

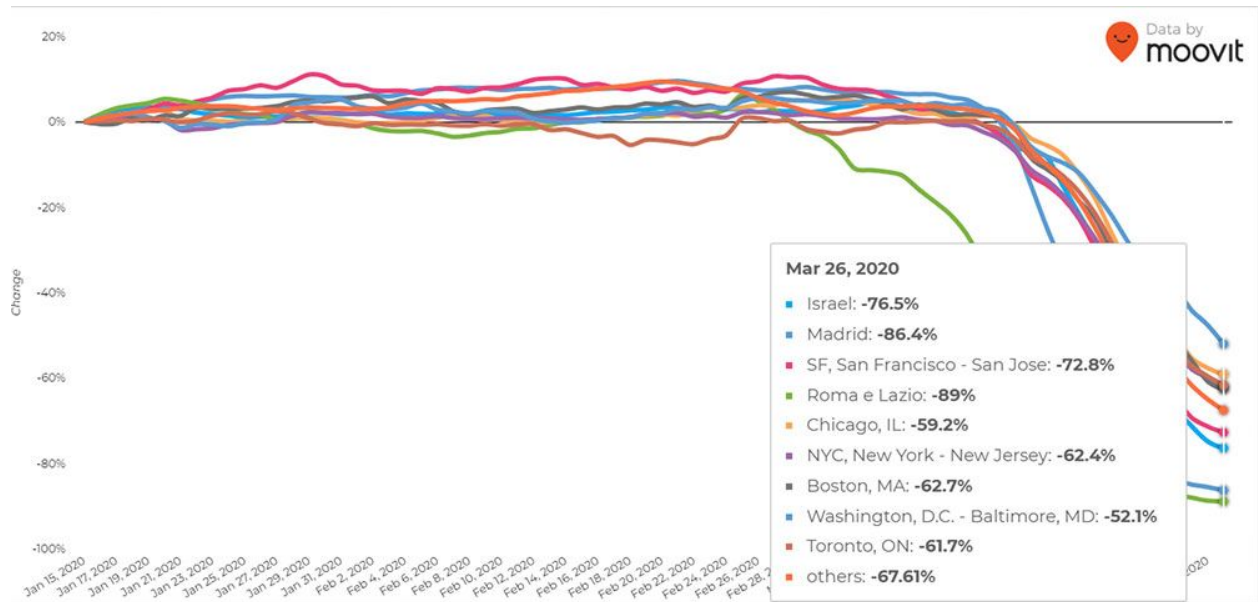
Safe and Efficient Transportation in the Face of the Covid 19 Pandemic

Problem statement

Our daily lives have been affected in diverse ways due to the Covid-19 pandemic. Many people lost not only their jobs and means of livelihood, but also their own loved ones. Due to the prevalence of such circumstances, a major part of society has been led into a state of confusion and apprehension towards transport. Although few kinds of jobs, like those in the software industry, could be easily performed from the comfort of a home, many other kinds of people with various professions like daily wage labourers, electricians, plumbers, delivery persons, cab drivers, bus drivers and conductors, auto drivers and medical students simply could not get their work done, without having a safe, reliable and efficient means of transport. As they hope for the city to reopen as soon as possible not having a safe and reliable means of transport may mean they won't get their pay for the day, leading to undesirable consequences.

While many countries, and cities in particular, are finding their way during their respective reopening stages, the transport sector needs careful thought and planning. Experts predict that as people try to avoid crowded places, people would prefer using their own vehicles or taxis, thus leading to a significant increase in traffic congestion. They also predict that the traffic congestion would be more severe than before the pandemic. In such a setting, public transport would take a very long time to recover. The local transport authorities would then become massively burdened as they would have to grapple with financial losses suffered by the public transport sector as well as handle the severe road congestions.

