Oulu Public Transport Data Analysis

Aviroop Mukherjee
Biomedical Engineering
University of Oulu
aviroop.mukherjee@student.oulu.fi
2643423

Gunjan Chandra
Computer Science and Engineering
Biomimetics and Intelligent Systems Group
University of Oulu
gunjan.chandra@oulu.fi
2576930

Kennedy Opoku Asare
Center for Ubiquitous Computing
University of Oulu
kennedy.opokuasare@oulu.fi
2510116

Mira Rauhala
Industrial Engineering and Management
University of Oulu
mira.rauhala@student.oulu.fi
2427270

***Abstract*—Public transport is one of the services that everyone has used at least once in their lifetime. It financially benefits communities, reduces air pollution and traffic congestion, saves money, and increases mobility. The main objective of transport providing agencies is to meet the needs of the customers while keeping the system organised. We analysed the data provided by the Oulu joukkoliikenne (city of Oulu's public transport) agency to understand how they are accomplishing those goals. We started by analysing the schedule structure of buses and similarities in the different routes and how we can improve the existing system. In the city of Oulu, bus services are at its peak in the afternoon period, and most routes are similar in terms of the number of stops and travel time and distance between each stop and for the entire trip. Moreover, using the skip-detour method, routes were modified to reduce the travel distance, time, fuel cost, and exhaust emission. In addition to customer satisfaction that these proposals in our paper will target, the results of reductions listed before will improve environmental and economic sustainability.**

***Index Terms*—Public transport, Route optimisation, Exhaust emission, Route correlation, Frequency of bus, Sustainability**

I. Introduction

We are writing this paper as a group for a course called Data Mining project at the University of Oulu. At the beginning of the course, students were able to wish for a data set based on provided data sets by the course organisers. After that students were divided into groups based on their preference for the data sets. During the course we had 3 presentations in class about the project and the stages of it.

Our data set contains data from Oulu joukkoliikenne (city of Oulu's public transport) agency. Public transport financially benefits communities, reduces air pollution and traffic congestion, saves money, and increases mobility. Our intention was to evaluate network performance, reliability, and accessibility of Oulu public transportation. Our analysis focuses on stoppage time of buses and their subsequent availability to passengers. Based on this information we are planning on creating a public transport route optimisation system. We are conducting this project from customer's perspective and thus we hoped to improve customer satisfaction as well as improving environmental and economic sustainability.

During our project, we faced some issues posed by the data set we were working with. We found out later during the project that the data in our data set was not real-time GPS data but based on the bus timetables. This means that the data in the data set was the same as the scheduled bus timetables, that is to say that there was no data of how long the bus stopped at a particular stop nor what happened with the bus schedules in real life e.g. if the bus was late or on time. Having discovered this, we had to modify our objectives and we were not able to conduct the type of data mining that we had intended in the beginning of our project. Eventually, we were able to work with the data set and mine helpful information from it.

We started by analysing the schedule structure of buses and similarities in the different routes and how we can improve the existing system. In the city of Oulu, bus services are at its peak in the afternoon period, and most routes are similar in terms of the number of stops and travel time and distance between each stop and for the entire trip. Moreover, using the skip-detour method, routes were modified to reduce the travel distance, time, fuel cost, and exhaust emission. In addition to customer satisfaction, this optimisation of public transport improved environmental and economic sustainability as well as social sustainability.

II. Related work

In public transport, the term accessibility is measured in ease by which a person can reach a mode of transportation [1]. Estimating the average time travelled from one neighbourhood to its nearest supermarket [2] provides us how a planned structure lacks accessibility in some areas and how it can be improved. Another study shows how services provided by public transport agencies differ from demand [3]. Whereas, accessibility is something we can not guarantee just with a transit timetable. The study shows that when a transit timetable is compared with a real-time feed, the transport gaps are even bigger [4].

For very long, researchers are trying to solve the issue of accessibility to improve the existing system. In a study

of public transport route optimisation by scaling down the available street network to a level where optimisation methods such as genetic algorithms can be applied [5] could be one of the economical solutions for the future.

