JOURNAL OF URBAN TECHNOLOGY
https://doi.org/10.1080/10630732.2020.1791679

Quantifying the Spatio-Temporal Potential of Drive-by Sensing in Smart Cities

Amin Anjomshoaa [a,b], Paolo Santi [a,c], Fabio Duarte [a,d], and Carlo Ratti[a]

[a]Senseable City Lab, Massachusetts Institute of Technology, Cambridge, MA, USA; [b]Lero—the Irish Software Research Centre, National University of Ireland, Galway, Ireland; [c]Istituto di Informatica e Telematica del CNR, Pisa, Italy; [d]Pontifícia Universidade Católica do Paraná, Curitiba, Brazil

ABSTRACT
Recently, portable sensors, with high accuracy and embedded communication technologies, have become available and affordable. By deploying such sensors on various urban vehicles that routinely navigate through city streets, vehicles can form a dynamic network for comprehensively and efficiently monitoring the urban environment. This drive-by sensing approach benefits also from the lower costs of sensor deployment and maintenance compared to stationary sensor networks. However, the data sampling frequency and spatial granularity of measurements are constrained by factors such as topology of the underlying street network and mobility pattern of sensor-equipped vehicles. In this paper we investigate the effect of street network topology on the quality of data captured through drive-by sensing. To this end, we first study the temporal aspects of drive-by sensing and present a quantitative method for comparing various street networks. Then, we consider the spatial aspects of drive-by sensing by defining a sensing-potential indicator for urban areas based on the geometrical properties of the street networks. This indicator is then combined with vehicle mobility patterns derived to measure the sensing potential of routes and cycles. In this context, we define the novel concept of Sensogram for describing the spatial sensing potential of network cycles using dedicated vehicles.

KEYWORDS
Environmental monitoring; mobile sensing; mobility patterns; smart city; spatio-temporal phenomena

Introduction

The use of mobile sensors has the potential of enabling a better understanding of urban phenomena and their spatio-temporal variation in the urban environment. Especially in environmental use cases, the spatial and temporal dimensions are of great importance for monitoring and understanding urban environments more effectively. Through drive-by sensing, it is indeed possible to create multiple spatio-temporal data sets using a small fleet of sensor-equipped vehicles such as taxis. This not only reduces the costs of sensor deployment and maintenance by an order of magnitude or more but also offers a possible way of democratizing data by bringing the data to the doorsteps of citizens.

CONTACT Amin Anjomshoaa amina@mit Senseable City Laboratory, MIT 9-216, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

Another advantage of drive-by sensing over traditional urban sensing approaches such as stationary and remote sensing is that it can effectively cover urban areas in both space and time dimensions. In fact, airborne sensing covers large areas of target cities at sparse time intervals (typically, once per day), whereas stationary sensors provide a high temporal coverage, but capture signals only in a few fixed points in space. Considering, as an example, air pollution, stationary sensors measure the ambient pollutants in precise locations which, in fact, could be very different in nearby streets and neighborhoods (Vardoulakis et al., 2005). On the other hand, satellite-based measurements can produce highly accurate spatial content, but with limited temporal coverage. Another example of remote sensing is using satellite imagery to measure surface temperature which, due to temporal limitations, relies on robust mathematical models (Rosenfeld et al., 2017) to predict temperature changes over time.

The drive-by sensing approaches fall under two major categories, namely *opportunistic* and *dedicated* sensing models. In opportunistic sensing, sensors are deployed on existing fleets of vehicles (e.g., buses and taxis) without having any influence on the routes of hosting vehicles. The dedicated sensing vehicles, on the other hand, are exclusively equipped with specific type of sensors and follow predefined routes and schedules.

The goal of this paper is to study the relationship between urban street networks and the spatio-temporal potential of drive-by sensing approaches. Furthermore, we are aiming to create indicators in order to quantify and compare the potential of drive-by sensing approaches in urban areas.

The rest of the paper is organized as follows. The notion of fitness and quality of drive-by data are introduced in the next section, where we also discuss why both topology and geometry of street network graphs should be considered in determining the effectiveness of drive-by sensing approaches. Then, related works will be discussed followed by two sections introducing the concepts and methods used for assessing temporal and spatial potential of drive-by sensing. After this, we add mobility element to the static street network and introduce the concept of *Sensogram* that can be used for planning and scheduling drive-by sensing applications using a dedicated fleet of vehicles. Finally, the last section concludes our work and describes potential applications.

Fitness of Drive-by Sensor Data

In the context of drive-by sensing, the quality of captured data can be interpreted as the number of recorded data points in a specific area and period of time. Figure 1 depicts the concept of spatio-temporal coverage where the required quality of sensing (number of data points inside the space-time box) depends on the phenomenon under study and the target use case. For instance, a high spatial density of data points is needed for capturing noise, whereas temperature can be captured with a lower spatial density. On the other hand, sensing street surface quality is much less sensitive to time compared to other phenomena such as air pollution.

There are a number of ways to increase the density of data samples in space, time, or both dimensions. For instance, increasing the number of sensor-equipped vehicles would increase the number of visits to street segments and decrease the time interval between visits. Furthermore, the mobility pattern of selected vehicles plays a crucial role in density of captured data points along space and time dimensions. Considering, as an

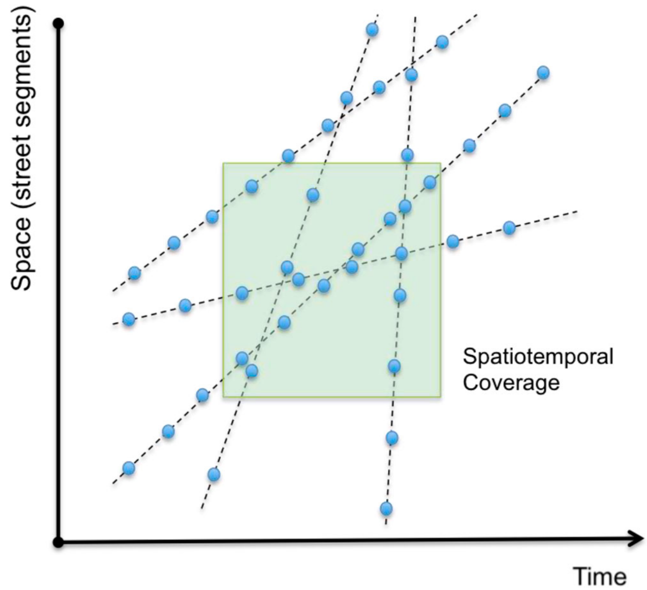

Figure 1. Spatio-temporal data quality in drive-by sensing. Dots are the measurements and lines represent vehicles that visit the selected area (e.g., a street segment) during a specific time window.

example, buses as sensing vehicles and increasing the number of buses on a specific route, the captured data will be denser on the time dimension for the same set of street segments. Whereas adding buses from other bus routes will result in coverage of a larger number of street segments while keeping the same temporal resolution.