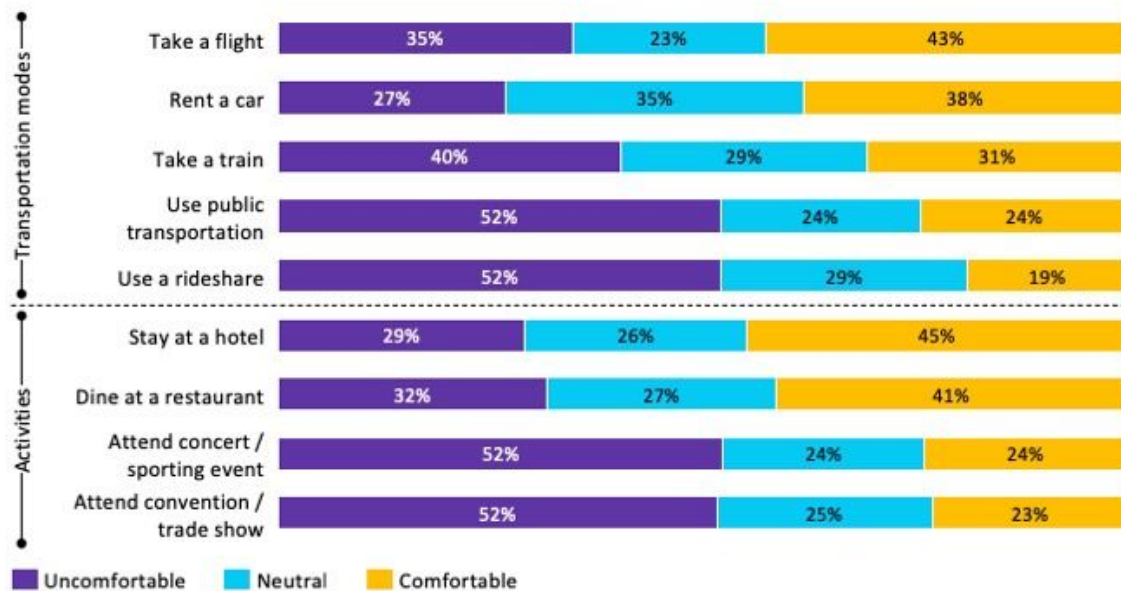
Many of us are unaware of when public transport is safe to use and when it isn't. We don't know when social distance is practiced (i.e when the population density is less) and we don't know when social distance isn't practiced (i.e when population density is high). While paying for a cab or auto or using a private vehicle for a couple of days a week may be alright, it may not be a feasible daily option for everybody. Hence, monetary cost, duration of travel and more importantly - risk of infection come out as the 3 most important factors while choosing a mode of transport.



COVID-19 CONTINUES TO MAKE CONSUMERS WARY OF SOME ACTIVITIES

After the COVID-19 outbreak ends and travel restrictions are lifted, how comfortable will you feel doing each of these activities?¹

Percent of respondents, (n=4,594)



1. Uncomfortable = "Very uncomfortable" or "Somewhat uncomfortable"; Comfortable = "Very comfortable" or "Somewhat comfortable"

