Collaborative Sensing Platform for Eco Routing and Environmental Monitoring

Markus Duchon[1], Kevin Wiesner[2], Alexander Müller[2], and Claudia Linnhoff-Popien[2]

[1] Siemens Corporate Technology, CT T DE IT 1
Otto-Hahn-Ring 6, 80200 Munich, Germany
markus.duchon.ext@siemens.com
[2] Ludwig-Maximilian-University Munich, Mobile and Distibuted Systems
Oettingenstr. 67, 82358 Munich, Germany
{kevin.wiesner,linhoff}@ifi.lmu.de,muelleral@cip.ifi.lmu.de

Abstract. During the past decades, ecological awareness has been steadily gaining popularity. Especially in so called Megacities, the burden caused by air pollution is very high, as millions of people live together in a localized manner. To be aware of the current pollution status, selective measuring stations where deployed in the past. The idea of this work is to enable the masses to participate in obtaining and using their own measurements, e.g. with future generations of mobile phones that are equipped with adequate sensors. The proposed platform allows for a high resolution environmental monitoring and provides additional services such as *Eco Routing* or visualization. Furthermore, we will present the results of the platform's performance as well as a comparison between the traditional (shortest/fastest) routing and the novel (shortest/fastest) *Eco Routing* approach.

Key words: collaborative sensing, monitoring platform, Eco Routing

1 Introduction

During the past decades, environmental problems have been gaining more and more attention, especially in highly populated areas. In Europe, many administrations raised the bar to observe certain specified pollution limits within urban areas. For this purpose, environmental zones have been defined in most of the bigger cities in Germany. The entry to these zones is restricted to certain types of vehicles, e.g. those that are equipped with state of the art technologies to minimize exhaust emissions. But, according to a preliminary study[1] of the German Federal Environment Agency (UBA), the values of fine particulate matter in 2011 exceeded the average of the last four years, although these zones have been established. Besides traffic, also combustion processes in industry and private households attract increasing attention in the area of environmental observation.

[1] http://www.umweltdaten.de/publikationen/fpdf-l/4211.pdf, last access 01.02.2012

Nowadays, measurement systems only allow for a selective monitoring of environmental aspects for certain places, which will not be sufficient according to comprehensive monitoring tasks. In this context, a high-resolution and real-time environmental monitoring platform is highly beneficial in order to determine problematic areas and to plan appropriate counter-measures. This information is not only valuable for administrative purposes, but also for the general public, e.g. services that provide up-to-date information about healthy regions to live or routes to get from one point to another.

Sensor networks have improved and leveraged the environmental observation especially in rough, dangerous and preservable areas. In this context, also mobile or people-centric sensing becomes increasingly popular. This progress led to cheaper production, improved measuring techniques, and small sized measuring units. We believe that in the near future mobile devices, like smart phones, will be equipped with environmental sensing units and therefore this information will be available ubiquitously.

In this work we present a *Collaborative Sensing Platform for Eco Routing and Environmental Monitoring*. On a related note, the necessary information is collected by mobile as well as stationary sensing units and will be processed at the platform to preserve high resolution and up-to-date information about the current air quality with the focus on urban environments. Furthermore, the platform is capable to provide certain pollution-related services like *Eco Routing* and visualization. The first one should help to improve administrative monitoring and (traffic) planning tasks, whereas the latter one can be advantageous for pedestrians and cyclists to minimize the exposure to air pollution and find more healthy routes through the town.

The remainder of this paper is structured as follows. In Section 2, surveyed related work is outlined before we discuss our approach of the platform in Section 3 including the main requirements, the system architecture as well as the relevant components for *Eco Routing* and visualization. The evaluation of our platform is examined in Section 4. Section 5 concludes the paper and provides an outlook on future work.

2 Related Work

In this section, we will give an overview on related research and projects. Within the first part, platforms to obtain, manage and evaluate sensor data will be addressed. The second part consists of different technologies for modeling environmental pollution data. Finally, two bicycle routing projects will be introduced, which allow for specific routing options.

Eisenman et al. proposed a mobile sensing system, `BikeNet` [1], for mapping cyclist experience. In this project, sensors are directly embedded in bicycles and gather information about rides and according environmental information. The data is either stored offline and will be transferred to the platform later on or it will be transfered via so called `Service Access Points` (SAP) in real-time. The platform can be used to share cycling related data, like favorite routes, as well

as data of more general interest, like pollution data. With regard to our system the proposed approach concentrates only on the pollution of the actual driven routes and therefore does not allow for high resolution monitoring capabilities.

`CarTel` [2] is a mobile sensor computing system, proposed by Hull et al., which was designed to collect, process, deliver and visualize data. The information is obtained from sensors located on cars. The gathered data is preprocessed on a so called `CarTel node`, before it is delivered to a central backbone for further analysis. The system provides a query-oriented programming interface and can handle large amounts of heterogeneous sensor data. The communication of the nodes primarily relies on opportunistic wireless connectivity to the Internet or to other nodes. `CarTel` applications use the collected data e.g. to optimize the travel distance of commuters or to localize areas with a high traffic density based on a low average speed. Nevertheless, the environmental information only addresses the emissions produced by the vehicle itself and not the pollutant concentration in the surroundings as proposed in our work.