As we have stated, it is important for us to take customers' needs into consideration. Related work about a demand-responsive transport service introduces us to the following major challenge: "In order to be able to compete with private cars, service should be available within a short period of time from the trip request. In order to ensure a sufficient level of service, the customers' waiting and ride times should be relatively limited." The paper suggests that the vehicle dispatching algorithms should be designed in a way that the constrained nature of the problem is taken into account. [11].

Regarding the quality of service and customer satisfaction, there has been conducted research about why customers have left the services of public transport and started using private cars. Research suggests that the main reason for leaving or returning to public transport use is a change in the work-place/school/residence. This means that when public transport became less convenient than the car, a customer is changing the way of travelling. Understanding this is important when rerouting buses and optimising bus systems. [13].

Achieving the environmental and economic sustainability mentioned in the objectives, we base these goals on the principles of sustainable business model elements that consist of economic, environmental, social and multidimensional characteristics [12].

As it is easy to understand the economic and environmental characteristics of sustainability through reducing fuel consumption and emission levels in public transport, the social aspect is also as important. It has been discussed that the social sustainability in public transport or mass transit is e.g. offering lower cost option for low-income customers and commuters [14].

III. Objectives

Our research consists of three main objectives that are connected. The first task is to conclude about the frequency of bus numbers and routes by the time of the day. This objective gives us an insight of all the buses and their number of stop-to-stop trips according to the time of the day which is the following: morning (05-11), afternoon (11-15), evening (15-19), night (19-05). This also gives further insight into how frequent a particular bus number or bus route is in the particular time of the day.

The second objective we aim to compute optimised bus routes by using alternative routes. The main objective is to optimised the bus routes and build an optimised system to reduce cost for agencies while underlining the customer demand. This system will reduce financial the costs by lowering fuel consumption. Additionally, this solution will reduce environmental impact by lowering Co2 emission. For customers this provides reduced travel time by alteration of bus routes, including commuter customers. A skip-stop / skip-detour method will be used to achieve this. The main idea is to explore the possibility to prove the hypothesis with the given data set.

For the third and final objective the task is to visualise our results such as travel distance, time, fuel cost, and exhaust emission. This will be conducted by comparing values before and after calculations. Our hypothesis thus is that the alterations on the public transportation we are suggesting, will reduce cost, travel time and travel distance as well as exhaust emissions and thus improve efficiency and sustainability.

To this end, we summarise the objectives of the project as follows:

- To conclude about the frequency of bus numbers and routes by the time of the day
- To compute optimised bus routes by using alternative routes
- To visualise travel distance and how close bus stops are

IV. Data

The data set used in this project is the regional schedule and route data produced by Oulu public transport agency [10]. The data is in GTFS (General Transit Feed Specification) format. GTFS is a data specification that allows public transit agencies to publish their transit data in a format that can be consumed by a wide variety of software applications. The material contains public transport information e.g. bus stops, public transit routes, and line schedules by bus stops. The data is supposed to be of last 30 days and is updated daily but it came out in the research that there more data available than just of last 30 days. Also, we tried sending an email to city of Oulu to get access to the real-time data but the respond for our email came too late since the schedule of this project was challenging and we had to settle for the data we originally had.

In the exploration of the data, we found that Oulu public transport data has two calendars: winter calendar and summer calendar. Summer calendar time is from 2019-06-03 to 2019-08-07, and it offered 212 services. The winter calendar is from 2019-080-8 to 2020-05-31, and has offered 325 services. The services are offered on 129 bus routes with 1943 bus stops. Each route, stop and trip have their unique ID with 2 directions. There are in total 6388 trips for all routes.

From the data we discovered, using R, that the top 5 most busiest bus stops are Pokkitörmä with 2802 stops, Kaupungintalo E with 2185 stops, Toivoniemi P with 2182 Stops, Raati P with 2182 Stops and Toivoniemi E with 2152 stops. For each trip, a bus stops at minimum of 10 stops, a median of 33 stops, a maximum of 87, and a mean of 45.59 stops (SD 18.33).