Theoretically, the optimal way of performing environmental sensing is to deploy a dense network of stationary sensors in the city. While giving both a high spatial and temporal resolution, this approach is severely limited by cost. A possible solution for enabling the benefits of high spatio-temporal sensing resolution while containing cost is drive-by sensing. In principle, if the fleet of vehicles is large enough (e.g., equal to or larger than the number of street segments) then, with a high probability, we will have one measurement per street segment for any given time window which guarantees an efficient coverage of a city area. However, thanks to vehicles' mobility (e.g., the stochastic mobility pattern of taxis), we do not need such a large fleet of vehicles to obtain a high-density spatio-temporal coverage similar to that provided by a dense stationary sensor network. Our previous research has shown that a remarkably small number of taxis can scan a large number of street segments (O'Keeffe et al., 2019). Furthermore, this analysis has shown that three simple parameters, namely number of taxis, average number of traveled road segments per day, and street segment popularity, can be used to obtain a very accurate estimate of the achieved street network coverage.

So far, the quality of sampled data introduced here is defined by density of data points within the space-time box as depicted in Figure 1 and is characterized solely based on the topology of the street network, the number of vehicles, and their mobility patterns. However, the impact of street network geometry is also a deciding and less-explored factor associated with drive-by sensing applications.

In various urban applications such as flow optimization and routing, a street network can be formalized as a graph representation where vertices represent the intersections and edges represent street segments. In this paper, we refer to the spatial relations between street network elements that remain unaffected by the continuous change of shape or length of street segments, as graph topology, and we use it in analysis of the temporal aspects of sensing potential. The graph topology is particularly well-suited for solving various topological analysis problems, but it does not take properties such as geographical coordinates, lengths, areas, shapes, and angles into account. These characteristics of street networks, which are pertinent to graph geometry, play an important role in the analysis of the spatial aspects of sensing potential.

If we ignore non-planar urban features such as bridges and tunnels, the street network graph can be considered as a geometric planar graph where the vertices are embedded as points in the Euclidean plane, and graph edges partition the city area into faces (also known as parcels). This representation of a street network which divides the city area into a number of regions, is analogous to the notion of Voronoi tessellation that is used extensively in the field of computational geometry for various applications (Aurenhammer, 1991). Using Voronoi tessellation, the area of parcels, which are shaped by the geometry of graph vertices, are further broken down into sub-parcels, based on the minimum distance of points inside parcel to the surrounding edges. In a drive-by sensing scenario, the captured data points are restricted to streets, and in order to extend the results to the whole city area, we need to assign the measurements along the street segments to their corresponding Voronoi neighborhoods.

In practice, methods such as spatial interpolation and extrapolation together with elaborated algorithms and models are used to extend the measurements of street segments to parcel areas and the uncovered street segments. For instance, existing research in the domain of air quality monitoring, proposes a spatial interpolation method for fine particulate matter that considers also the impact of wind-field on the distribution of the particulate matter (Li et al., 2014) or presents an inverse distance-weighted method to estimate daily concentrations of fine particulate matter (Ramos et al., 2016).

Figure 2 depicts the collected air quality data (PM2.5 pollution indicator) from our previous drive-by experiment (Anjomshoaa et al., 2018; deSouza et al., 2020) in the City of Cambridge (MA, US) where the sensors were deployed on trash trucks. In this context, interpolation is a common practice for producing a continuous air quality map in order to cover the parcel areas as well as the uncovered street segment, e.g., the street segment in the middle of highlighted parcel in Figure 2.

Another important factor in drive-by sensing approaches is the effective range of measurements which highly depends on the type of sensor and the phenomenon under study. For instance, while air pressure is known to be consistent across large urban areas, other properties such as air quality and distribution of pollutants may vary considerably within a few meters. Accordingly, sensors are characterized by a spatial *effective range*, and their measurements are not reliable for the points out of their effective range. In other words, if the diameter of a face compared to the effective range of specific sensors is too big, the measurements on its surrounding street segments cannot be extended to all points inside that face. In addition, sometimes the effective range of sensors is further obstructed due to various barriers. For instance, the effective range of a visual camera is limited by line of sight. To clarify the relationship between the effectiveness range of the sensor and the geometry of the street network,

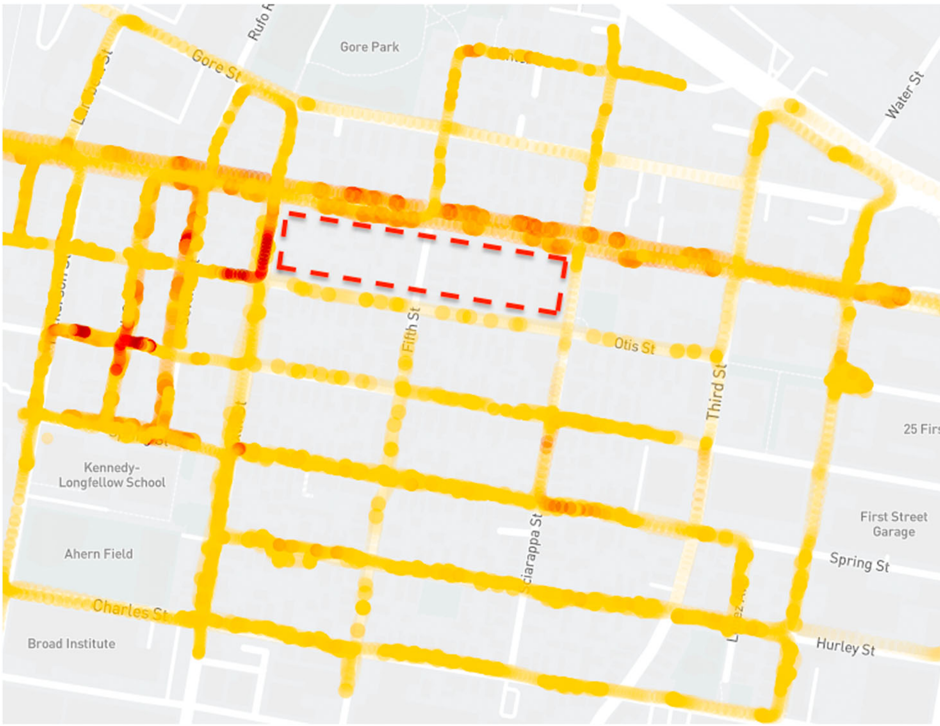

Figure 2. The PM2.5 measurements from a drive-by sensing experiment in City of Cambridge (MA, US)

consider the two highlighted parcels in Figure 3. The two depicted urban parcels have different form factors but are topologically equivalent and have roughly equal areas. The Voronoi tessellations of these parcels, which are depicted by blue lines, show that the distance between central areas of narrower parcel and the actual drive-by data points is smaller than the analogous distance in the wider parcel. As a result, the drive-by measurements may provide a more reliable approximation for the points inside the narrower parcel.