Source: Oliver Wyman Traveller Sentiment Survey, Oliver Wyman analysis

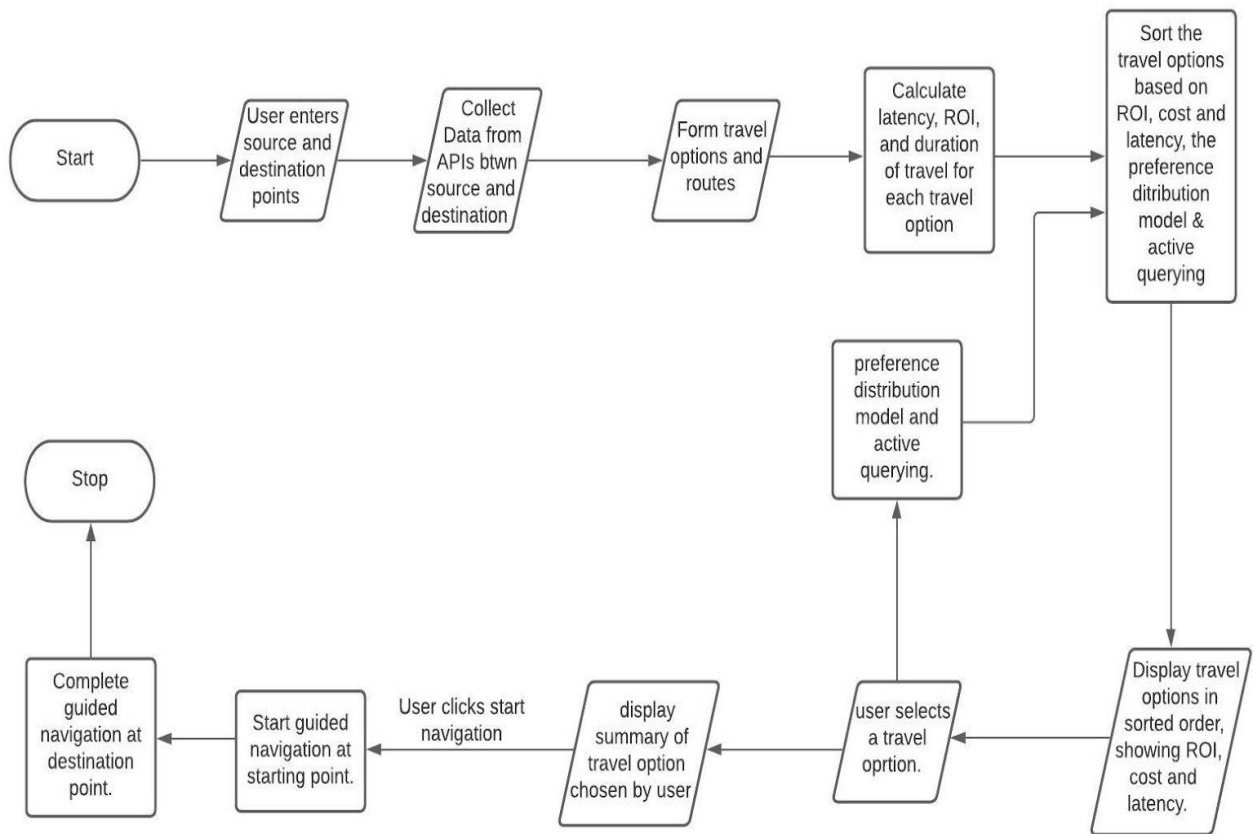
© Oliver Wyman

As seen in the above data, public transportation has taken a major plunge in ridership and is also one of the most uncomfortable options for travel for most people. As a student, studying computers at college, I want to be a part of the solution to revive transportation in the city and empower people to use transport in a safe and reliable way.

Problem Solution

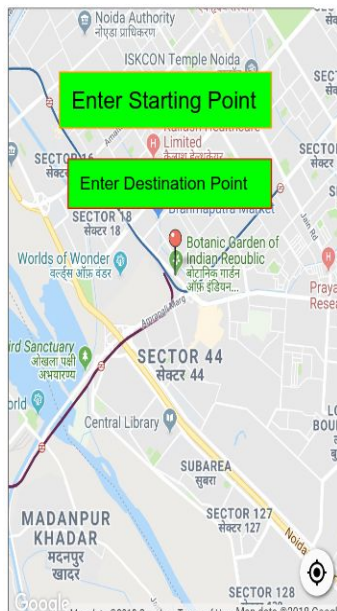
The problem at hand is clear - we want to transport ourselves from one place to another, but lack a safe and feasible system to do so. Different cities have tried to come up with different solutions. London and Paris have begun working on plans to create more space for bicycling and walking lanes, whereas Istanbul implemented a model of alternate working hours for public servants and schools. But, especially in an Indian context, this may not be feasible, and may not even be possible. So we need to address the challenge of modelling how Indian user preferences have evolved through the pandemic and how these preferences must affect routing, public transportation and ride hailing services.

My inspiration for this solution comes from [this](#) paper by Beliaev et al. The paper gave me deep insights into what it takes to build and implement an end-to-end transportation network model, preference distribution model, how to learn user preferences and how to optimize for safety and efficiency. My goal is to implement a few of the mentioned concepts in the paper in an Indian context of transportation, along with some of my own ideas - like predicting the population density at a future time in a public transport setting using a machine learning predictor like a neural network. Below is the flowchart for the working of the system I would like to build. The sections after that explain the steps involved at each stage.



ROI - Risk of Infection

User enters source and destination points



The interface that greets the user would be much similar to Google Maps, Ola or Uber - a prompt for the user to enter their starting point and destination point. As a minor upgrade, the map in the background could show live traffic data, as well as areas with a high risk of infection in specific colour coding, like red for areas with a high risk of infection, and a green colour for areas with a low risk of infection. The maps could be easily incorporated into the application using the Google Maps SDK.