Unfortunately we discovered, that the data in this data set is not real-time data with GPS information which poses challenges for the purpose of our research. This means that the data in the data set was the same as the scheduled bus timetables, that is to say that there was no data of how long the bus stopped at a particular stop nor what happened with the bus schedules in real life e.g. if the bus was late or on time. Having discovered this, we had to modify our objectives.

From the raw data, information was segregated based on bus numbers following which considering each bus number, the information was further segregated based on the direction that they travel: P (North) or E(South). This followed the process wherein for each such direction, its respective day-night time table was considered. Simultaneously, the day-night time table was divided into the sections which gave us sections of morning, afternoon, evening and night with respective time-slots of 0500 hrs - 1100 hrs, 1100 hrs – 1500 hrs, 1500 hrs – 1900 hrs and 1900 hrs and above. This formed the basis on which further analysis and conclusions were made concerning the frequency of buses by the time of the day.

For the pre-processing of the data, there was a need for information about identifying detours on the map, calculating number possible of routes for this research and sampling the data after detours were removed from the route.Tools used for pre-processing listed: Google Maps; Python (with libraries of googlemaps, pandas, numpy, json, tKinter, glob, os); R (with library of dplyr); Full support of Google Cloud platform; Matlab (with R2015a).

V. Data pre-processing and Methods

A. Frequency of bus

The methodology followed here is an exemplification of timetable classification wherein after having subjected the raw data to primary segregation followed by specific segregation of the same according to the time of the day (morning: 0500-1100 hrs; afternoon: 1100-1500 hrs; evening: 1500-1900 hrs and night: 1900 hrs and beyond), we orient the morning to night data of each bus route into separate spreadsheets and export them in the Comma Delimited CSV format. A developed code is then used to run the data and plot the same according to the specified time sections. The distinct colour coding has also been used in order to make it easier to identify the respective correspondence of data according to the time of the day.

B. Feature Extraction and Cluster Analysis

To further understand our data set, beyond data pre-processing, we analysed our data set with feature extraction and Cluster Analysis. The feature extract and Cluster Analysis was used to highlight the underlining features, similarities, differences and clusters within our data set. To achieve this aim, we extracted 14 different features for all 92 routes using estimations such as count, mean, maximum, median, standard deviation and entropy, of the distances and time between consecutive bus stops of each route. Table I lists the features extracted from the travel time and distances between successive bus stops of a route. We used the bus stop sequence to determine the correct order of bus stops for a trip on a route.

Using the features extracted from the travel time and distances between bus stops of a route, we computed the Euclidean distances between each pair of route features. The euclidean distances between each pair of route features, thus the distance similarity matrix, would highlight the similarity between the pairs of the routes, in that, similar route have closer and smaller Euclidean distances between them. We

TABLE I
Features extracted from time and distance between successive bus stops of a route

n_stops	
time based features	
total_time	mean_time_btn_stops
sd_time_btn_stops	median_time_btn_stops
max_time_btn_stops	entropy_time_btn_stops
distance based features	
total_distance	mean_distance_btn_stops
sd_distance_btn_stops	median_distance_btn_stops
max_distance_btn_stops	entropy_distance_btn_stops
norm_entropy_distance_btn_stops	max_distance_btn_stops

computed and visualised the Euclidean distances using the *factoextra* R package [8] for multivariate data analysis. Figure 1 visualises the similarity between routes. The x and y axis of Figure 1 shows the route pairs and the colouring on the x,y points shows the their Euclidean distances