Related Works

Monitoring urban phenomena using mobile sensors is widely studied recently in the context of smart cities. One of the earliest researches in mobile sensing is Wikle and Royle (1999), where the authors formulate methods for constructing optimal spatial sampling designs for environmental monitoring and consider both spatial and temporal variability of environmental phenomena. Thanks to the development of wireless sensor networks and the availability of various portable sensors with high accuracy, it is now possible to monitor urban properties such as road quality (Lu et al. 2013; Wang et al., 2014), air quality (Deville Cavellin et al., 2015; McKercher et al., 2017), and traffic conditions (Bhoraskar et al., 2012; Li et al., 2008).

There is also a handful of research work that discusses the sensor coverage issues in the urban context. For instance, Du et al. (2018) provides a general formulation for deployment of stationary sensors in order to cover target areas, and Pasqualetti et al. (2012)

Figure 3. The impact of geometry on sensing quality

focuses on the problem of patrolling an environment with a team of autonomous agents in order to cover the target areas cooperatively. In their research, the authors define "refresh time" as the time gap between any two visits of the same region. They prove that the problem of minimizing the refresh time is NP-hard and propose a constant factor approximation algorithm for this problem.

There is abundant research about route optimization problems in urban street networks that cover a broad range of topics including the route inspection problem and various node/edge traversal problems (Bruno et al., 2011).

Temporal Sensing Potential

The optimal traversing of street networks is a well-studied problem. Considering the urban street network as a graph whose vertices represent intersections and edges represent street segments, it is possible to address various topological analysis problems such as minimizing the travel time or distance between points in an urban network. In this context, the path planning and routing use cases are among the most common applications of graph theory in cities.

Although the urban street network is often perceived as a graph, its spatial properties cannot be captured using common graph notation alone. Spatial properties are crucial for understanding and studying urban phenomena and are used in various urban domains such as transportation and epidemiology. To this end, the street networks are represented as geometric graphs where both vertices and edges are spatial elements associated with geometric objects and located in a space equipped with a certain metric.

In the drive-by sensing scenarios, the sensor-equipped vehicles visit street segments along their path and capture data for selected phenomena such as air quality, temperature, etc. In order to achieve a comprehensive overview of the city using a single dedicated vehicle, the vehicle needs to visit all street segment in the shortest possible time intervals. This can be translated into the problem of finding a Eulerian cycle (Biggs et al., 1986) such that each edge is visited exactly once. However, given that the topology of an urban street network is usually very complex, finding an Eulerian cycle is either impossible (in non-Eulerian graphs), or otherwise computationally too expensive to calculate. In particular, the necessary condition for the existence of Eulerian cycle is that all vertices in the graph have an even degree, which happens very rarely in real world street networks. As such, it is more realistic to simplify the conditions and allow the sensor-equipped vehicle to visit some edges more than once. In graph theory, the latter case is known as "route inspection problem" (RIP) (Kwan, 1962).

Clearly, if the street network forms an Eulerian graph (Biggs et al., 1986), the graph itself will be the solution to the route inspection problem which means the minimum length of a route in best case is equal to the total length of the graph edges.

Solving the RIP problem for small graphs with fully directed or fully undirected edges is rather simple and can be solved in polynomial time (Eiselt et al., 1995), but for larger street networks it gets computationally very expensive (solvable in $O(|V|^2|E|)$ time). Furthermore, urban street networks are commonly a mixed graph which includes both directed and undirected edges and solving RIP in context of a mixed graph is an NP-complete problem (Papadimitriou, 1976).

In order to provide a scalable indicator for the temporal coverage of cities, we use an existing RIP algorithm (Edmonds and Johnson, 1973) and apply it to the undirected version of city's street network. Furthermore, since the street networks often consist of thousands of nodes and edges, we first partition the city area into smaller rectangles and then intersect these rectangles with the city's street network to find the corresponding sub-graphs. To this end, the bounding box of a target city (longitudes and latitudes of two opposite corners of a city's bounding rectangle) is calculated and then we walk on both latitude and longitude direction with a fixed step-size to create the covering rectangles. In the process of breaking the street network into sub-graphs, the edges that cross the boundary of rectangles are eliminated. So, in order to minimize the number of eliminated edges, the rectangles cannot be too small. On the other hand, if the step-size is too big, the sub-graphs will be also large and the RIP algorithm requires a much longer time to run. For the purpose of this study, the step-size of 0.03 degrees, on both latitude and longitude directions, is used. Accordingly, the area of created rectangles will be around 6.14 square kilometers which in dense city areas may include up to 2,000 street segments and the RIP problem for the corresponding sub-graphs can be solved in a reasonable time. After solving the RIP for each sub-graph, the lengths of resulting routes are added up and considered as the minimum route to cover the whole city. Then, the ratio of RIP route to the total length of street segments will be used as the temporal sensing potential indicator of the drive-by sensing for the target urban area. Although by breaking the large street network into smaller sub-graphs, some edges are discarded, thus introducing some approximation in the computation of the temporal coverage indicator, the RIP results of undirected sub-graphs seem to provide a rather good estimation of total RIP of a whole street network. To clarify this, we may consider the whole sub-graphs as single nodes and then add the discarded edges to connect those nodes. If we apply the RIP algorithm to this simplified graph, we end up with a RIP solution which might not be optimal (compared to minimum possible number of edges in a route inspection problem) but still traverses all edges of the underlying street network. As a result, our proposed approach is partially optimized (within sub-graphs) and its resulting route is longer than the optimal RIP solution for the whole graph.

We have applied this method to the street network of six cities/borough with different sizes and densities: Manhattan (NY, USA), San Francisco (CA, USA), Cambridge (MA, USA), Malden (MA, USA), Vienna (Austria), and Parma (Italy). The results are shown in Table 1. The last column includes the RIP to total network length ratio, that in the case of a Eulerian graph should be equal to one. Since complex street networks are unlikely to contain a Eulerian path, this ratio is commonly larger than one. Also, in the worst case, a vehicle should visit each street segment twice (e.g., in the case of street networks with star topology), which implies that the upper bound of this ratio is two.

A closer look into the values of the ratio and the number of nodes and edges in sub-graphs reveals some interesting properties of RIP in various street networks. For instance, Figure 4 compares the distribution of RIP to network-length ratios for larger cities in our list, namely Vienna, Parma, San Francisco, and Manhattan. If we ignore the last column, which is mostly generated by the sub-graphs of small suburban areas (star-shaped topology), the distribution of the ratio, to some extent, is affected by the structure of the street network. More regular street networks, such as San Francisco's, tend to result in relatively lower values of the temporal coverage indicator, while the opposite is true for older and more irregular street networks such as Parma's.

