On Providing Healthy Routes: A Case for Fine-Grained Pollution Measurements Using Mobile Sensing

Srinivas Devarakonda, Ruilin Liu, Badri Nath
Department of Computer Science
Rutgers University, USA
e-mail: {skd70, rl475, badri}@cs.rutgers.edu

Abstract— **Extended periods of exposure to air pollution is a health hazard. For commuters, avoiding polluted areas to the extent possible can minimize long-term pollution exposure. Hence, it would be desirable to have an option in navigation systems to choose a route based on the pollution index of the route. However, sufficient data is not available to make an informed choice of routes based on pollution index. This choice is predicated upon availability of ubiquitous pollution measurements along the route segments. In this paper, we present a healthy route recommendation schema that uses spatially and temporally dense pollution measurements to recommend route options that avoid polluted road segments. These healthy routes were evaluated on a neighborhood scale using measurements from vehicular-based mobile sensors. Experiments using data generated by these mobile sensors demonstrate that significant reduction in pollution exposure can be achieved by taking a healthy route instead of the shortest route or the quickest route.**

Keywords- Air Quality; Mobile Sensing; Healthy Routes; Participatory Sensing; Navigation System.

I. INTRODUCTION

According to the factsheet published by the World Health Organization [1], outdoor air pollution in cities and rural areas was estimated to have caused 3.7 million premature deaths worldwide in 2012. Major health effects associated with outdoor air pollution are respiratory and cardiovascular disease, lung cancer, asthma exacerbation and chronic bronchitis. Predominant outdoor airborne pollutants that contribute to these health effects are particulate matter (PM), ozone (O3), nitrogen dioxide, sulfur dioxide and carbon monoxide (CO). The effects of pollution depend on many factors – the concentration of the pollutant, the state of health of the person, the activity of the person and the duration of exposure.

During the 1996 Olympic games in Atlanta, efforts were made to reduce downtown traffic congestion. These efforts resulted in a prolonged reduction in O3 pollution and significantly lower rates of childhood asthma events [7]. The results show that reduction in traffic volume reduces air pollution measurably and improves the health. Another study reveals that utilizing a route away from motorized traffic could reduce bicycle commuter’s exposure to particle number concentrations [5]. It was also seen that the inhaled dosage of pollution not only depends upon the pollution levels but also on the activity of the individual [8].

These studies show that, to reduce the effects of pollution one must either reduce the pollution or make an informed decision to avoid polluted areas. The latter serves as the motivation for our work in building a navigation system that uses spatially and temporally dense (fine-grained) pollution measurements in suggesting a healthy route choice instead of the shortest or the quickest route choices.

The healthy route choice is predicated upon the availability of accurate pollution measurements along the route segments. The difficulty in providing a healthy route choice is the lack of accurate ground-truth pollution measurements. The existing air pollution measurements are available for a non-representative sample of urban areas. Pollution is measured using expensive equipment located at a few select locations. Measurements from these stations are extrapolated over a large area using dispersion models. This data may not truly reflect the ground-truth measurements of pollutants. Localized variations in pollutant concentrations may not be truly represented by the published measurements based on modeled data. Consequently, individual exposure to pollutants on this basis is not fully known. Particularly in urban and metropolitan areas, an individual's daily pollution exposure levels are not truly quantified.

The availability of inexpensive sensors and the ubiquity of reliable cellular bandwidth have provided an impetus towards building and using mobile sensor systems to measure fine-grained pollution concentrations. The increase in the availability and usage of smartphones has seen an increase in the availability of personal pollution sensing devices. This gave rise to a new sensing paradigm – participatory sensing. The mobile sensing systems and personal sensors together are now providing a fine-grained pollution sensing opportunity.

In this paper, we present a navigation schema that generates a healthy route choice using fine-grained pollution measurements. We evaluated this schema on a neighborhood scale. Specifically, we evaluated the following:

- Is there a measurable reduction in pollution exposure on a healthy route as compared to the pollution exposure on the shortest route or the quickest route?
- What is the cost in terms of time and distance if a commuter chooses a healthy route recommendation?
- Is there a temporal variability to a healthy route or does it stay the same every day in a neighborhood?
- How does the choice of the pollutant affect the healthy route recommendation?
- Does the route recommendation differ depending on the mode of transport?