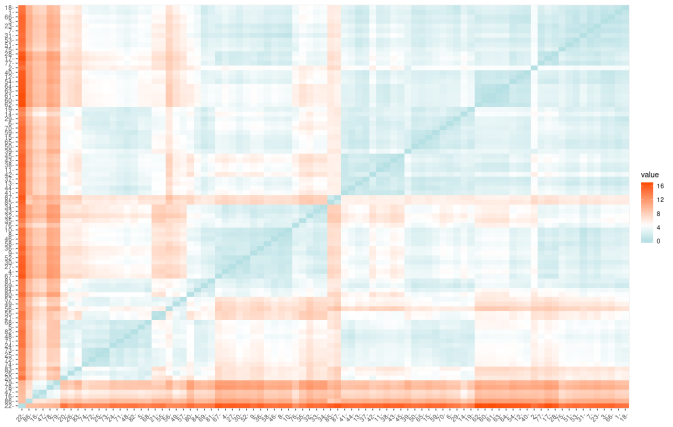


Fig. 1. Visualisation of Euclidean distance similarity matrix of routes. Red colour gradient denotes dis-similar routes, and teal colour gradient denotes similar routes.

Furthermore, we used an unsupervised machine learning method, K-means clustering, to determine the clusters within bus route features. The best choice of K in K-means clustering, thus, the number of groups or clusters to determine from the features data is essential in revealing the actual clusters in the data set. In view of this, we used multiple K values , k=2,3,4 and 5 in our cluster analysis, to highlight the clusters for each case. Figure 2 presents a visualisation of the clusters in the computed route features.

C. Optimisation

Finally to propose a more optimised bus route system, we used a variant of Skip-stop method [6]. From the raw data, we calculated the number of unique routes. Oulu pubic transport data have 92 unique routes out of 129 so far in transit timetable. Out of those 92, we used 13 for our model. For each unique route of selected 13 routes, we manually determined detours and calculated a direct distance from start to end of detour using Google maps API request. Figure 3 represent an example of detour. We also calculated the distance between

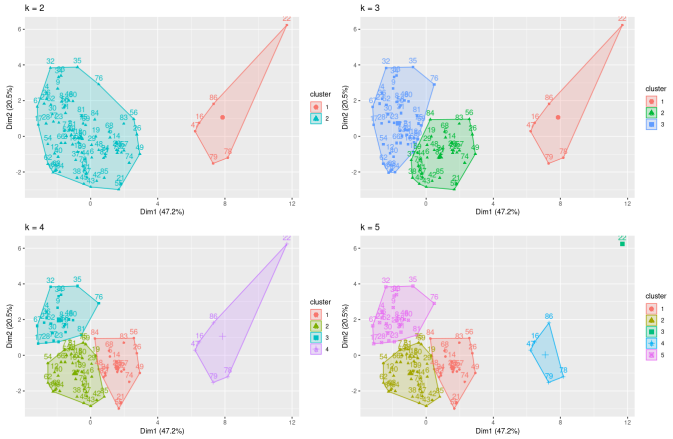


Fig. 2. A visualisation of the clusters in the computed route features, using K-means clustering for K=2,3,4 and 5

each bus stops in similar manner as mentioned earlier to calculate the total distance of a route with and without detours. Finally we calculated the estimate time taken by the bus to complete each route.

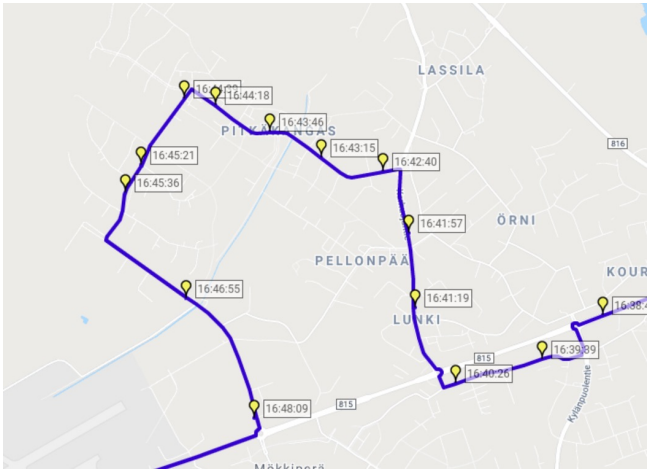


Fig. 3. Example of a detour. For a bus travelling from the right hand side of this figure to the left had, side, the bus could go straight to its destination without going through the branching or detour stops

Oulu public transport uses buses from various companies, each company has its models, and each model has its variations. For this study, we selected three bus manufacturing companies; Volvo, VDL, and Scania, which provide buses to the city of Oulu. From the before specified companies, we chose the most commonly used bus model, which has three axles. As per European emission standards, emission levels permitted to a bus are Euro 6 standard. Figure 4 represents the exact permitted numbers of emission.