In this context, another interesting and unexpected result is that as density in street networks increases and we get more nodes and edges in sub-graphs bounded by equally sized blocks, the RIP to network length ratio gets closer to one. This finding is shown in Figure 5, where the size of the circles depicts the RIP to network-length ratio and as we move to the right-top corner, the circles for all four cities get smaller and smaller. A possible

Table 1. Route Inspection coverage and temporal coverage indicators

City Name	Edges	Nodes	Sub-Graphs	Network Length (m)	RIP Length (m)	RIP to Network Ratio
Manhattan, USA	4,573	8,134	16	802,214.44	907,724.62	1.13
San Francisco, USA	9,559	15,829	26	1,599,175.12	1,914,576.88	1.20
Cambridge, USA	1,808	2,678	5	216,732.49	274,197.40	1.27
Malden, USA	1,448	1,987	6	159,598.28	209,895.54	1.32
Vienna, Austria	16,043	24,052	72	2,303,142.61	2,982,201.15	1.29
Parma, Italy	6,086	8,307	51	715,825.76	1,065,927.43	1.49

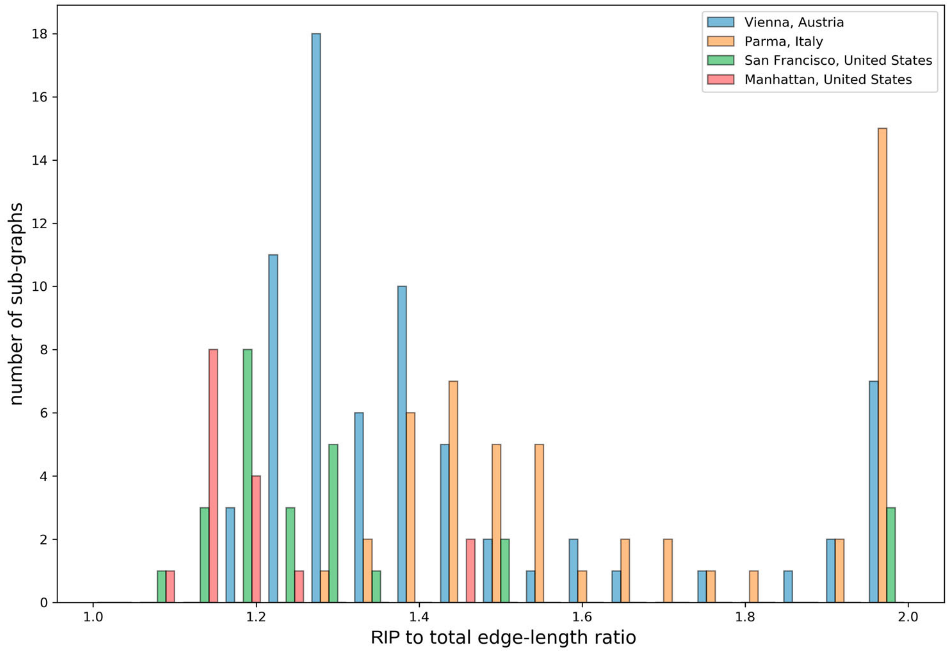

Figure 4. Distribution of urban network coverage of four cities (Vienna, Parma, San Francisco, and Manhattan) based on RIP to network-length ratios

explanation for this observation could be a kind of self-organizing feature in the street networks that emerges as density of nodes and edges in the network increases.

Also, by comparing the ratios of different cities it can be observed that ratio values of cities with more structured and grid-shaped street networks such as Manhattan or San Francisco, are relatively smaller compared to those of traditional and unbalanced street networks like those of Parma or Vienna. As a matter of fact, this finding is in line with the reality and complies with routing and navigation scenarios in practice.

The RIP to network-length ratio has an implication for the temporal potential of urban sensing because by calculating this ratio and considering the average speed of vehicles in an urban area, we can estimate the required time to cover the whole street network using a dedicated sensor-equipped vehicle. This estimated time, which is analogous to the sampling rate of drive-by measurements, is commonly large for bigger cities. Consequently, to cover urban areas with a higher sampling rate using dedicated vehicles, a larger number of vehicles should be employed. In addition, implementing the drive-by sensing method using a stochastic fleet of vehicles such as taxis is another possible solution that covers a rather large portion of street segments (O'Keeffe et al., 2019), but does not guarantee covering each single street segment as in the case of the coverage achieved by a RIP route and dedicated vehicles.

Spatial Sensing Potential

The sensors that are used to monitor continuous phenomena such as air quality, generally have an effective spatial range (hereafter denoted by ρ) that can range from few

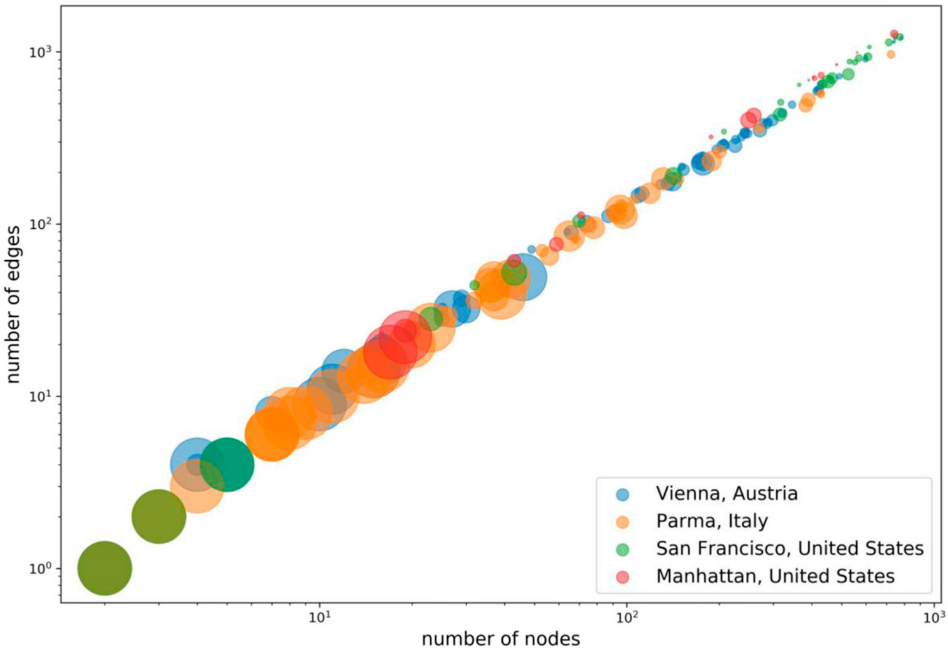

Figure 5. Urban network coverage of four cities (Vienna, Parma, San Francisco, and Manhattan) based on Route Inspection Problem of sub-graphs

centimeters to many kilometers. For instance, the effective range of an air quality sensor is a few meters. In order to have a reliable set of measurements and uniform coverage of a city area, we would need a dense sensor network with maximum distance of ρ between its sensors. Since deployment of such a sensor network, for sensors with relatively small effective ranges, is neither economic (due to the large number of sensors required) nor feasible (due to the urban constraints for positioning sensors in a perfect grid form), drive-by sensing offers a trade-off between number of sensors and quality of data. In this section, we study the spatial coverage of city area by mobile sensors and discuss how drive-by measurements, which are spatially restricted to a street network, can be extended to a city area.

As mentioned earlier, street networks are spatial graphs that in addition to topology include also geometry information such as geographical coordinates, lengths, areas, shapes, and angles. Due to the complexity of street networks such as bridges and tunnels, some street networks cannot be drawn on a two-dimensional plane and as a result they cannot be considered as planar graphs (Boeing, 2018). In this research however, we use the planar simplification of complex street networks, which for the purpose of this paper provides a good approximation of sense-ability measures.