SenseWeb [3], presented by Luo et al., is an open and scalable infrastructure for sharing and geocentric exploration of sensor data streams. It allows for indexing and caching of sensor data as well as processing spatio-temporal queries on real-time and historic data. Furthermore, aggregation and visualization of the obtained information is provided by a web-interface. In this work only stationary sensors or wireless sensor networks are utilized, whereas changes in position and density are not considered so far.

In the domain of statistics, there exist several approaches for data preparation and prediction which are often used in medical studies, e.g. the influence of air pollution on live expectation. In this context, several models already exist to describe phenomena like air pollution. Jerrett et al. compared five different classes of models in their work [4]: proximity, interpolation, regression, dispersion and meteorological models. These models were also subject to certain improvements and specializations within several publications [5, 6, 7, 8, 9], which will not be explained in more detail due to space limitations.

Cyclevancouver[2] is a a free web-based service that allows for bicycle route planning in the city of Vancouver (Canada). The objective is to encourage citizens in cycling. The service provides optimized routes especially for bikes and allows for additional options like low ascent, abundant vegetation, and low air pollution. However, the pollution-based route generation utilizes a statical model of the year 2003 that was only calculated once and therefore does not consider the current environmental situation.

Fahrradies[3] is also a web-based service which offers an interactive navigation system for cycle tours within the Oldenburger Münsterland (a region in North Rhine-Westphalia, Germany) [10]. The multifaceted routing options are very flexible and the navigation engine also allows for several intermediate stops

[2] http://cyclevancouver.ubc.ca, last access 01.02.2012
[3] http://www.fahrradies.net/, last access 01.02.2012

(sights, shops, etc.). By using the *green route* option only the environmental surroundings in terms of the vegetation and development are considered.

3 Collaborative Sensing Platform

In this section, we will introduce the requirements, a collaborative platform for environmental monitoring and the provisioning of related services must fulfill. In this context the term *collaborative* refers to the collaboration of users in collecting measurements. Afterwards, we present the architecture and explain the main components in more detail, whereby the basis was the final thesis of Alexander Müller [11]. Finally, we outline a service called *Eco Routing*, which utilizes the obtained and preprocessed sensor data to allow for healthier and less polluted routes. The second service concentrates on the visualization of pollution data for different purposes.

3.1 Requirements

To enable high resolution measurements, not only stationary sensors will be considered, but also mobile ones. In order to use the provided data in a geospatial context, information about the location must be obtained. The location of stationary sensors is known by definition, whereas mobile sensor devices require additional positioning capabilities, e.g. GPS receiver. For environmental monitoring in the context of traffic and transportation, suitable sensors for NO_2, CO, CO_2, etc. should be used.

Scalability and robustness are also important requirements to preserve high resolution pollution information, wherefore large amounts of sensory data must be handled and processed. In this context, a viable storage and update structure is essential. Moreover, the provided information certainly is subject to wild fluctuations. These fluctuations do not only arise from the quality of different measuring devices, but also by the dimensions of space and time. During a single day, there will be more crowded areas leading to more measurements and a higher density as well as there will be certain areas that are less crowded, leading to sparseness of measurements. Therefore, suitable techniques are necessary to allow for an equalization of the data and provide appropriate information about the degree of pollution at a certain place and time.

Besides the information basis, also the service provisioning demands certain needs. The most important one is to map the pollution information onto the corresponding road segments as weights that can be used during the routing process. Additional information about the street network itself is essential, which should allow for routing and can be easily separated into the mentioned road segments.

3.2 Architecture

Before we explain the details of the proposed system to enable environmental monitoring including an *Eco Routing* and a visualization service, we will give

an overview of the architecture and the involved components. We decided for a layered approach, which is illustrated in Figure 1.

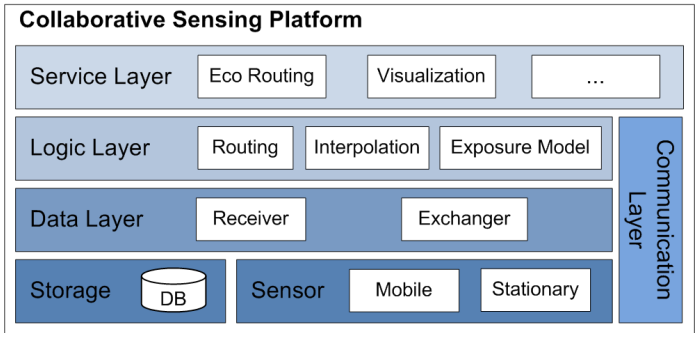


Fig. 1: System architecture overview

The bottom layer consists of two parts, namely sensor and storage. The sensor sector represents the mobile and stationary sensor devices, which gather information about the environment in form of pollution data as mentioned in the requirements section. The storage sector is responsible for storing the persistent information like e.g. the street network, results received by the upper layers, or as archive for historical purposes.