For calculating the level of emission by a bus for every trip, we took an average of the engine power of three selected bus models. Based on estimated time taken for a bus to complete the route we calculated the estimated emission based on an average level of emission permitted per hour. Other factors

European emission standards for heavy-duty diesel engines, g/kWh

Tier	Date	Test cycle	CO	HC	NO_x	NH_3 [ppm]	PM	PN [#/kWh]	Smoke [m^{-1}]
Euro I	1992, < 85 kW	ECE R49	4.5	1.1	8.0		0.612		
	1992, > 85 kW		4.5	1.1	8.0		0.36		
Euro II	October 1995		4.0	1.1	7.0		0.25		
	October 1997		4.0	1.1	7.0		0.15		
Euro III	October 1999 EEVs only	ESC & ELR	1.5	0.25	2.0		0.02		0.15
	October 2000		2.1	0.66	5.0		0.10 0.13*		0.8
Euro IV	October 2005		1.5	0.46	3.5		0.02		0.5
Euro V	October 2008		1.5	0.46	2.0		0.02		0.5
Euro VI	31 December 2012[15]	WHSC	1.5	0.13	0.4	10	0.01	8×10^{11}	
		WHTC	4.0	0.16	0.46	10	0.01	6×10^{11}	

Fig. 4. European emission standards for bus [7]

such a fuel consumed for each trip was also calculated to support the optimised model in terms of cost.

VI. Results

The principle aim of the above was to present a technical melange of ideas relating to the concerning data of Oulu bus transport data set, wherein determination of frequency of bus routes according to the time of the day, computation of optimised bus routes using alternative routes and pollutant emission study were the main running themes. Based on the about 10 percent of workable data, following were the results pertaining to each theme.

A. Cluster Analysis

Our results from the cluster analysis shows two main cluster of routes. As depicted in Figure 2, the clusters does not change significantly when the K was increased from 2 to 5. Examining the features of the clusters, we find that Cluster 2 contained only 7 routes which had higher number of bus stops (*mean number of stops =62.285*), as compared to Cluster 1 routes with relatively smaller number of stops (*mean number of stops= 47.65*).

Additionally, routes in Cluster 2 had higher distances between stops (*mean total travel distance between stops = 81.571 kilometres*) and longer total travel times (*mean total travel time between stops = 62.285 minutes*) as compared to routes in Cluster 1 with total distances between stops (*mean total travel distance between stops = 65.66 kilometres*) and to travel times (*mean total travel distance between stops = 48.085 minutes*)

Surprisingly, route 22 starting from OSAO PP and ending at Haukiputaan ammattikoulu was found to be quite dis-similar to the rest of the route as show in Figure 1, and also quite dissimilar to the Cluster 1 and Cluster 2. Route 22 was found to have 59 stops, 69.6 km total distance and 65 minutes total travel time.

B. Frequency of Bus

Figure 5 depicts the number of trips by the time of the day across all bus numbers and in respective directions (0 and 1). Figure 5 is the graphical representation of all the buses and their number of stop-to-stop trips according to the time of the

day. This helps us understand that having divided the day into four distinct time-classifications and having further divided all the available bus routes into these classifications, each bus number has a certain frequency, for example Bus No. 3 has a trip-frequency of 48 times in the morning whereas the same bus has a trip frequency of 52, 38 and 30 in the afternoon, evening and night times. This helps the user comprehend how frequent a bus is during a particular time of the day. With further research, this concept forms the base of Consumer Economic Benefit Analysis of bus travellers in the City of Oulu.