In Section II, we present related work. The details of the sensor systems deployed on public transportation, as well as the sensors used as participatory mobile sensing devices contributing to the fine-grained pollution measurements used in our study are presented in Section III. Details of route generation and the experiments are discussed in Sections IV and V, respectively. Results are presented in Section VI. During this study, there were several ideas that we identified were relevant to the generation of healthy routes but could not be investigated as part of the current work. These are discussed as future work items in Section VII. We provide our conclusions in Section VIII of the paper.

II. RELATED WORK

In an earlier work, Ribeiro et al. developed a healthy route planning system for pedestrians and cyclists to promote less polluting, economical and more equitable modes of transportation [11]. In this work, the healthy routes were calculated based on data collected and estimated through the simulation of noise levels and pollution indices derived from sparse measurement stations using dispersion models. In another study, Beheshtitabar et al. built a system that uses an alternate cost function to predict the bicycle route choice of a commuter based on cost function attributes chosen by the commuter [4]. The authors considered two types of cost function attributes – link-level factors, such as riding surface, riding incline, etc., and route level factors, such as travel time, presence of stop signs, etc. Both groups provided an alternate routing strategy specifically targeting pedestrians and cyclists. In our approach, we developed a system that generates fine-grained pollution measurements and used them to evaluate healthy navigational route choices for cyclists and drivers. We also evaluated the impact of choosing a healthy route in terms of increased time or distance. To the best of our knowledge, this is the first time a healthy route navigation system is developed based on multiple pollutants using fine-grained measurements and evaluated the trade-off between a healthy route choice and the quickest or shortest route choices.

III. HEALTHY ROUTING SYSTEM

In this paper, we evaluate healthy route choices for cyclists and drivers with the aim of minimizing a commuter's long-term pollution exposure. The factors that influence this decision are: the on-road pollution concentrations and the mode of transport. Pollutants that are found in higher concentrations near roads include PM, CO, oxides of nitrogen (NOx), and O3. Fine-grained measurements of these pollutants are required to provide a healthy route choice. The pollution considerations are different for cyclists and drivers. A cyclist is concerned with the ambient air pollution whereas a driver is concerned with the pollution inside the vehicle during their respective commutes. Irrespective of the modes of transport, a user may prefer to minimize exposure to a specific pollutant – for example one may choose to minimize PM exposure or to minimize O3 exposure. So, the considerations in providing a healthy route to a commuter are: on-road pollution measurements, inside the vehicle pollution measurements, pollution inventory as an input in route calculations and route generation.

A. *Pollution Measurements*

Two pollution-sensing models were used to collect the measurements in this study: public transportation sensors that are designed for use on public transportation vehicles to measure ambient pollution and personal mobile sensors that are used as participatory sensing devices to measure pollution inside vehicles. The sensor deployment schema is shown in Fig. 1.

1) *Public Transportation Sensors*

Public transportation sensor units are custom designed and built for use on public transportation infrastructure. We assembled sensor units for use on Rutgers University campus buses. One of the units was mounted on a car for experiments in areas not covered by the campus buses (Fig. 2).

These units are equipped with sensors to measure PM, CO, O3, nitrogen dioxide (NO2), temperature, humidity and pressure (see Table 1 for sensor specifications). The unit is powered by the vehicle battery and starts measuring pollution when the vehicle is powered up and put into service. Global Positioning System (GPS) location data, speed, date and time are attached to each pollution data point. Accumulated data points are uploaded every minute to a server deployed in the cloud. The data upload is done using the data channel of the cellular modem in the unit. These units provide the ambient pollutant measurements relevant for computing the healthy routes applicable to cyclists.