Consider a planar graph $G = (V, E)$ that partitions the urban area into parcels or blocks (analogous to faces in graph terminology). The spatio-temporal phenomena χ should be modeled based on the measurements on the graph edges. As such, the area of a city is decomposed into a finite set of faces $\{f_1, f_2, \ldots, f_n\}$ where the number of faces, n, for a graph with v vertices and e edges, can be calculated according to Euler's formula as $n = 2 - v + e$.

For a given sensor type, we define the *sensing neighborhood* of face f_i as the circle centered at the center of mass of f_i with radius ρ, the effective range of the target sensor, and denote it by $\mathcal{N}_i|_\rho$. In this context, the sensor-equipped vehicles that visit the street segments located within the neighborhood of a specific face, will contribute to the estimation of sensor data for that face. If the effective range of a sensor is large (e.g., in the case of a barometer sensor), the sensing neighborhood may include the edges of the face itself as well as some edges from neighboring faces. Figure 6 shows three sensing neighborhoods corresponding to different sensor types and their effective ranges. The small circle C_1 corresponds to a sensor with a small effective range and the circle is not covered any edges, which indicated that parts of the selected parcel cannot be covered by the drive-by sensing approach with this specific sensor type. On the other hand, the C_2 neighborhood covers a number of edges which means the measurement on those edges can contribute to the sensing of a selected parcel. Finally, the C_p neighborhood shows the smallest possible circle that includes all the edges of selected parcel.

Now, we define effective edges as the collection of street segments that contribute to the sensing of a selected parcel. To this end, the edges that are completely covered by the sensing neighborhood of a selected parcel are considered as effective edges. In addition to these trivial edges, some other street segments which partially intersect with the sensing neighborhood can also contribute to the sense-ability of graph faces. So, we extend the definition of effective edges by including the intersecting edges whose

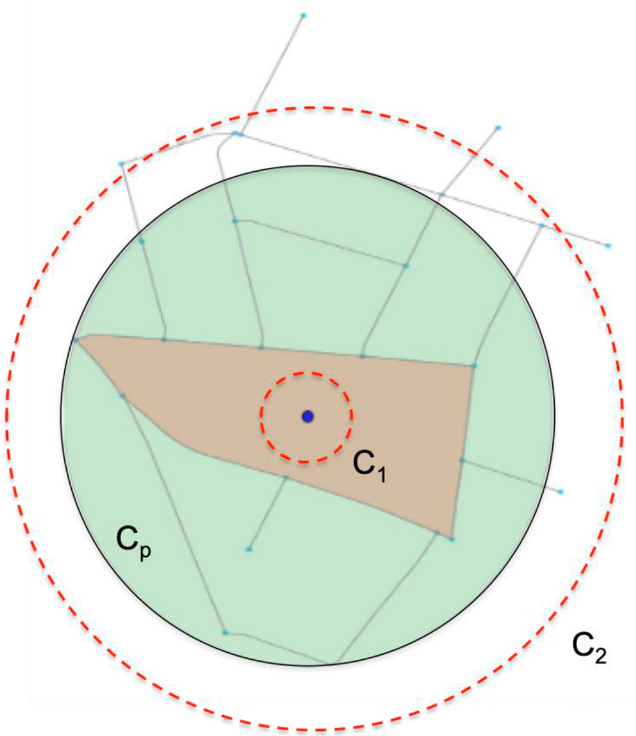

Figure 6. Sensing neighborhoods of a face f_i based on various sensors with different effective ranges

overlap with a neighborhood is larger than a given intersection threshold α :

$$\langle \mathcal{C} \rangle_{f_i} = \{e \in \mathrm{E} \mid e \cap \mathcal{N}_i|_\rho > \alpha\} \tag{1}$$

The sensing potential of face f_i is then calculated based on the area of f_i and the length of its effective edges. In order to normalize this indicator and make sensing potential of different faces comparable, we introduce an equivalent circle for all face f_i that has the same area. Since the perimeter of a circle has the smallest length among all planar shapes with a fixed area, this will give us a uniform measure to compare the sensing potential of graph faces and find out how better their effective edges contribute to the sensing of faces compared to the worst case of a circle perimeter, which has the shortest sensing path with the same area. Consequently, the sensing potential of a face f_i is denoted by $\mathcal{S}_{f_i}$ and defined as follows:

$$\langle \mathcal{S} \rangle_{f_i} = \frac{1}{\mathcal{P}_i} \sum_{e_n \in \langle \mathcal{C} \rangle_{f_i}} \|e_n\| \tag{2}$$

where $\mathcal{P}_i$ is the perimeter of the circle whose area is equal to the area of face f_i, denoted by $\mathcal{A}_{f_i}$, and is defined as follows:

$$\mathcal{P}_i = 2\sqrt{\pi \mathcal{A}_{f_i}} \tag{3}$$

According to this definition, a face with a longer total length of effective edges has a bigger sensing potential because the drive-by sensing vehicle will have a longer path to do the measurements for that specific face.

Figure 7 shows the sensing potential of graph faces, which is characterized by the effective edges of each face and their contribution to the estimation of sensor values for the corresponding face. The higher the sensing potential of a given face via its corresponding effective edges, the darker the depicted face color is. Clearly, smaller faces receive enough measurements from surrounding street segments, and we would have a higher sensing potential. On the other hand, the larger areas require more measurements in order to provide a reliable estimation of sensor data for their inner points. The total sum of sensing potentials across all city faces provides the maximum sensing potential of a city area. In order to achieve this value, we would need one measurement per street segment and per given time window, which is feasible by either having a stationary sensor at each street segment, or a rather large number of sensor-equipped vehicles. The next section introduces the notions of time and mobility, and the total sum of sensing potential of faces will be considered as the higher bound of sensing potential that is desired to achieve by means of sensor-equipped vehicles.

Furthermore, some edges might contribute to the sensing potential of many faces which makes them more significant in a drive-by sensing scenario. As such, the definition of sensing potential provides a measure to assign a geo-significance attribute to the street segments, which in turn can be used as a weight in route planning of drive-by sensing scenarios in order to use the more significant street segments when possible.