Collect data using various APIs

Population Density to calculate flow of people (in buses, trains and footpaths) - GMaps API

Travel Route Options - GMaps API

Traffic Data (like estimated duration of travel) - GMaps

Ola(or any other ride-hailing service) auto ride estimate - Ola API

Ola cab ride availability - Ola API

Cost of the ride - Ola API

The incorporation of APIs from other ride hailing services like Uber can be easily accomplished. Here Ola is used only for the purpose of demonstration.

A sample of the parameters of the Ola API which could be used in our application to get the ride availability details - estimated time of arrival, cost of the ride etc can be found [here](#).

A sample response of Gmaps API response message which includes the parameters that we would need can be found [here](#).

Below is a screenshot of a [heatmap based on population density in Google maps](#) which can be obtained from the GMaps API.



Form Travel Options and Routes

After collecting the data from the APIs, it is important to store them in suitable data structure, so that it can be easily and meaningfully accessed and retrieved. The implementation for this can be seen in the GetFromGmapsAPI and GetFromOlaAPI functions in the [code](#) I wrote.

In my implementation, I have considered 6 different modes of transport -

1. Railways
2. Buses
3. Private Vehicle - 4 Wheeler
4. Private Vehicle - 2 Wheeler
5. Cabs/Auto and
6. Walking

Calculating Latency, Cost and Risk of Infection for each travel option

Latency, which is basically the duration of travel, and cost of each travel option can be easily obtained from the APIs mentioned above. We now have to calculate the Risk of Infection.

As mentioned in the paper, the risk of infection for all people in the network is given by -

$$\text{Risk: } R(f_v, f_r, f_p) = R_v(f_v) + R_r(f_r) + R_p(f_p)$$

Where $R_p(f_p)$ is Risk of Infection for pedestrians given the flow of pedestrians,
 $R_r(f_r)$ is Risk of Infection for railway passengers given the flow of railway passengers,
 $R_v(f_v)$ is Risk of Infection for vehicles given the flow of vehicular movement on the road, and
 $R(f_v, f_r, f_p)$ is the total risk of infection in the network given flow of people in vehicles, footpaths and railways.

The math to find risk of infection for each component i.e railways, pedestrians and vehicles is clearly described in the paper and is easy to implement. It is possible to use computation of ROI (Risk of infection) in railways to compute the ROI in buses as well. Similar calculations are done to calculate latency i.e duration of travel.

So for each travel option we obtained using the GMaps API, we find the total risk of infection. Now for each travel option, we have 3 different output parameters - monetary cost, duration of travel/latency and Risk of infection.

It is important to note that we need to calculate ROI ahead of time. ROI needs to be predicted. The Google maps API gives us a lot of information, but DOES NOT give us information like - how many people **will** get on the bus/train, how many people **will** get off the bus/ train at intermediate stops. Since ROI depends on the number of people present at that particular location, this is a factor that would greatly determine the reliability of the application.

Let us suppose the local/national authority for buses and trains released the data of how many people get on and get off the bus/train in the intermediate stops (the bus stops between the onboarded stop and alighted stop). We could then use that data to train a machine learning model, like a neural network, to predict ahead of time, how the flow of people on the same bus/train would change with time, and use that to calculate the risk of infection. Unfortunately,

such data isn't available yet. On the flip side, it would not be hard for the authorities to obtain such data as most bus and train bookings happen through the internet and can be easily retrieved.

















Sorting the travel options and updating preference distribution model and active querying

We first now pass the 3 parameters into an optimization function, where we find a weighted sum of these 3 parameters i.e cost, latency and ROI. We do this for each travel option and sort the list of travel options based on the output value of the optimization function for each travel option. The objective is to sort the travel options in ascending order of optimization function output. Therefore, the travel option which has the least weighted sum of the 3 parameters would have a higher preference.

An optional 4th parameter for sorting the recommended travel options comes from the preference distribution model. This step is explained further in the next few steps.