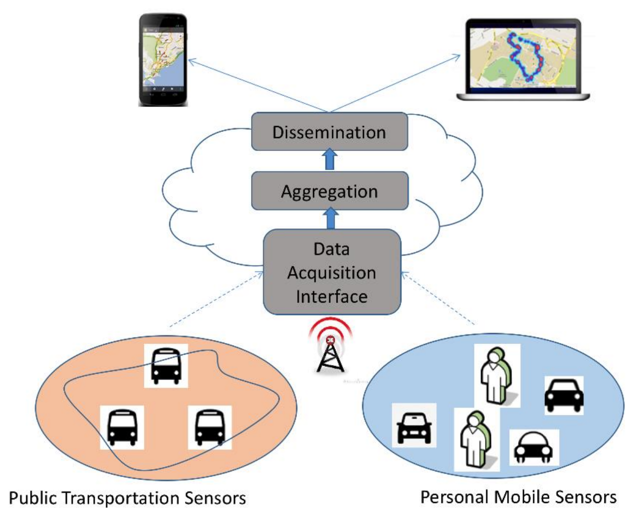

Figure 1. Sensor deployment schema.

Figure 2. Public transportation sensor mounted on a car for experimentation.

Figure 3. Personal mobile sensor mounted in a car near the vent.

TABLE I. PUBLIC TRANSPORTATION SENSOR SPECIFICATION

Sensor	Measures	Range
Figaro TGS5042	CO	0-10000 PPM
MICS 5525	CO	0-1000 PPM
MICS 2710	NO2	0.05 – 5 PPM
MICS 2610	O3	10-1000 PPB
Shinyei PMS1	PM	0.3 µm and larger
Shinyei PPD42NS	PM	1 µm and larger
Sharp GP2Y1010AU0F	PM	0.3 µm – 10 µm

2) *Personal Mobile Sensors*

As part of the participatory sensing model, we use NODE sensors from Variable Technologies for data collection [2]. These sensors connect to the user's IOS or Android smartphone over Bluetooth. We developed applications for IOS and Android platforms to communicate with and manage the NODE devices. GPS location data, speed, date and time are added to the pollution measurements collected by the smartphone application and are uploaded to the cloud server every minute over the phone's data channel. The user may choose to upload this data over Wi-Fi if data bandwidth is a constraint.

For our experiments, we use the NODE device to measure CO, temperature, pressure and humidity (see Table 2 for sensor specifications). The personal mobile sensor, when conveniently mounted inside a vehicle in front of the vent during the user's commute, can measure CO levels in the user's personal space (Fig. 3). In an earlier work, we have shown that the measurements inside the vehicle correlate well with outdoor values [6]. The personal mobile sensors provide the measurements relevant for computing the healthy routes applicable to drivers.

TABLE II. PERSONAL MOBILE SENSOR SPECIFICATION

Sensor	Measures	Range
NODE	CO	0-400PPM

The public transportation sensors and the personal mobile sensors send pollution data to a cloud server. We use Amazon Web Services (AWS) platform [12] to host our server in the cloud. The pollution measurements are stored in a PostgreSQL database [13]. The database has PostGIS [14] extension installed to add support for geographic objects and to allow location queries to be run. Data is post-processed to remove outliers and to apply calibration curves. The resulting dataset consists of pollution measurements, time, GPS location, and speed at the time of the measurement. Information whether the measurement is from inside a car or outside is also stored so that relevant measurement can be used in route calculations for cyclists and drivers. The cloud server provides these pollution measurements to users through a Web portal, as well as through the Android and IOS applications.

IV. ROUTE GENERATION BASED ON POLLUTION MEASUREMENTS

A neighborhood scale road segment graph was created for the Rutgers University College Avenue campus in New Brunswick, New Jersey. This is a directed graph with road intersections as graph nodes and roads as graph edges. Each road segment has a cost associated with it which includes the distance of the segment used for the shortest path calculation, the segment travel time used for the quickest path calculation and the average pollution index used for the healthiest path calculation. In our schema, we used PM and CO measurements for the pollution exposure values so that the healthy route can be calculated based on average PM or CO concentrations per road segment.

 ISBN: 978-1-61208-598-2

For each road segment, we stored the segment length, travel time, average PM and average CO concentrations per unit time. These values are obtained directly from the pollution inventory generated by the public transportation sensors and the personal mobile sensors. Segment lengths and segment travel times are used to calculate the shortest paths and the quickest paths, respectively. To calculate the healthiest path, the pollution load on each segment is required. A segment's pollution load at any given time is calculated based on the average pollution on the segment per unit time and the travel time on the segment.