Figure 8 shows the geo-significance of graph edges for the city of Melrose, MA, USA, where the significant edges are highlighted in red and the width of street segments depicts its geo-significance value. It is important to note that the street segments with lower geo-

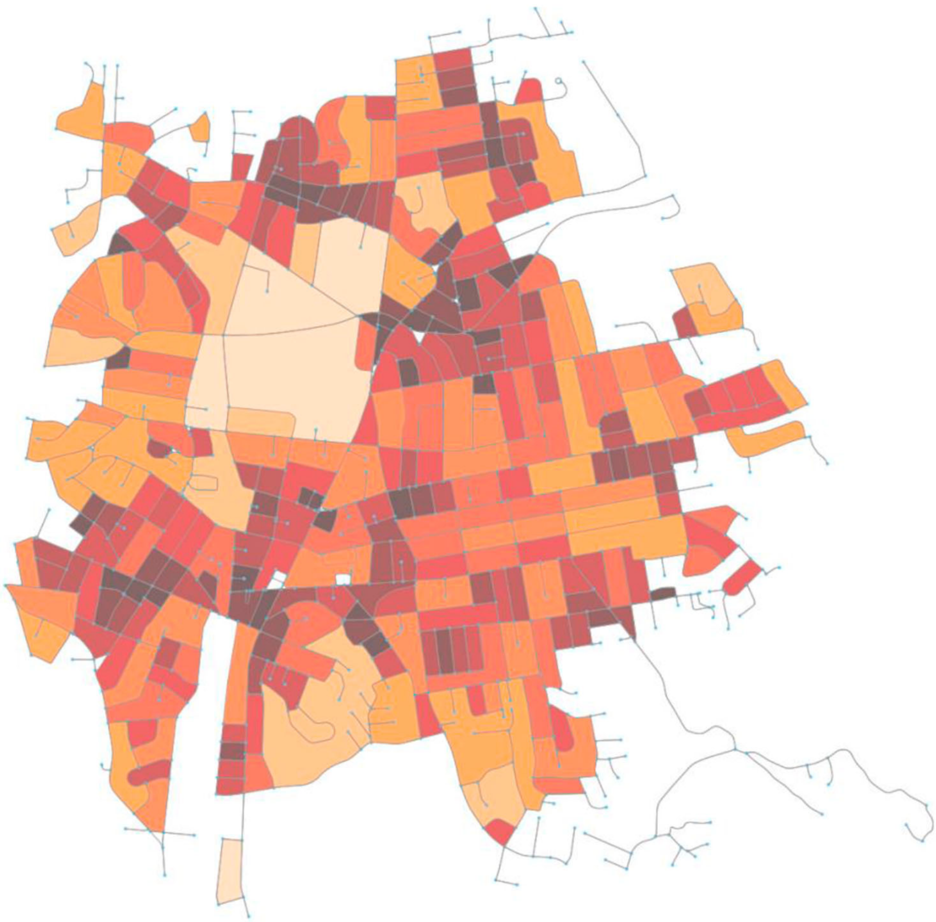

Figure 7. Sensing potential of graph faces in City of Melrose, MA, US

significance are also needed in a drive-by sensing scenario because, although these segments do not contribute greatly to the total sensing coverage of the city, they might still be needed to ensure coverage of the entire city area. Furthermore, a side-product of the geo-significance concept could be the optimal positioning of stationary sensors in locations that most contribute to the urban sensing of selected phenomenon.

Mobility and Route Planning

So far, the sensing potential of an urban area is calculated merely based on the static features of the underlying street network. In this section we will discuss the mobility aspect and explore its impact on spatio-temporal sensing of urban areas.

Suppose we have a sensor-equipped vehicle that visits various street segments based on a random walk model (e.g., a taxi). As discussed in the previous section, each street segment serves as an effective segment of one or more faces of a street network. So, as the vehicle touches a street segment, the corresponding faces will be activated and their sensing potential will increase in accordance with the length of the street segment and

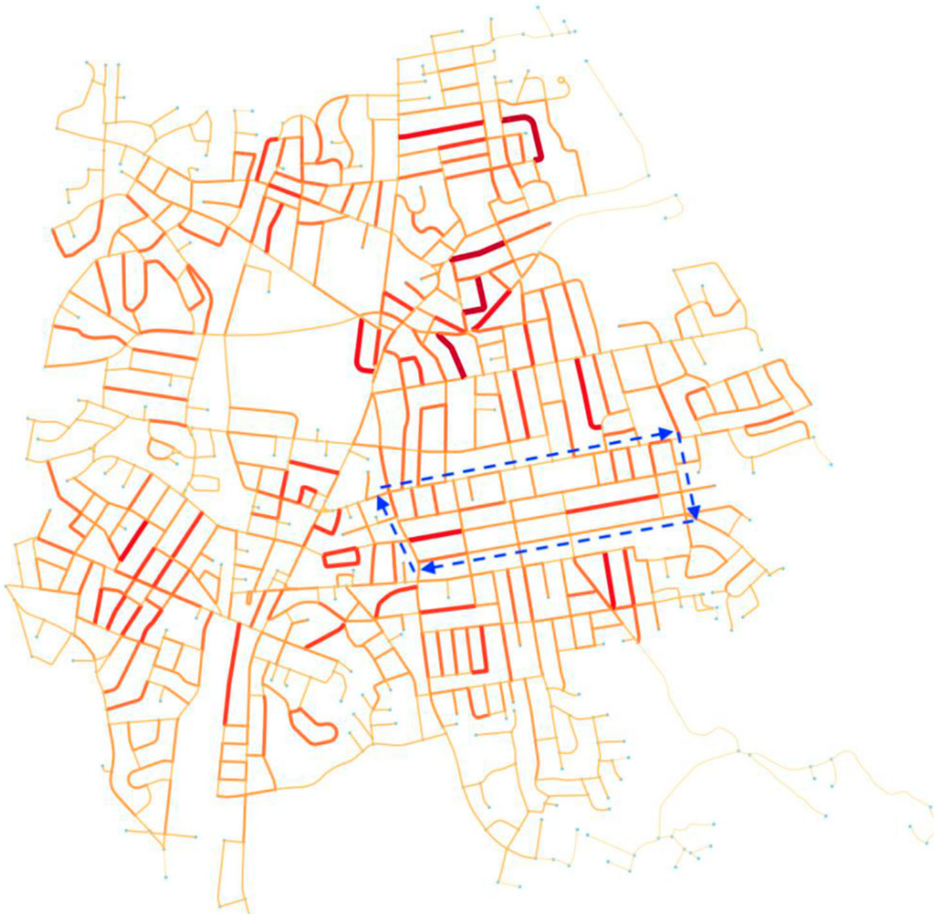

Figure 8. Geo-significance of street segments in City of Melrose, MA, US. The edge colors and edge widths depict the contribution of a specific edge to the total sensing coverage of the target city

the area of the affected faces. This increased level of sensing potential, remains unchanged until either one of these two cases happens:

(1) the vehicle visits another effective edge of this face and the sensing potential increases accordingly
(2) the measurements which are captured at the effective segments get outdated according to a decay function (e.g., air quality measurements may be considered outdated after five seconds), and as a result the sensing potential of the face decreases.

Now, if the sensor-equipped vehicle travels in a cycle (e.g., a bus traveling in a predetermined cycle), we can use the sensing potential of neighboring faces in order to calculate the total sensing potential of a selected cycle. As the vehicle travels in this cycle, the time-series of sensing potential values can be calculated based on the visited street segments and the decay of previously measured data points. If the vehicle manages to visit the street segment before the previously measured data points get expired, then the actual sensing potential will not drop and gets closer to the maximum sensing potential of affected faces. In other words, if the time required to complete the cycle is less than the decay time of the sensor data, the actual sensing potential can reach its maximum value.

