the 1970s as they are easy to implement and show higher accuracy compared to other statistical methods. It's a classical statistical approach to analyzing past events and predicting future ones. It observes data that is collected from a

series of regular time intervals and assumes that past patterns will repeat in the future.

However, traffic flow is a complex structure with many variables that can't be effectively processed with the help of the univariate ARIMA models.

Machine learning approach

Machine learning (ML) allows you to create predictive models that consider large masses of heterogeneous data from different sources. Numerous studies have been conducted on the application of ML algorithms to forecast road traffic. Here are some successful examples.

The **random forest** algorithm creates multiple decision trees and merges their data to obtain accurate predictions. It's quite fast and can produce effective results given sufficient training data. When applied to the road congestion problem, this method showed an accuracy of 87.5 percent. In this case, weather conditions, time period, special conditions of the road, road quality, and holidays are used as model input variables.

Deep learning approach

Deep learning (DL) methods have proved highly effective in predicting road traffic in comparison to ML or statistical techniques, consistently showing about 90 percent forecasting accuracy and higher. Deep learning algorithms are based on neural networks.

Neural networks (NN) or artificial neural networks (ANN) consist of interconnected nodes (neurons) that are arranged in two or more layers and are designed to function similar to the human brain. There are many types of neural networks developed for different purposes. Here are some that were used in traffic analysis and prediction.

How to implement traffic prediction

If you run a logistics business, most likely you don't need traffic prediction by itself, but rather its impact on your operations. As we've already mentioned, accurate prediction is important for routing and scheduling purposes. If this is the case, there are three main ways to get those forecasts and [build optimal routes](#) (check our related article for more ideas and information).

Off-the-shelf solutions

There are lots of ready-made software solutions on the market developed for any type of business. If your company is small or medium-sized and your operations (be it field service, [last-mile delivery](#), taxi, moving, or long-haul transportation) are more or less standardized, you can find a tool that meets your needs and has routing capabilities to support your business activities.

[OptimoRoute](#), [Fixlastmile](#), [Badger Maps](#), [Route4Me](#), or [Routific](#) – a myriad of platforms offer route planning and optimization functionality (especially in the short-term perspective). The choice depends on your industry and specific business demands.

Custom development and API integrations

If you operate a large enterprise and have unique business requirements, consider building a custom model to solve your specific needs and implement it into your platform. Be prepared that it would require significant investment, skilled [data specialists](#), and a great deal of time to connect to diverse data providers and train those fancy ML/DL algorithms. On the bright side, you'll get your own predictions and stay independent of software vendors.

MAIN TRAFFIC PREDICTION DATA PROVIDERS

	Traffic prediction functionality	Frequency of updates	Routing functionality
Google Maps Platform	Only current traffic	Frequently	✓
Waze	Only current traffic	Every 2 minutes	✓
TomTom	Up to 24 hours ahead	Every 30 seconds	✓
HERE	Up to 12 hours ahead	Every minute	✓
ArcGIS	Up to 4 hours ahead	Every 5 minutes	✓
PTV	Up to 1 hour ahead	N/A	✓

Google Maps Platforms. If you want to partner with the biggest mapping data provider, keep in mind: you'll only be able to add the current traffic layer to your map, but no forecasts. Moreover, their documentation warns: "*Traffic information is refreshed frequently, but not instantly. Rapid consecutive requests for the same area are unlikely to yield different results.*" They also offer a diverse suite of APIs to add their map or routing functionality to your location-based product.

Waze. Another Google product and second most popular navigation app gets information on traffic, accidents, jams, and other road conditions only from their users driving around. The information is updated every 2

minutes and can be accessed via localized XML and JSON Geo RSS [data feeds](). Note, that this information will be reliable only if there is a sufficient number of drivers in the area. You can also embed [Waze Live Map]() with search and routing functionality on your website. As for predictions, the app doesn't make and provide any.

TomTom. [TomTom's Traffic RESTful APIs]() give access to historical and real-time data related to traffic incidents and flow. TomTom makes use of over 600 million GPS and floating car data probes to collect up-to-date information (updated every 30 seconds), analyze it, and make predictions up to 24 hours ahead. You can leverage this data to create routes with your own app or get their comprehensive [routing product]().

Project Conclusion:

In the Phase 2 conclusion, we will summarize the key findings and insights from the innovation to solve the problem. We will reiterate the impact of these techniques on improving the traffic data and machine learning algorithms.