We implemented Dijkstra lowest cost path algorithm with segment distance, segment travel time and segment pollution index as the costs to compute the shortest, quickest and healthiest paths, respectively.

V. EXPERIMENTS

Experiments were conducted in Rutgers University College Avenue campus in New Brunswick, New Jersey. Besides Rutgers University campus, the area has a transit train station, downtown businesses and a residential area. The neighborhood has a mix of pedestrians, cyclists and vehicular traffic. Because of this variability in the traffic mix, we chose this neighborhood for our experiments.

We measured CO concentrations inside the car using the personal mobile sensors placed near the car vent. The vent was set to a constant speed throughout the experiments. All the car windows were closed during data collection. PM, CO, NO2 and O3 outside the car were measured using a public transportation sensor mounted on the car. Inside-car measurements were used for route calculations for drivers and outside PM and CO measurements were used for route calculation for cyclists. Vehicular travel time on each segment was obtained from the measurements. Due to lack of transit data for cyclists in the test area, we assumed a constant speed of 10 mph in our route calculations for cyclists. Data was collected over a period of three days.

VI. RESULTS

During the tests, we observed significant temporal and spatial variations in the pollution measurements. Hence, a choice exists in selecting routes that avoid polluted road segments. The CO measurement data from two test runs on two different days shows the spatial and temporal variations in measurements (Fig. 4). On the map in Fig. 4, we display the CO data in three ranges instead of showing the discrete measurements so that spatial variations in pollution can be clearly differentiated. The graph in the inset shows CO measurements in parts per million (PPM) over time. Periodic data identifiers from the x-axis of the plot in the inset are marked on the map to show the corresponding location of the measurements.

The PM measurements were relatively steady due to a series of snowstorms in the area (Fig. 5). We believe that the few high readings observed were due to the re-entrained deicing treatment on the roads. Even though these readings had very little variation, we observed that average segment level PM concentrations showed variability so we went ahead and used these measurements in our healthy route evaluation. On the map in Fig. 5, we display the PM data in three ranges instead of showing the discrete measurements so that spatial variations in pollution can be clearly differentiated. The graph in the inset shows PM measurements in milligrams per cubic meter (mg/cum) over time. Periodic data identifiers from the x-axis of the plot in the inset are marked on the map to show the corresponding location of the measurements.

A. Route choice and Pollution

Using the inside-car CO measurements, we computed the shortest, quickest and healthiest paths between an arbitrary