Display travel options in sorted order, showing ROI, cost and latency, and the user selects their travel choice

Select Travel Flow Choice

		<u>ROI</u>	<u>Cost</u>	<u>Duration</u>
1	 →  →  →  → 	3.6	27	1hr 30 minutes
2		4.0	50	1 hr
3		0.0	70	45 minutes
4	 →  →  → 	3.75	67	1hr 40 minutes
5		0.0	90	37 minutes
6	 →  →  → 	3.75	95	1hr 47 minutes

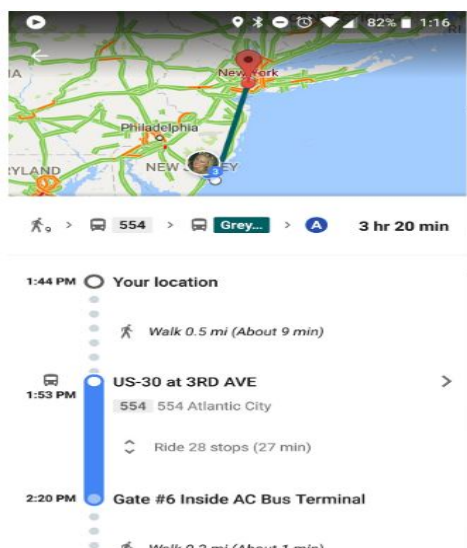
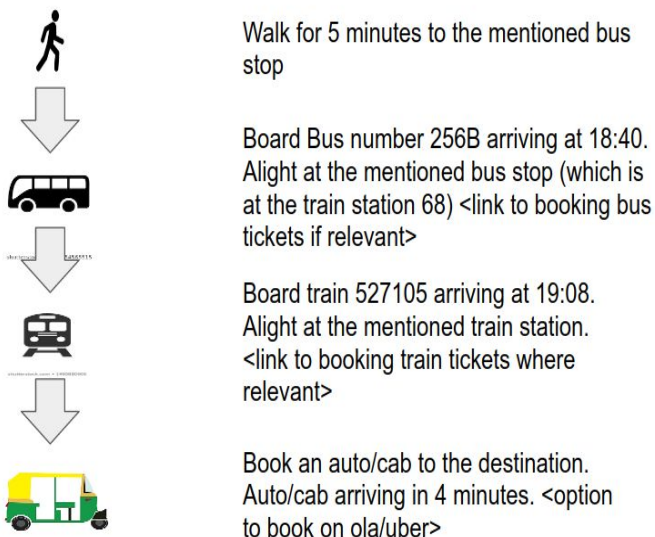
We now display the recommended travelling options in sorted order and wait for the user to select one of the travel options.

After the user selects their choice, we update the preference distribution model and the active querying model. This offers a more personalized order of recommendations based on the users past preferences. For example, when we know that a user prefers to book a taxi instead of travelling by a private car (maybe because they would like to read while travelling), then the model recommends taxis above private vehicles. The math involved in building the model is clearly explained in the paper and can be easily implemented.

Display summary of travel option chosen by user

After the user has selected their choice, we then display the travel flow summary wherein we specify the transitions (if any), booking options, specific bus numbers, specific train numbers, time left for arrival of a cab/auto and how long we need the user may need to walk for.

Travel Flow Summary



The interface would be similar to that on the left i.e the same way transitions between different modes of transport are displayed in google maps.

Start guided navigation at starting point

After the user clicks the start button, the assistant will then start the guided navigation just like how google maps works.

Scope for Improvement and Conclusion

The implementation of a machine learning model to predict the number of people getting on and off intermediate stops would be crucial in determining the risk of infection. Although this can be done accurately only after the data is made available by the authorities in question, for the time being, we could use a simulator to train and evaluate our end to end transportation model.

Also, additional features, like recommending safety precautions at certain places could be incorporated. We could also try to obtain information regarding cabs that was recently used by a person who tested positive for covid-19, and their primary contacts, and use that to calculate the degree of risk on infection based on factors related to - if the car was disinfected regularly or not etc.

Also, I look forward to implementing the entire transportation model, including the preference distribution model and active querying, including all the math and the flow of processes involved in calculating the risk of infection and recommending travel options.

Although this idea isn't necessarily 'groundbreaking' per se, a real world implementation of the idea would go a long way in helping our cities revive their transportation services as well as keeping the road congestion under control. This would mean a lot to cities, especially those cities having high levels of road congestion and population, like the city of Bangalore, where I live. Our goal is to empower people to achieve more. I hope this idea conveys hope, and I look forward to being a part of our answer to the transportation problems of our time, so that everyone is empowered to travel safely and efficiently in the face of the pandemic.

A basic implementation - the prototype of the model can be found on my github [here!](#)