The data layer mainly receives the sensor data, adds appropriate timestamps, does minor plausibility checks and partitions the data samples according to their location before it is passed to the storage sector for further processing. Another task of this layer is to exchange data from the lower layer to the upper ones and vice versa. This becomes important when the platform is deployed in a distributed manner, whereby different communication channels can be used to exchange data with spatially separated components.

Within the logic layer, there are three main components, which are responsible for routing interpolation, and model generation. These three components will be explained in more detail later on. The vertical communication layer provides different communication standards to allow for an independent data exchange using various communication channels. Thereby, also a distributed interaction between the layers (storage, data, logic) is feasible.

The top layer represents services, which can be provided by the platform. Examples are a service, called *Eco Routing*, which allows for route calculations based on the current environmental situation in order to minimize the exposure to air pollution or a simple visualization or monitoring service, which can be used to receive up-to-date information about the pollution within a desired area.

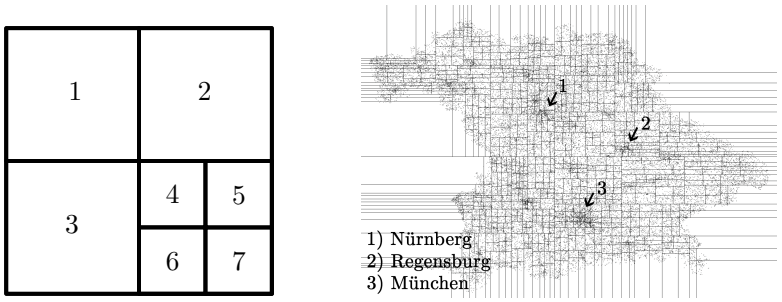
Data Interpolation In the first place, the sensor platform needs to collect and organize all sensor data received from the stationary and mobile sensors. This data includes:

– The GPS coordinates where the measuring took place

– The sensor type, e.g. information about the measuring error
– The measuring values of the air pollutant concentrations

The platform adds a server-side timestamp to the measuring samples on reception and uses a different database for each pollutant afterwards.

As the data is expected to be non-uniformly distributed in space and time, we require a special data structure to organize the samples for further processing. For that reason, we propose a spatial quadtree, where each cell has a maximum capacity. If the capacity of a cell exceeds, it is divided it into four child cells, each containing the same amount of measuring samples inside (0.5 quantile split). Figure 2a illustrates the logical layout of the quadtree cells and Figure 2b an example of a quadtree with quantile splitting method, showing a map of Bavaria with simulated sensor data. The measuring density in cities like Munich, Nuremberg or Regensburg is significantly higher than in rural areas.



(a) Logical layout of the quadtree

(b) Quadtree with simulated data of Bavaria

Fig. 2: Quadtree data structure

For the interpolation part, we utilize a geostatistical interpolation technique called kriging [12, 13], which models a spatial area as a multi-dimensional random process. To ensure real-time capabilities, we are only using the most recent measuring samples. As this method requires the underlying stochastic process to be intrinsically stationary, we use the quadtree cells as window for localized semivariogram estimations and a subsequent local kriging that is based on the nearest neighbor data samples. If a quadtree cell is spatially too large to assume stationarity, we try to subsequently split it using the same quantile based method as mentioned above. If this is not possible, e.g. there are not enough measuring samples inside, we try to use older measuring data. The semivariance of a local area is strongly influenced by geographic features as roads, landscape, buildings, etc., that do not change very quickly. So we are looking for a time slice, where the measuring density was high enough for a solid variogram estimation and use that data for the process. If this also fails, we use the whole cell for the semivariogram estimation and are forced to tolerate a small amount of non-stationarity instead of having no data at all for the according area.

As both, the variogram estimation and the kriging, require the data to be normally distributed, we need to transform the input data first by using a log-transformation. For the semivariogram estimation we use different variogram models, that are put in a candidate list in the first place. One by one, we try to fit them to the measuring data and if the fitting fails, the candidate gets removed from the list. Finally, the candidate with the lowest residuals is selected as model for the subsequent kriging interpolation. The candidate models are: (i) Spherical model, (ii) Exponential model, (iii) Gaussian model, and (iv) Matern model with M. Stein's parameterization.

We support two different types of kriging, each having a different objective. The first approach uses localized block kriging to estimate a raster map of the air pollutant concentration over a large area. This way we can gather information not only about the street network, but also about the neighborhood, i.e. residual areas, parks, etc. Using block kriging additionally smooths the interpolation result, thus reducing the influence of potential outliers in the input data.