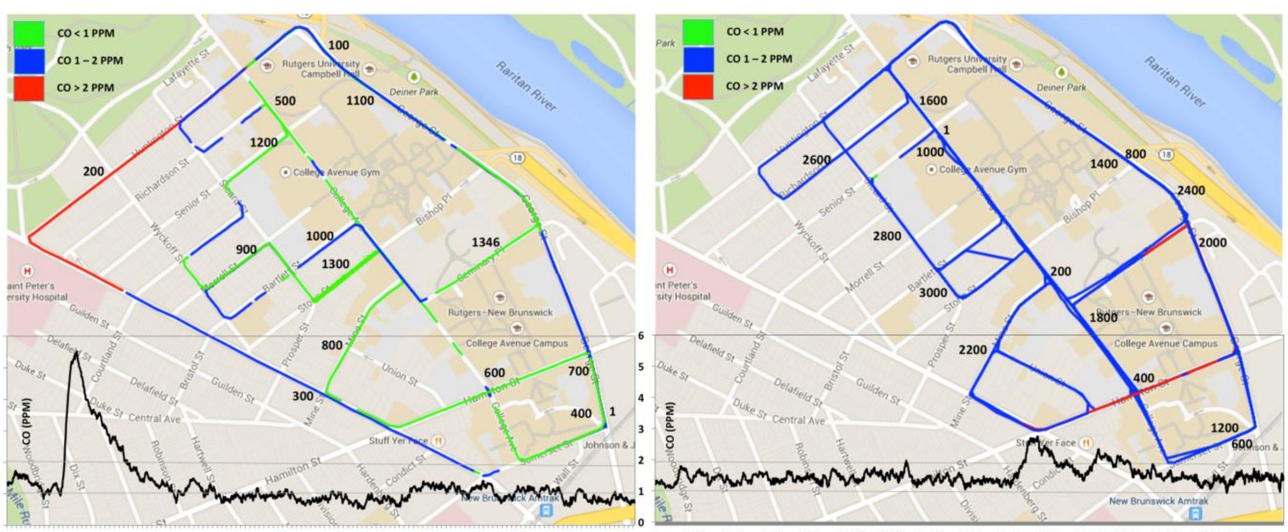

Figure 4. CO measurements in College Avenue campus during two test runs on two different days. Inset in each of the maps shows the time series plot of CO measurements. Periodic X-axis values are marked on map to correlate measurements shown in inset to corresponding location.

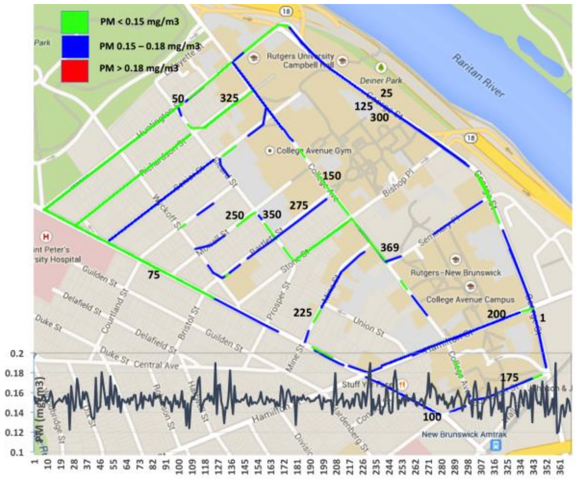

Figure 5. PM measurements in College Avenue Campus. Inset shows the time series plot of PM measurements. Periodic X-axis values are marked on map to correlate measurements shown in inset to corresponding location.

set of origin and destination (O/D) pairs. The exposure cost for the three route choices for 30 O/D pairs was calculated and plotted as scatter plot (Fig. 6). The CO exposure in the case of the shortest path and the quickest path has always been more than the exposure on a healthy route. The exposure is higher on the shortest path as compared to the exposure on the quickest path. It can be seen from Fig. 6 that, the healthy route's CO exposure of 10.19 PPM is much less than the corresponding exposure values of 30.2 PPM for the shortest path (data point marked as 1) and 17.0 PPM for the quickest path (data point marked as 2).

We quantified the percentage increase in the cumulative CO exposure on the shortest path and the quickest path as compared to the healthiest path for the 30 O/D pairs we evaluated (Fig. 7). It is seen that, on average there is an

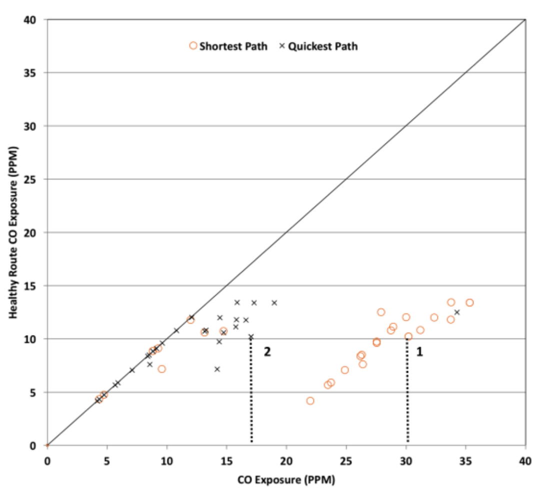

Figure 6. CO exposure on healthy route vs. CO exposure on shortest and quickest route. X-axis values for points 1 and 2 show the CO exposure values for shortest and quickest paths for a healthy route's CO exposure of 10.19 PPM.