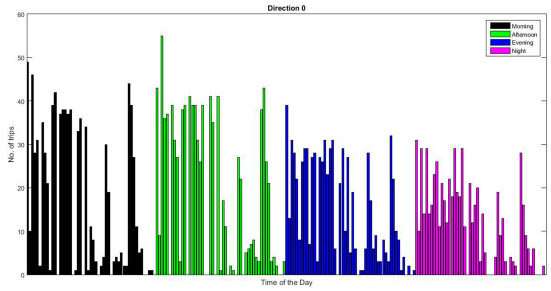


Fig. 5. Frequency of Bus Numbers/Routes showing the number of trips for each bus route occuring according the Time of the Day

C. Route Optimisation and Emission Studies

In this section, the optimisation of routes using the concept of "implementation of alternating detours" was applied wherein having optimised the studied routes based on distance travelled, time travelled, fuel consumption and cost and emission studies, the results, as depicted in Figure 7 are obtained.

Difference before and after removal of detours

Route_id	Distance difference (KM)	Time difference (Min)	Fuel (%)	Emission (%)
1	2.1999999999999957	3.7499999999998934	3.832752613240416	6.818181818181614
8	0.29999999999999716	0.7666666666667012	0.393184796854527	1.7037037037037805
9	4.5	7.25	4.712041884816756	11.153846153846146
10	5.899999999999999	4.20	13.470319634703202	7.924528301886795
14	0.0	0.0	0.0	0.0
34	0.0	0.0	0.0	0.0
48	0.30000000000000426	0.6499999999999986	0.8174386920981078	1.4444444444444429
53	8.900000000000006	7.3333333333333	8.356807511737102	13.333333333333258
54	8.899999999999991	6.149999999999899	9.026369168356993	15.374999999999744
57	0.0	0.0	0.0	0.0
106	0.40000000000000036	1.0999999999999996	3.80952380952381	7.857142857142847
244	6.799999999999997	6.983333333333306	6.087735004476272	8.952991452991412
568	8.8	12.75	40.0	39.84375

Fig. 6. Difference before and after removal of detours

As seen in the Figure 6 above, it is clear that the distance and travel time in most cases are reduced since the detours have been removed. When regarding to reduce cost, lower fuel consumption equals less cost. This seems to be achieved in most of the routes as well as cutting down emissions.

VII. Discussion

The determination of bus frequency forms a base of qualitative analysis concerning a consumer-controlled estimation of the availability of such buses in respective or desired locations. Alongside, the route optimisation system seems to improve efficiency concerning consumer logistics and decrease fuel cost. Having said that, a more profound research would require more detailed information about the data and other allied factors. Besides, the removal of detours as a part of route optimisation would require more data about factors e.g. consumer density, traffic data and ticket sales. All in all, based on the findings presented, it can surely be said that the potency of such is deep-seated and is an apt exhibition of the cynosure and novelty of this research title. Discussing the results of this project it could be that the emission decrease of our findings in reality is greater than what research proposes because the buses do not have to stop for all the stops that were in the detour and this decreases emissions.

Besides not having the real time data, there were also other limitation for our research. These limitations are referring to factors that would have been beneficial for this research in order to achieve a more profound optimisation. It would have been beneficial to have knowledge about purchased tickets for each route and bus stop, so that we would get information about the demand of the buses and to get knowledge of what is the most optimal in terms of altered routes, skipped stops etc. In addition, not having data about he stopping times, as in how long a bus stayed at a particular bus stop, limited our research. This would be interesting subject of research for the future since it can give information about the schedule reliability of buses. The skip stop method mentioned in our research could be especially be altered for rush hours, for Airport bus routes (flight information required) and for commuters in the morning and evening, when these research limitations mentioned above are conquered. Also, a questionnaire for customers about desired changes would give a good insights of customers' needs.