The second approach is to use a localized ordinary kriging to interpolate fixed points on the street network, which are determined using a geospatial data set. For each road we define a maximum segment length s. We split the road into small segments with maximum length s by dividing the total length l of the road. The segment count is given by $ceil(\frac{l}{s})$. We apply the ordinary kriging to the end points of the road segments that have been determined by the previous step. The second approach can only give information about the air pollution that is present directly on the street network itself. It does not offer further information about adjacent areas.

Both kriging types can account for known measuring errors by decomposing the variogram's nugget into a model-based semivariance and an error-based one. The interpolation result is then back-transformed to the original distribution.

Exposure Model After the interpolation of the pollutant concentration, we need to apply an exposure model to calculate an edge weight for the routing graph for each road. We assume that the total exposure of a road user is basically proportional to the retention time t and the concentration k on the road. Using this simplifying assumption, we can calculate the total exposure E using

$$E = \int_0^T k(t)\ \mathrm{d}t$$

for a continuous function k or

$$E = \sum_i \Delta t_i \cdot k_i$$

for a discrete case, which can easily be derived from the interpolated data. The process is shown in Figure 3. The exposure E is equal to the total area of the gray surface.

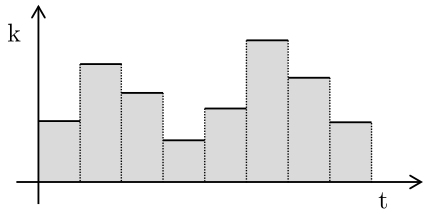


Fig. 3: Function $E_s(s)$ as discrete function

The calculation of the total exposure E is split into two separate steps. We interpret the road as a one-dimensional function from 0 (start point) to l (end point), so we can define a function

$$E_s : [0, l] \mapsto \mathbb{R} \ ,$$

which gives us the pollutant concentration for all points on the road. We can calculate E_s using the discrete data gathered by the interpolation step. With the raster interpolation, we can intersect the road with the raster cells to get an ordered list of road segments, each having a particular concentration value. In the point interpolation case, we already precomputed the road segments, which were used to determine the interpolation point coordinates. Using these points, we define new segments, each having one interpolation point in the center, which represents the concentration value for the whole segment. In doing so, we can interpret the segments and the associated pollutant concentration as a discrete representation of the function $E_s(s)$.

To incorporate the time into the model, we define a function

$$E_t : [0, T] \mapsto \mathbb{R} \ , E_t = E_s \circ s(t) \ ,$$

with T being the total driving time on that road and $s(t)$ giving the location of the vehicle at time t.

T and $s(t)$ are dependent of the velocity v on that specific road. In the simplest case, we can assume a constant speed v_{const}, that is either determined by an explicit speed limit or an implicit average speed for that kind of road. In some advanced cases, e.g. if we use floating-car data to estimate the actual speed, the velocity v can be given as coordinate-dependent function $v(s)$. To estimate E_t, we need to calculate the function $s(t)$, which can be derived from $v(s)$ by:

$$\mathrm{d}t = \frac{\mathrm{d}s}{v(s)} \quad \Rightarrow \quad t(s) = t_0 + \int_0^s \frac{\mathrm{d}\bar{s}}{\mathrm{v}(\bar{s})} \quad \Rightarrow \quad s(t) = t^{-1}(s)$$

And in the case of a constant velocity: $s(t) = v_{const} \cdot t$.

Eco Routing The generated model, explained in the previous section, is the basis for the routing component. Thereby, we propose an approach that uses the exposure value calculated beforehand, which is static when the routing process

starts. Hence, no dynamic edge weights will be used and in case of a long distance route, it should be recalculated after defined timeouts. But, as the routing service mainly addresses the inner-city traffic the occurrences of long distance travels is quite rare.

In general, routing requires a weighting function for the edges in order to determine the costs for a specific way. Classical routing strategies like the *shortest path* metric utilize the distance as weight, whereas the *fastest path* uses the allowed speed in combination with the distance to utilize the driving time as edge weight. Later on, these two approaches will be used for comparison during our evaluation in Section 4.3.

The *fastest eco path* metric is directly based on the exposure value E, as it is directly proportional to the driving time as it is a product of the driving time and the exposure to pollution for the according road segment. In case a faster vehicle drives along the road segment, the exposure will be lower than for a slower one.

The *shortest eco path* metric utilizes the calculated exposure value E only as basis for the edge weight. Hence, E will be normalized by dividing it by the driving time of the corresponding road segment. The interim value represents the average pollutant concentration, which is multiplied with the length l of the road segment to receive the final edge weight ω:

$$\omega = \frac{E \times l}{t}$$

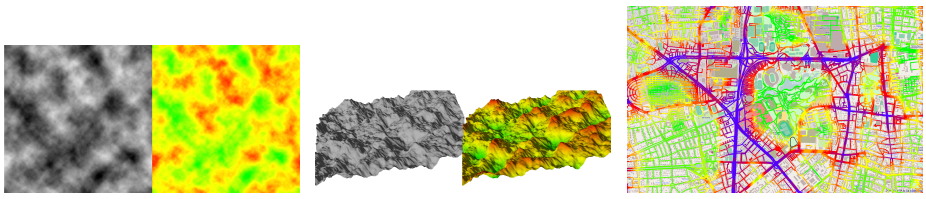
Its advantage is that ω is independent of the driving time and therefore also independent of the vehicle's speed and the speed limit on that road.