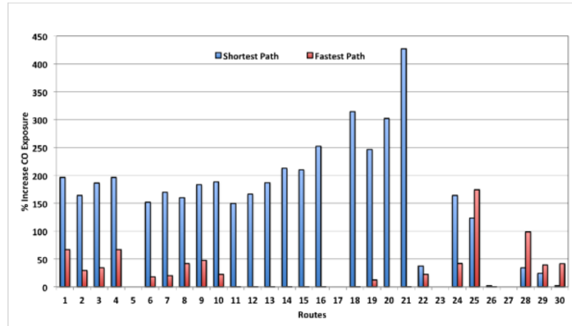

Figure 7. Increase in cumulative CO exposure in shortest and quickest paths over healthiest path in 30 origin-destination pairs.

increase of 160% in CO exposure if a shortest path is chosen. There was one instance where a 400% increase in CO exposure was recorded. In the case of the quickest route, the increase in CO exposure is not as predominant as in the case of the shortest path. On average, there is an increase of 45% in CO exposure when the quickest route is chosen. There were several instances of the quickest route where only a marginal increase in CO exposure was observed over the corresponding CO exposures on the healthiest route.

B. *Cost of Healthy Route*

In this section, we discuss the cost of the healthy route choice in terms of increase in time and distance of travel. We will use the same data set from the 30 O/D pairs we used before.

It is observed that in all the 30 O/D pairs we evaluated, there has been an increase in the distance and time of travel (Fig. 8). On average, there is an increase of about 8-9% in the distance travelled and about 30% in the time of travel. The maximum increase in the distance is about 50% and in time it is about 85%.

There is an additional cost to a healthy route choice due to the increased time and distance of travel. However, it is the choice of an individual – whether the benefits of a healthy route outweigh the additional costs of increased travel time and distance. There could be another choice provided to the user by the routing implementation to accept a threshold for the increased travel time and/or distance, beyond which the healthy route recommendation is not provided.

C. *Healthy Route – Temporal Dependency*

Pollution index for a route varies over time. To establish this temporal dependency, we calculated the healthy routes for the same O/D pairs (Fig. 9) using the CO measurements inside the car taken over two days.

It is seen that the healthy routes are different on the two days. Hence, static profiles of routes are not sufficient. Tests with different O/D pairs yielded similar results. The quickest routes were also different but we have not shown them in Fig. 9 for clarity.

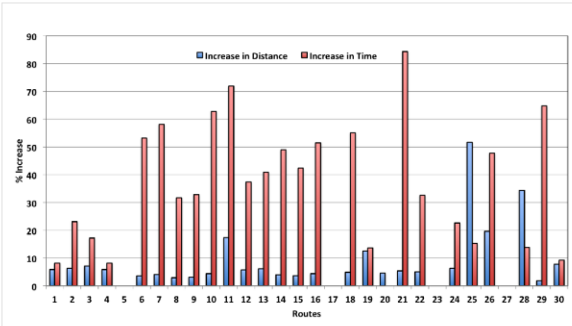

Figure 8. Percentage increase in distance and time on a healthy route corresponding to the shortest and quickest routes.

Figure 9. Healthiest route recommendations on two different days. Shortest route is also shown. Quickest route is omitted for clarity.

D. Healthy Route – Dependency on Pollutant

A user may prefer to base their healthy route choice on a specific pollutant. We evaluated this scenario using pollution measurements outside the car using PM and CO measurements for a cyclist. Currently, we only have CO measurements inside the car, so we used outside PM and CO measurements to evaluate this scenario.

Fig. 10 shows the route recommendation for a cyclist based on CO and PM measurements. We show the results for one O/D pair only, even though we observed similar results for other O/D pairs. A similar recommendation could not be evaluated for drivers due to lack of fine-grained measurements for pollutants other than CO.

E. Healthy Route – Mode of Transport

The healthy route recommendations for cyclists and drivers did not show any variation in the routes. The pollution exposure depends on the amount of time spent on a road segment. So, we think that the use of a constant speed of 10 mph in our calculations may be affecting the route calculations. Additional evaluation needs to be done when transit data becomes available for cyclists. Alternatively, we

Figure 10. Distinct healthy route recommendation for cyclists based on based on different pollutants.

need to choose a test location where transit data for cyclists is available.