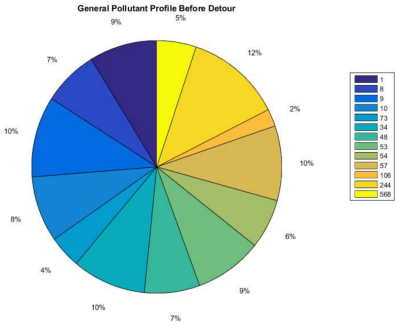
Other factors to be taken into consideration would be the rush hours and traffic jam sport and weather data. The information about rush hours are important as they give insight on the effects of bus timetable reliability and exhaust emission because stopping in traffic lights increases emissions. Also, the weather has its effects on public transportation: on a very cold or rainy day it is expected to have a high demand for buses.

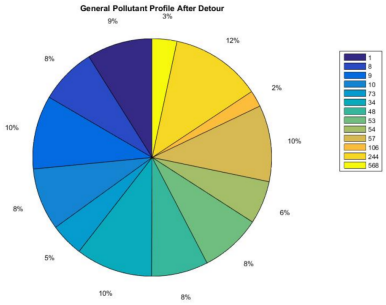
VIII. Reflection on Group Work

We as a group worked really up to the mark as expected from each other. While the course of things, the most interesting part turned out to be when after the initial struggle regarding orienting and being able to comprehend the raw data, we started seeing results in the form of plots and visuals. Having said that, the most challenging part would definitely be scheduling the meetings as well as having to set new objectives for the project. Also, the fact that this course was held only during one period of studies, posed challenges schedule-wise.

before

Route_id	Distance (Km)	Time (Min)	Fuel	Cost	CO g/KWh	HC	NOx	NH3	PM
1	57.4	55.0	19.133	25.447	2.475	0.133	0.394	9.167	0.009
8	76.3	45.0	25.433	33.826	2.025	0.109	0.323	7.5	0.0075
9	95.5	65.0	31.833	42.338	2.925	0.157	0.466	10.833	0.0108
10	43.8	53.0	14.6	19.418	2.385	0.128	0.380	8.833	0.009
14	16.1	26.0	5.367	7.138	1.170	0.063	0.186	4.333	0.004
34	65.4	60.0	21.8	28.994	2.7	0.145	0.43	10.0	0.01
48	36.7	45.0	12.233	16.270	2.025	0.109	0.3225	7.5	0.0075
53	106.5	55.0	35.5	47.215	2.475	0.133	0.394	9.167	0.0092
54	98.6	40.0	32.867	43.713	1.800	0.097	0.287	6.667	0.00667
57	71.6	60.0	23.867	31.743	2.7	0.145	0.43	10.0	0.01
106	10.5	14.0	3.5	4.655	0.630	0.034	0.100	2.333	0.00233
244	111.7	78.0	37.233	49.520	3.510	0.188	0.559	13.0	0.013
568	22.0	32.0	7.333	9.753	1.440	0.0773	0.229	5.333	0.00533





after

Route_id	Distance (Km)	Time (Min)	Fuel	Cost	CO g/KWh	HC	NOx	NH3	PM
1	55.2	51.250	18.400	24.472	2.306	0.124	0.367	8.542	0.009
8	76.0	44.233	25.333	33.693	1.990	0.107	0.317	7.372	0.007
9	91.0	57.75	30.333	40.343	2.599	0.140	0.4139	9.625	0.00963
10	37.9	48.8	12.633	16.802	2.196	0.118	0.350	8.133	0.00813
14	16.1	26.0	5.367	7.138	1.170	0.0628	0.186	4.333	0.00433
34	65.4	60.0	21.8	28.994	2.7	0.145	0.43	10.0	0.01
48	36.4	44.35	12.133	16.137	1.996	0.107	0.318	7.392	0.00739
53	97.6	47.667	32.533	43.269	2.145	0.115	0.342	7.944	0.00794
54	89.7	33.850	29.900	39.767	1.523	0.081	0.243	5.642	0.00564
57	71.6	60.0	23.867	31.742	2.7	0.145	0.43	10.0	0.01
106	10.1	12.9	3.367	4.477	0.581	0.0312	0.0925	2.15	0.00215
244	104.9	71.017	34.967	46.506	3.196	0.172	0.509	11.836	0.0118
568	13.2	19.25	4.400	5.852	0.866	0.047	0.138	3.208	0.00321