Within our implementation, we utilize the A* routing algorithm, which was first mentioned by Hart et al. in [14]. The algorithm applies a heuristic that guesses the minimal distance between two vertices within a weighted graph $G = (V, E, \omega)$ consisting of vertices V and edges E and according edge weights ω. Vertices in a street network carry additional geographic information, which can be used to calculate the beeline as heuristic for A*, which also fulfills the triangle inequality in our case.

Visualization The calculated pollutant concentration can be visualized according to the desired purpose. In the following, we present 3 methods, whereby the first two require raster data and the latter one operates on the pollutant concentration on the street network.

`Pollution data as Bitmap:` The visualization of raster information is often realized by using bitmaps. In our example, the interpolated values are mapped onto the value range $[0, N_{max}]$ with N_i as brightness value, resulting in a monochrome bitmap, where black represents areas with low pollution and white areas with a high pollution value. It is also possible to colorize the bitmap in a second step. Figure 4a illustrates monochrome and colorized pollution bitmaps using red for areas with a higher pollution and green for areas with a lower one.

`Pollution data as 3D Model:` In comparison to the bitmap, the generation of a 3D model is much more computation-intensive. Thereby, the pollution



(a) monochrome/colored Bitmap (b) monochrome/colored 3D Model (c) Street Network

Fig. 4: Pollution visualization as Bitmap, 3D-Model, and Street Network

value will be used as height information or the z-axis respectively. For rendering purposes a Point Cloud as well as a Mesh is reasonable. Optionally, a bitmap of the aerial perspective can be used as texture. Figure 4b shows the 3D model generated with the same pollution values as the bitmap before.

`Pollution mapping on Street Network:` In case the exposure values were calculated for the street network, also a direct mapping on the according road segments is possible. Thereby, the color information is directly derived from the current values. One option is to use the exposure value E and the other one is the utilization of the exposure function E_s. In the first case, the exposure value must be normalized onto the length of the road segment, otherwise longer road segments with an equal exposure value will have a higher (color-)value than shorter ones. For both cases, similar to the bitmap approach, the exposure values must be initially mapped onto an according (color-)value range. Figure 4c shows a colorized part of the street network close to the Olympic Park of Munich, whereby blue indicates the highest pollution and green the lowest.

4 Evaluation and Results

In this section, we present the evaluations conducted for our platform for routing and environmental monitoring. Therefore, we first describe how we obtained the sensor data needed for our evaluations, followed by the analysis of our interpolation methods and the *Eco Routing* approach.

4.1 Sensor Input

The proposed platform assumes that large-scale sensor information from distributed mobile sensors is available. Sensors could be integrated in user devices, such as smartphones, or mobile vehicles, such as cars or bicycles. However, since this kind of sensors are not yet available in large quantities, we had to generate realistic sensor input data for the following evaluations.

For this purpose, we used *Land-use regression* (LUR). Since LUR requires a broad data basis, which was not available in our case, we applied an adapted model [6] that integrates information about the street network, traffic density and absolute altitude. The street network information is extracted from *OpenStreetMap* project, the traffic density is estimated based on the type of road and

the amount of lanes, and the absolute altitude was determined with the ASTER GDEM dataset[4].

The following regression model is used for generating synthetic NO_2 data:

$$\mu(NO_2) = \beta_0 + \beta_1 \cdot (15 \cdot TD_{0-40} + TD_{40-300}) + \beta_2 \cdot log_{10}(Alt)$$

TD_{n-m} denotes the traffic density in the area between n and m meters around the measuring point. The traffic density for a road segment is calculated by multiplying its length with a factor for the type of road. TD_{n-m} is the sum of densities for all road segments in the specified area. β_0-β_3 are weighting coefficients, which are mostly based on values used by Briggs et al. [7] in Huddersfield (UK), though β_1 was adapted to receive feasible results in our setting. For our simulations, we used the following values: $\beta_0 = 38.52$, $\beta_1 = 0.0003705$ and $\beta_2 = -5.673$.

4.2 Data Interpolation

As mentioned in Section 3.2, the proposed platform supports two interpolation methods: point and raster interpolation. We evaluated those methods by comparing the difference of exposure obtained with raster interpolation (exp_{raster}) and point interpolation (exp_{point}) for each street segment and determining the ratio $q = \frac{exp_{raster}}{exp_{point}}$. The analysis was conducted for various maximum distances of measuring points: 15, 25, 35, 45, 55 and 75 meters.