F. Trade Off – Healthy Route and its Cost

The left part of Fig. 11 shows the comparison between the increase in travel time on a healthy route and the increase in pollution on the quickest route. There is no clear distinction to choose one option over the other. A threshold function can help wherein an increase in travel time above a threshold on a healthy route - chooses the healthy route over the quickest route.

Fig. 11 shows the tradeoff between pollution exposure, distance and time. The right side of Fig. 11 shows the increase in pollution exposure along the shortest path plotted along with the increase in distance on the healthiest path. The proportional increase in pollution exposure far out-weighs the increase in travel distance on a healthy route.

VII. Discussion and Future Work

In this paper, we presented fine-grained pollution data as a choice in route selection. However, we have not discussed validation methods for our healthy route approach. Validation needed a large team divided into producers of fine grained pollution measurements and consumers of healthy route recommendations with a view to validate the correctness of the route recommendations. We did not have a large team at our disposal during these tests to conduct validation of our approach during our experimentation. However, during our tests we observed that the segment level pollution load and segment travel time remained relatively constant for a period of about 15-20 minutes. So, we use this to observe that our route recommendations are at least valid for this duration. A rigorous validation of our route recommendation will be taken up as part of a future work using data from sensors mounted on campus buses and participatory sensing using student groups.

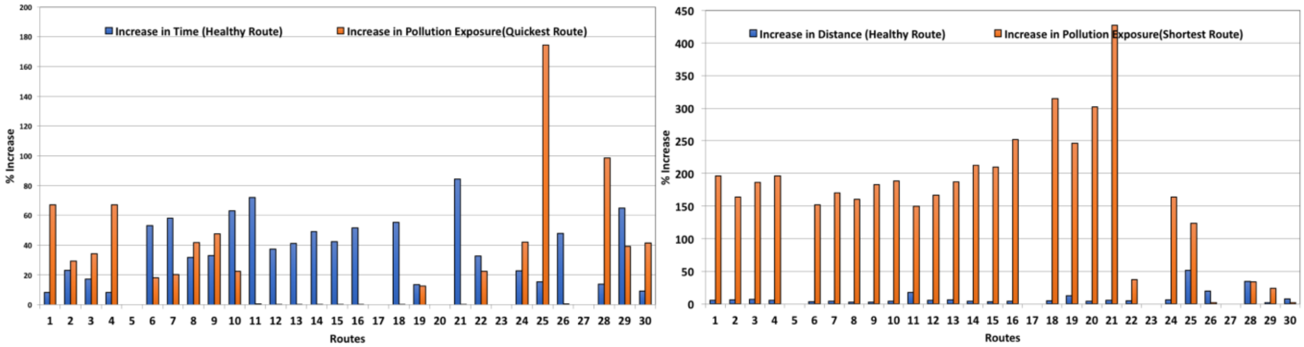

Figure 11. (Left) Percent increase in travel time for healthy route vs. percent increase in pollution for quickest route. (Right) Percent increase in distance for healthy route vs. percent increase in pollution for shortest route.

The healthy route recommendation depends on the availability of fine-grained pollution measurements. For areas with no pollution data, the available measurements in the near vicinity can be used to extrapolate on a micro scale based on the approach discussed by Ulyanik et al., [10].

Participatory sensing may not be contributing to fine-grained measurements in all areas. Using the model developed by Hudda et al., [9] outside the car pollution measurements can be used to derive the pollution exposure inside the vehicle. The inside-car measurements depend upon the comfort settings in the vehicle (ventilation settings, opening of windows, etc.) and the age of the vehicle. The model caters to all these variables. Similarly, it must be possible to derive outside car measurements based on measurements from inside the car. However, this needs further work in the future.

A healthy route recommendation need not be a static route generated at the start of the trip. Current navigation systems are capable of dynamic route updates based on road and traffic conditions. With the availability of fine-grained pollution measurements, it is trivial to reuse the existing dynamic route update capability to healthy routes as well.

Our study has been confined to a neighborhood scale due to lack of fine-grained pollution data. Even though our work is based on neighborhood scale data, we feel it can be easily extended as fine-grained measurements become available over a larger area. Consequently, the demand on existing navigation systems is additional storage to store per segment pollution attributes and the additional computation to generate healthy routes.