Fig. 7. Tabular and Chart Representation of Route Optimization Paramters and Individual Pollutant Profiles

The biggest thing that we learnt from this project is how to work public transport data, identify ideas that could appear up and actually dig out a potent idea that could make things better and more efficient with respect to the data set. We had multiple meetings wherein in the first meeting we all gave in our own individual ideas from where we collectively identified the main research question of this project. We as group members came from very various university programs and that is why we divided the responsibilities between the team members according to their abilities and capacities. It was great to discover that despite of our multi-disciplinary backgrounds we were able to communicate our thoughts and ideas and divide the workload. In addition, it is always good to have people with various knowledge so that you get to learn different types of ways to work and think which will be very important in working life as well as in life in general.

As for the particular contributions, Kennedy Opoku Asare and Gunjan Chandra took care of the route optimisation and pollutant emission ideation and worked on the coding and designing of the same. Aviroop Mukherjee took care if the ideation concerning the bus frequency findings and worked on the segregation to the end-result visualisation of the same. Mira Rauhala took care of the class presentations, theoretical findings, composition and putting together the whole theory of the project.

REFERENCES

[1] Bok, Jinjoo and Kwon, Youngsang. (2016). Comparable Measures of Accessibility to Public Transport Using the General Transit Feed Specification. Sustainability. 8. 224. 10.3390/su8030224.

[2] Farber, Steven; Morang, Melinda Z.; Widener, Michael J. (2014-09-01). "Temporal variability in transit-based accessibility to supermarkets". Applied Geography. 53: 149–159

[3] Fransen, Koos; Neutens, Tijs; Farber, Steven; De Maeyer, Philippe; Deruyter, Greet; Witlox, Frank (2015-10-01). "Identifying public transport gaps using time-dependent accessibility levels". Journal of Transport Geography. 48: 176–187

[4] Wessel, Nate; Allen, Jeff; Farber, Steven (2017-06-01). "Constructing a routable retrospective transit timetable from a real-time vehicle location feed and GTFS". Journal of Transport Geography. 62: 92–97

[5] Philipp Heyken-Soares, Christine L Mumford, Kwabena Amponsah, and Yong Mao. 2019. "An Adaptive Scaled Network for Public Transport Route Optimisation". (2019). journal Public Transport.

[6] Cao, Z., Yuan, Z. and Li, D. Estimation method for a skip-stop operation strategy for urban rail transit in China. J. Mod. Transport. 22, 174–182 (2014).

[7] European emission standards, Wikipedia, accessed 20th Feb, 2020

[8] Extract and Visualize the Results of Multivariate Data Analyses. (n.d.). Retrieved March 1, 2020, from https://rpkgs.datanovia.com/factoextra/index.html

[9] A. Mukherjee, S. Dey and P. Muthu, "Programmed footprint analyser for detection and analysis of diabetic peripheral neuropathy," 2018 IEEE 14th International Colloquium on Signal Processing and Its Applications (CSPA), Batu Feringghi, 2018, pp. 19-24.

[10] Oulun joukkoliikenteen liikennöintidata (GTFS)—Dataportaali. (n.d.). Retrieved March 1, 2020, from https://tinyurl.com/shorter-link-to-oulu-liikenne

[11] Häme, L. (2013). Demand-responsive transport: Models and algorithms. Espoo: Aalto University, School of Science, Department of Mathematics and Systems Analysis.

[12] Maassen, M. A. (2018). Sustainable Business Models: An Imperative in the Strategic Management of Companies and Organizations. Management Dynamics in the Knowledge Economy, 6(2), 323-335.

[13] Vicente, Paula and Reis, Elizabeth. (2018). Ex-regular Users of Public Transport: Their Reasons for Leaving and Returning. Journal of Public Transportation, 21 (2): 101-116.

[14] Abdallah, T. (2017). Sustainable mass transit: Challenges and opportunities in urban public transportation. Amsterdam: Elsevier.