Distance d [m]	Min	Max	μ	σ	Segment length [m]
15	0.654	1.558	0.9995	0.0120	12.03
25	0.657	1.558	0.9995	0.0120	18.09
35	0.659	1.558	0.9993	0.0122	23.28
45	0.662	1.558	0.9993	0.0123	27.91
55	0.662	1.558	0.9992	0.0125	31.92
75	0.660	1.558	0.9991	0.0130	35.47

Table 1: Minimum, maximum, mean and standard deviation of q

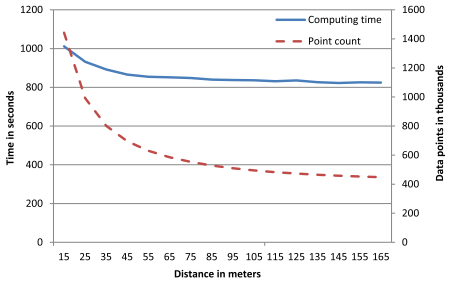
The results for q are listed in Table 1. It can been seen that the mean value (μ) for q is in all case very close to the ideal value of 1.0, which shows that both interpolation methods perform almost similar. The standard deviation (σ) is also very low and ranges from 0.0120 (for $d = 15$) to 0.0130 (for $d = 75$), so that for 99% of the tested roads both values deviate less than 5%.

We also analyzed the performance of both approaches with a standard PC from the premium price segment. Even though the absolute numbers will be different on other computers, the relative results indicate the performance of the approaches for different settings.

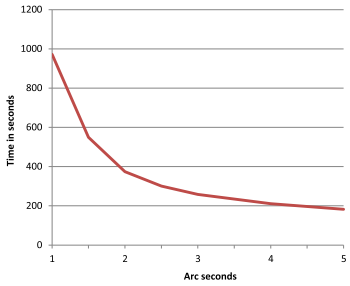
[4] http://www.gdem.aster.ersdac.or.jp/, last access 01.02.2012

Figure 5a shows the time measurement results for point interpolation with varying distances between measuring points. It can be observed, that the duration rapidly decreases with greater separation of measuring points. In our setting, a more fine-grained interpolation than 45m should be avoided, as the run time sharply increases for lower distances.

The raster interpolation is only dependent on the raster size. Thus, we examined the time duration for this kind of interpolation with different raster sizes. In our simulations, we analyzed the run times for rasters with varying side length, ranging from 1 to 5 arcseconds. The results are illustrated in Figure 5b. It can be seen, that the required time is greatly reduced with greater side lengths. However, it does not scale in a quadratic way as one could imagine because of the quadratic scaling of the raster area, which mainly results from additional exposure calculation.



(a) Run time evaluation for point interpolation



(b) Run time evaluation for raster interpolation

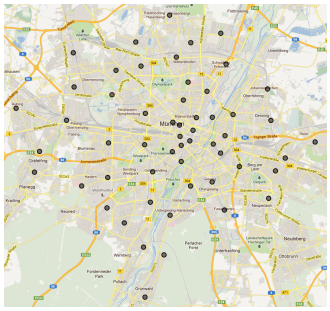
Fig. 5: Performance analysis for point and raster interpolation

4.3 Eco Routing Comparison

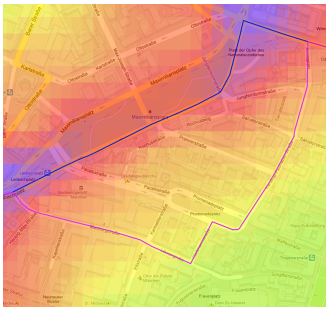
In this section, we analyze the effects of the proposed *Eco Routing* on driving distance, driving time and exposure to air pollution exposure. We therefore compared the traditional routing algorithms *shortest path* and *fastest path* with the results of the corresponding *Eco Routing* algorithms.

Simulation Setup To evaluate the different routing approaches, we simulated realistic air pollution with the above mentioned LUR-based approach for the simulation area (Munich, Germany) and compared the resulting routing decisions based on this data.

For the routing evaluation, we generated a large set of different routes. Therefore we selected a set of 55 starting and end points in Munich. All selected points represent some point of interest (e.g. metro station, popular square, etc.) and are reachable by car and bicycle (e.g. no pedestrian area or highway). Further,



(a) Selected starting and end points in simulation area



(b) *Shortest path* with / without *Eco Routing*

Fig. 6: *Eco Routing* simulation

all points are well connected to street network, i.e. routes in all directions are possible without long detours. The selected points are marked in Figure 6a.

For each of these starting points, routes to all other points were calculated, resulting in 2970 different starting point/end point combinations. Subsequently, 6 routes were determined for each of this combinations:

- Two routes determining the shortest path for cars, once with the traditional approach and once with our proposed *Eco Routing*. An example is shown in Figure 6b.
- Two routes determining the fastest path for cars, again without and with applied *Eco Routing*.
- Two routes for bicycles (using the fastest path). As we assume a constant speed of 16 km/h of cyclists, irrespectively from the chosen road segment, the shortest and fastest path are the same. Consequently, for the bicycle profile, we evaluated only the fastest path with and without *Eco Routing*.