We now define the concept of "Sensogram" as the accumulated sensing potential of faces that are activated by the visited street segments in a drive-by sensing application. In fact, the sensogram concept exploits the previously introduced sensing potential indicator in order to demonstrate the sensing power of scheduled vehicles for a specific route or cycle. To clarify this concept, sensograms of a 3 km cyclic route (located in Melrose, MA, USA as depicted in Figure 8, blue dotted lines) for varying numbers of vehicles, are demonstrated in Figure 9. As the number of patrolling vehicles increases, the sensing potential of covered faces gets closer and closer to the ultimate sensing potential of surrounding faces. For the selected cycle of this example, when the number of vehicles is increased to 15 (purple line in Figure 9), the sensogram turns into a constant horizontal line which is characterized by the ultimate sensing potential of affected faces. The sensograms also exhibit a periodic behavior which corresponds to the decay time of sensor data and the time interval between successive vehicles. For this example, we used 20 seconds as the decay time of sensor data and 30 seconds distance between successive sensing vehicles. Based on the ascending and descending phases of

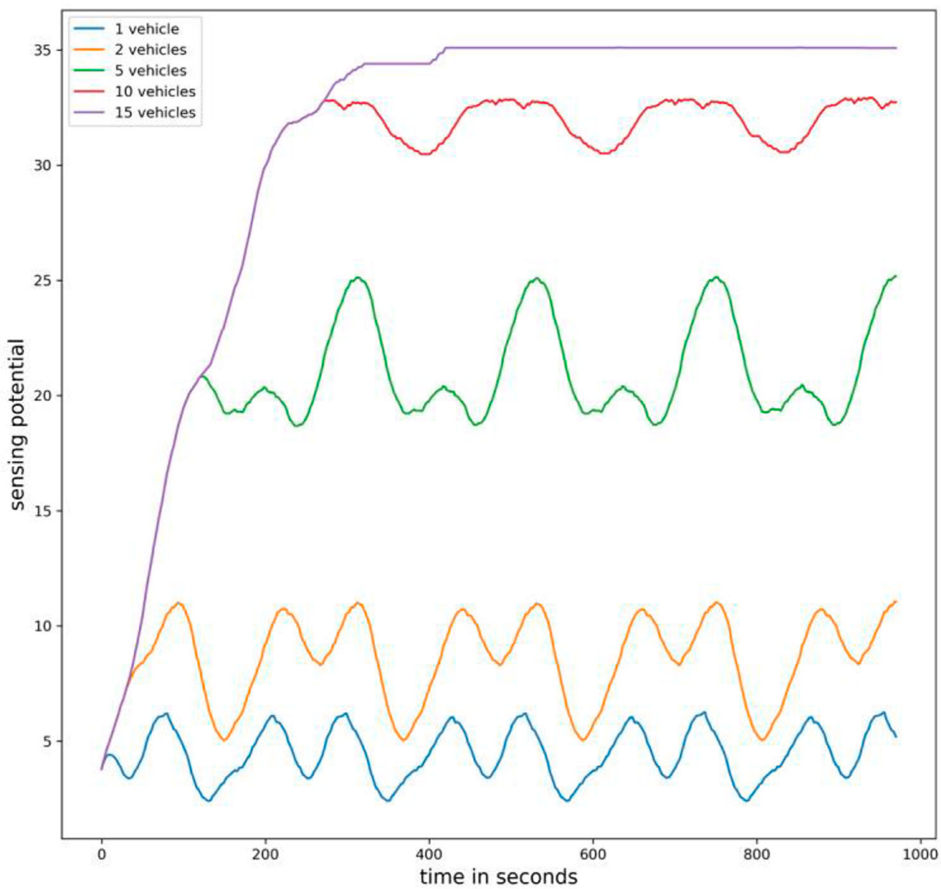

Figure 9. Sensograms of the cycle depicted in Figure 8

sensograms and the decay time of captured data, one can determine the appropriate distance between successive vehicles to maintain a uniform sensing potential during the sensing period. For instance, the sensogram corresponding to a single vehicle in Figure 9 implies that adding a second vehicle with a phase difference of around 150 seconds (instead of the assumed 30 seconds intervals shown in Figure 9), can achieve a more uniform sensing potential.

Conclusions

The ambitious goal of this research is to provide a theoretical framework for studying the impact of spatial complexity and texture of street networks on the efficiency of drive-by sensing in cities. The proposed framework in this research is not tailored to the sensing requirements of specific urban phenomena or sensor types. Depending on the spatio-temporal sensitivity of the target phenomena and the effective range of corresponding sensors, drive-by sensing use cases should be designed in such a way that address the required temporal and spatial densities of collected data points. Furthermore, in some sensing scenarios the full coverage of space may not be feasible or necessary. For instance, in the case of visual cameras, the effective range of cameras can be limited due to various barriers (e.g., buildings, trees, etc.), or in the case of road quality monitoring, we do not need to cover the space between street areas at all. So, in addition to collecting data points along the street network, we would need elaborated and customized models that consider various properties of urban parcels such as building footprints and land use to understand and evaluate the phenomenon under study. Since this research and the proposed framework are sensor-agnostic and do not target a specific urban phenomenon, these domain specific models are not addressed in this initial effort and remain as future work.

Drive-by sensing has the potential of effectively capturing environmental phenomena in our cities with a small number of urban vehicles. We envision a paradigm of modular sensing components and their corresponding cloud services that enable cities to deploy their required set of sensors in order to create elaborate applications for their inhabitants in a cost-effective manner. In this paper, we discussed the spatio-temporal aspects of drive-by sensing and introduced some indicators to measure the potential of drive-by sensing based on the underlying street network. For the case of temporal sensing potential, we studied the topology aspects of a street network and provided a method for calculating the RIP to street network ratio. This ratio can be used to compare the required time for covering street networks using dedicated vehicles and also showed some interesting features about the relationship between city form and a street network traversing in urban areas. Regarding the spatial sensing potential, we defined a sensing potential indicator for the graph faces which exploits the geometry aspects of street networks. The sensing potential indicator is then used to assess the geo-significance of street segments with potential applications in route planning and optimization of dedicated sensing vehicles as well as quantifying the effectiveness of a stochastic fleet of vehicles. Finally, the mobility and time aspects are introduced into the street network context, defining the novel concept of a sensogram based on the introduced sensing potential indicator. A sensogram is then used to explore the relationship between a street network and a fleet of vehicles for drive-by sensing scenarios. The tools and metrics introduced in this paper will enable a better understanding and planning of drive-by sensing applications in smart cities.

Acknowledgments

The authors would like to thank the members of the MIT Senseable City Lab Consortium for supporting this research.

Funding

The work of A. Anjomshoaa is supported in part by Science Foundation Ireland [grant 13/RC/2094] and co-funded under the European Regional Development Fund through the Southern & Eastern Regional Operational Programme to Lero—the Science Foundation Ireland Research Centre for Software (www.lero.ie).