The sensitivity of healthy routes to seasonal variations needs additional study. The dependency on local weather conditions and its influence on the healthy route choice will require additional work in the future.

We did not include pedestrians in the current study. A pedestrian’s pollution exposure depends on the mobility patterns and wait patterns in an urban setting. A separate study is required to provide similar insight into a pedestrian’s pollution exposure during a day.

Once a healthy route recommendation is made, it will be interesting to study how users behave. For example, given a healthy route choice, how many cyclists and drivers adopt it at the cost of increased distance or time?

VIII. Conclusion

This paper presents a navigation system to provide healthy route recommendations using fine-grained pollution measurements contributed by participatory mobile sensors and public transportation sensors. We evaluated this model on a neighborhood scale and showed that the healthy route provides significant improvement in the pollution load on an individual when compared to the pollution loads on the shortest route or the quickest route. In addition, we also show that fine-grained measurements are essential for providing the healthy route recommendations daily. The healthy route recommendation need not be based on a single pollutant but can be a choice of pollutants. We evaluated this model for two modes of transportation, namely bicycles and automobiles.

Acknowledgment

This work has been supported in part by National Science Foundation (NSF) Grant CNS-1111811.

References

[1] WHO factsheet http://www.who.int/mediacentre/factsheets/fs313/en/ [retrieved: September, 2017]

[2] NODE sensors http://www.variableinc.com/gas-sensing [retrieved: September, 2017]

[3] D. Allemann and M. Raubal, "Towards health-optimal routing in urban areas," In Proc. Second ACM SIGSPATIAL International Workshop on the Use of GIS in Public Health, pp. 56-59, 2013.

[4] E. Beheshtitabar, S. Aguilar Ríos, D. König-Hollerwöger, Z. Svatý and C. Rydergren, "Route choice modeling for bicycle trips," International Journal for Traffic and Transport Engineering 4, pp. 194-209, 2014.

[5] T. Cole-Hunter et al., "Utility of an alternative bicycle commute route of lower proximity to motorised traffic in decreasing exposure to ultra-fine particles, respiratory symptoms and airway inflammation–a structured exposure experiment," Environmental health 12, no. 1, 2013.

[6] S. Devarakonda et al., “Real-time air quality monitoring through mobile sensing in metropolitan areas,” In Proc. Second ACM SIGKDD International Workshop on Urban Computing, 2013.

[7] M. Friedman, K. E. Powell, L. Hutwagner, L. M. Graham, and W. Gerald Teague, "Impact of changes in transportation and commuting behaviors during the 1996 Summer Olympic Games in Atlanta on air quality and childhood asthma," Jama 285, no. 7, pp. 897-905, 2001.

[8] K. Hu, T. Davison, A. Rahman, and V. Sivaraman, "Air Pollution Exposure Estimation and Finding Association with Human Activity using Wearable Sensor Network," In Proc. MLSDA 2014 2nd Workshop on Machine Learning for Sensory Data Analysis, 2014.

[9] N. Hudda and S. A. Fruin, "Models for predicting the ratio of particulate pollutant concentrations inside vehicles to roadways," Environmental science & technology 47.19, pp. 11048-11055, 2013.

[10] I. Uyanik, A. Khatri, P. Tsiamyrtzis and I. Pavlidis, “Design and Usage of an Ozone Mapping App,” In Proc. Wireless Health 2014 on National Institutes of Health (WH '14). ACM, New York, NY, USA, DOI=10.1145/2668883.2668885 http://doi.acm.org/10.1145 /2668883.2668885

[11] P. Ribeiro and J. F. G. Mendes, "Route planning for soft modes of transport–Healthy routes," In 17th International Conference on Urban Transport and the Environment, pp. 677-688, 2011.

[12] Amazon Web Services details https://aws.amazon.com/ec2/details/ [retrieved: October, 2017]

[13] PostgreSQL database details https://www.postgresql.org [retrieved: October, 2017]

[14] PostGIS spatial and geographic objects for PostgreSQL database http://postgis.net [retrieved: October, 2017]