In total this leads to 17820 calculated routes in our simulation.

Simulation Results In order to compare the results from the traditional routing with our *Eco Routing*, we captured the relative difference for driving distance (len_q), driving time ($time_q$), and pollution exposure (exp_q) with

$$q := \frac{v_e - v_n}{v_n} = \frac{v_e}{v_n} - 1,$$

where v_n and v_e denote the resulting values from the traditional routing approach and the exposure-based approach respectively. As a result, negative values for q represent a decrease of the corresponding parameter when applying the proposed *Eco Routing*, positive values for q indicate a lower value for the traditional approach.

In the setting of routing for cars with the *shortest path* approach, the traditional routing outperformed the *Eco Routing* approach. Figure 7 shows that the average of all three evaluated values increased. Remarkable is, that even

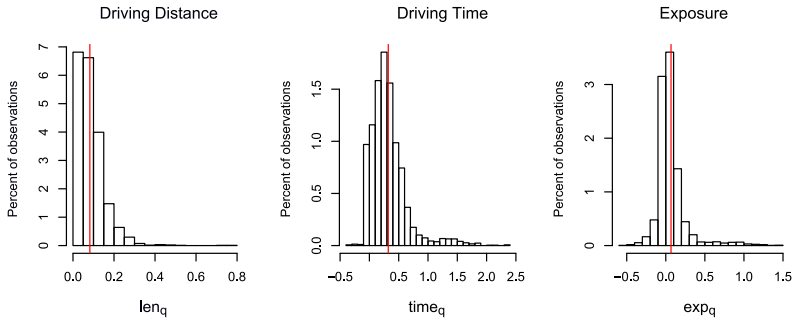


Fig. 7: Resulting relative difference for the *shortest path* approach for cars

the pollution exposure increased by 6.7%, which results from the fact, that the exposure is calculated as a product of time and pollutant concentration. Even though the average exposure might be reduced, the total exposure is increased due to a 32.2% increase of the average driving time.

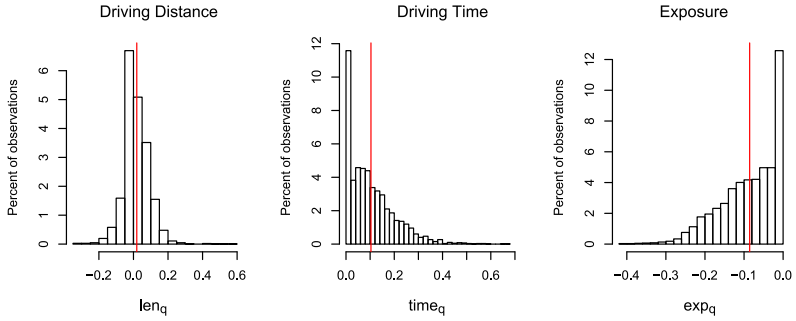


Fig. 8: Resulting relative difference for the *fastest path* approach for cars

The results of the *fastest path* simulations for cars (cf. Figure 8) show that the average pollution exposure could be significantly reduced. In our simulations, a decrease of 8.6% of pollution exposure could be observed when using the *Eco Routing* approach. However, the driving distance increased by 1.9% and the driving time increased by 10.3% in average. The latter is mainly the result of routes with a lower speed limit.

The simulations for the *fastest path* for bicycles led to the biggest decrease of pollution exposure (cf. 9). The average exposure was 12.6% lower with the proposed *Eco Routing* compared to the traditional approach. This was reached with only a moderate increase of driving distance and time, which both increased by 7.1%. A more detailed examination shows that for 10% of the simulated routes the exposure was decreased by even 25%, and for 30% routes it still reduced the exposure at least by 17%.

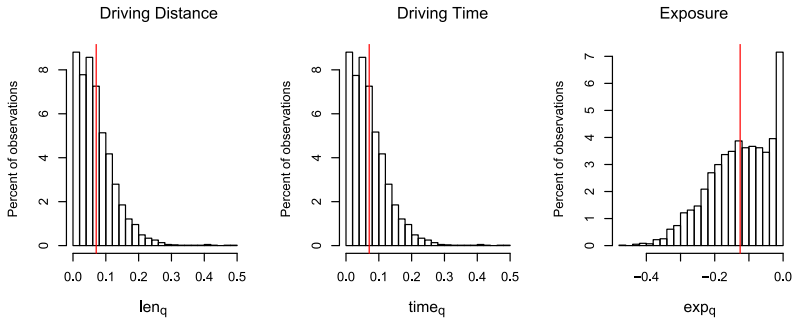


Fig. 9: Resulting relative difference for the *fastest path* approach for bicycles

5 Conclusion and Future Work

In this work, we presented our approach of a *Collaborative Sensing Platform for Eco Routing and Environmental Monitoring*. The proposed architecture enables a scalable and robust collection and processing of ubiquitous sensor data. The obtained information consists of traffic-related pollutants like NO_2, CO, CO_2 and allows for accurate and high resolution environmental monitoring in urban environments. Furthermore, we presented two services including their evaluation: *Eco Routing* and visualization. The former one calculates routes with the objective to minimize the exposure to harmful emissions. The results obtained by simulation indicate that exposure for cyclists was 12.6% lower compared to a traditional approach with a moderate increase by 7.1% of both the driving distance and the driving . For the latter service, we illustrated different techniques to visualize the pollution information for monitoring purposes. In this context, the interpolation process conduces as a basis, wherefore we carried out a performance analysis and presented the corresponding results.