Notes on Contributors

Amin Anjomshoaa is a senior research scientist at Senseable City Lab, Massachusetts Institute of Technology, Cambridge, US. He is also associated with Lero—the Irish Software Research Centre, National University of Ireland. Anjomshoaa defined the problem, designed the solution and models, and wrote the paper.

Paolo Santi is a principal research scientist at Senseable City Lab, Massachusetts Institute of Technology, Cambridge, US. He is also associated with Istituto di Informatica e Telematica del CNR, Pisa, Italy. Santi supervised the research and contributed to the writing.

Fabio Duarte is a principal research scientist at Senseable City Lab, Massachusetts Institute of Technology, Cambridge, US. He is also associated with Pontifícia Universidade Católica do Paraná, Brazil. Duarte contributed to the writing and the visualization of the results.

Carlo Ratti is the director of Senseable City Lab, Massachusetts Institute of Technology, Cambridge, MA, US. Ratti secured the financial support and supervised the research.

ORCID

Amin Anjomshoaa http://orcid.org/0000-0001-6277-742X
Paolo Santi http://orcid.org/0000-0002-8942-8702
Fabio Duarte http://orcid.org/0000-0003-0909-5379

References

A. Anjomshoaa, F. Duarte, D. Rennings, T. J. Matarazzo, P. deSouza, and C. Ratti, "City Scanner: Building and Scheduling a Mobile Sensing Platform for Smart City Services," *IEEE Internet of Things Journal* 5: 6 (2018) 4567–4579.

F. Aurenhammer, "Voronoi Diagrams: A Survey of a Fundamental Geometric Data Structure," *ACM Computing Surveys* (CSUR) 23: 3 (1991) 345–405.

R. Bhoraskar, N. Vankadhara, B. Raman, and P. Kulkarni, "Wolverine: Traffic and Road Condition Estimation Using Smartphone Sensors," paper presented at the Fourth International Conference on Communication Systems and Networks (Bangalore, January 3, 2012).

N. Biggs, E.K. Lloyd, and R.J. Wilson, *Graph Theory 1736-1936* (New York: Oxford University Press, 1986).

G. Boeing, "Planarity and Street Network Representation in Urban form Analysis," *Environment and Planning B: Urban Analytics and City Science* (2018) <https://doi.org/10.1177/2399808318802941> Accessed May 22, 2020.

G. Bruno, A. Genovese, and G. Improta, "Routing Problems: A Historical Perspective," *Journal of the British Society for the History of Mathematics* 26: 2 (2011) 118–127.

P. deSouza, A. Anjomshoaa, F. Duarte, R. Kahn, P. Kumar, and C. Ratti, “Air Quality Monitoring Case Study Using Mobile Low-cost Sensors mounted on Trash-Trucks: Methods Development and Lessons Learned,” *Sustainable Cities and Society* (2020) <https://doi.org/10.1016/j.scs.2020.102239>.

L. Deville Cavellin, S. Weichenthal, R. Tack, M. S. Ragettli, A. Smargiassi, and M. Hat-zopoulou, “Investigating the Use of Portable Air Pollution Sensors to Capture the Spatial Variability of Traffic-related Air Pollution,” *Environmental Science & Technology* 50: 1 (2015) 313–320.

R. Du, P. Santi, M. Xiao, A. V. Vasilakos, and C. Fischione, “The Sensable City: A Survey on the Deployment and Management for Smart City Monitoring,” *IEEE Communications Surveys & Tutorials* 21: 2 (2018) 1533–1560.

J. Edmonds and E. L. Johnson, “Matching, Euler Tours and the Chinese Postman,” *Mathematical Programming* 5: 1 (1973) 88–124.

H. A. Eiselt, M. Gendreau, and G. Laporte, “Arc Routing Problems, Part i: The Chinese Postman Problem,” *Operations Research* 43: 2 (1995) 231–242.

M. Kwan, “Graphic Programming Using Odd or Even points,” *Chinese Mathematics* 1 (1962) 273–277.

L. Li, J. Gong, and J. Zhou, “Spatial Interpolation of Fine Particulate Matter Concentrations Using the Shortest Wind-field Path Distance,” *PloS One* 9: 5 (2014) e96111.

X. Li, W. Shu, M. Li, H.-Y. Huang, P.-E. Luo, and M.-Y. Wu, “Performance Evaluation of Vehicle-based Mobile Sensor Networks for Traffic Monitoring,” *IEEE Transactions on Vehicular Technology* 58: 4 (2008) 1647–1653.

Y. Lu, Y. Zhang, Y. Cao, J. G. McDaniel, and M. L. Wang, “A Mobile Acoustic Subsurface Sensing (MASS) System for Rapid Roadway Assessment,” *Sensors* 13: 5 (2013) 5881–5896.

G. R. McKercher, J. A. Salmond, and J. K. Vanos, “Characteristics and Applications of Small, Portable Gaseous Air Pollution Monitors,” *Environmental Pollution* 223 (2017) 102–110.

K. P. O’Keeffe, A. Anjomshoaa, S. H. Strogatz, P. Santi, and C. Ratti, “Quantifying the Sensing Power of Vehicle Fleets,” *Proceedings of the National Academy of Sciences* 116: 26 (2019) 12752–12757.

C. H. Papadimitriou, “On the Complexity of Edge Traversing,” *Journal of the ACM* 23: 3 (1976) 544–554.

F. Pasqualetti, J. W. Durham, and F. Bullo, “Cooperative Patrolling via Weighted Tours: Performance Analysis and Distributed Algorithms,” *IEEE Transactions on Robotics* 28: 5 (2012) 1181–1188.

Y. Ramos, B. St-Onge, J.-P. Blanchet, and A. Smargiassi, “Spatio-temporal Models to Estimate Daily Concentrations of Fine Particulate Matter in Montreal: Kriging with External Drift and Inverse Distance-weighted Approaches,” *Journal of Exposure Science and Environmental Epidemiology* 26: 4 (2016) 405–414, <https://doi.org/10.1038/jes.2015.79>.

A. Rosenfeld, M. Dorman, J. Schwartz, V. Novack, A. C. Just, and I. Kloog, “Estimating Daily Minimum, Maximum, and Mean Near Surface Air Temperature Using Hybrid Satellite Models Across Israel,” *Environmental Research* 159 (2017) 297–312.

S. Vardoulakis, N. Gonzalez-Flesca, B. E. Fisher, and K. Pericleous, “Spatial Variability of Air Pollution in the Vicinity of a Permanent Monitoring Station in Central Paris,” *Atmospheric Environment* 39: 15 (2005) 2725–2736.

M. Wang, R. Birken, and S. S. Shamsabadi, “Framework and Implementation of a Continuous Network-wide Health Monitoring System for Roadways,” paper presented at the SPIE Nondestructive Characterization for Composite Materials, Aerospace Engineering, Civil Infrastructure, and Homeland Security (San Diego, California, March 8, 2014).

C. K. Wikle and J. A. Royle, “Space: Time Dynamic Design of Environmental Monitoring Networks,” *Journal of Agricultural, Biological, and Environmental Statistics* 4: 4 (1999) 489–507.