As our sample measurements were generated using a *Land-use regression* model it would be beneficial to obtain more realistic sensor input data. Therefore, a prototypical low power sensor board connected to a smartphone will be applied in a coming field test. On the one hand, future work concentrates on the implementation of additional services like a dynamic pricing system considering current environmental situation. Based on this, a flexible adaptation of the city toll or of fees to rent bicycles can be realized in order to improve the air quality. On the other hand, the collection of data is also in the focus, as continuously taking and sending measurements is cost-intensive in several aspects (energy, traffic, money). Another future task is the performance enhancement of the system by the utilization of cloud computing technologies.

References

1. Shane B. Eisenman, Emiliano Miluzzo, Nicholas D. Lane, Ronald A. Peterson, Gahng-Seop Ahn, and Andrew T. Campbell. Bikenet: A mobile sensing system for cyclist experience mapping. *ACM Trans. Sen. Netw*, 6:6:1–6:39, 2010.

2. Bret Hull, Vladimir Bychkovsky, Yang Zhang, Kevin Chen, Michel Goraczko, Allen Miu, Eugene Shih, Hari Balakrishnan, and Samuel Madden. Cartel: a distributed mobile sensor computing system. In *Proceedings of the 4th international conference on Embedded networked sensor systems*, SenSys '06, pages 125–138, New York and NY and USA, 2006. ACM.
3. L. Luo, A. Kansal, S. Nath, and F. Zhao. Sharing and exploring sensor streams over geocentric interfaces. In *Proceedings of the 16th ACM SIGSPATIAL international conference on Advances in geographic information systems*, page 3, 2008.
4. Michael Jerrett, Altaf Arain, Pavlos Kanaroglou, Bernardo Beckerman, Dimitri Potoglou, Talar Sahsuvaroglu, Jason Morrison, and Chris Giovis. A review and evaluation of intraurban air pollution exposure models. *Journal of Exposure Analysis and Environmental Epidemiology*, 15(2):185–204, 2004.
5. Sarah B. Henderson, Bernardo Beckerman, Michael Jerrett, and Michael Brauer. Application of land use regression to estimate long-term concentrations of traffic-related nitrogen oxides and fine particulate matter. *Environmental Science & Technology*, 41(7):2422–2428, 2007.
6. D.J Briggs, S. Collins, P. Elliott, P. Fischer, S. Kingham, E. Lebret, K. Pryl, H. van Reeuwijk, K. Smallbone, and A. van der Veen. Mapping urban air pollution using gis: a regression-based approach. *International Journal of Geographical Information Science*, 11(7):699–718, 1997.
7. David J. Briggs, Cornelis de Hoogh, John Gulliver, John Wills, Paul Elliott, Simon Kingham, and Kirsty Smallbone. A regression-based method for mapping traffic-related air pollution: application and testing in four contrasting urban environments. *Science of The Total Environment*, 253(1-3):151–167, 2000.
8. Walter F. Dabberdt, F.L. Ludwig, and Warren B. Johnson jr. Validation and applications of an urban diffusion model for vehicular pollutants. *Atmospheric Environment (1967)*, 7(6):603–618, 1973.
9. Nelson L and Seaman. Meteorological modeling for air-quality assessments. *Atmospheric Environment*, 34(12-14):2231–2259, 2000.
10. M. Ehlers, S. Jung, and K. Stroemer. Design and implementation of a gis based bicycle routing system for the world wide web. *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences*, 34(4):425–429, 2002.
11. Alexander Müller. Plattform für Echtzeit-Umweltdaten zur Navigation basierend auf ubiquitären Sensorinformationen, February 2012.
12. Daniel G. Krige. A statistical approach to some basic mine valuation problems on the witwatersrand. *Journal of the Chemical, Metallurgical and Mining Society of South Africa*, 52(6):119–139, 1951.
13. Noel A. C. Cressie. *Statistics for Spatial Data*. Wiley series in probability and mathematical statistics. Wiley, New York and NY, rev edition, 1993.
14. Peter Hart, Nils Nilsson, and Bertram Raphael. A Formal Basis for the Heuristic Determination of Minimum Cost Paths. *IEEE Transactions on Systems Science and Cybernetics*, 4(2):100–107, February 1968.