PGA: Physics Guided and Adaptive Approach for Mobile Fine-Grained Air Pollution Estimation

Xinlei Chen

Carnegie Mellon University
xinlei.chen@sv.cmu.edu

Jiayou He

Beijing Experimental High
School Attached to Beijing
Normal University
jiayouhejiayou@gmail.com

Xiangxiang Xu

Tsinghua University
xuxx14@mails.tsinghua.edu.cn

Hae Young Noh

Carnegie Mellon University
noh@cmu.edu

Xinyu Liu

Tsinghua University
liuxinyu13@mails.tsinghua
.edu.cn

Lin Zhang

Tsinghua-Berkeley Shenzhen
Institute
linzhang@tsinghua.edu.cn

Shijia Pan

Carnegie Mellon University
shijiapan@cmu.edu

Pei Zhang

Carnegie Mellon University
peizhang@cmu.edu

Abstract

Large-scale fine-grained air pollution information has both financial for city managers and health benefits for all city residents. Sensors installed on fleet of vehicles (like taxis) to collect air pollution data provides a low cost, low-maintenance and a potentially high coverage approach. The challenges here are: 1) Sensing coverage is sparse in some areas, and 2) it changes over time.

This paper presents PGA, a physics guided and adaptive approach to estimate fine-grained air pollution with vehicle fleets. We combine the advantages of a physics guided model and a data-driven model to achieve high accuracy with high spatial-temporal resolution. To evaluate the system, we deploy our air pollution sensing hardware on 29 taxis in the city of Shenzhen, and collected around 26.3 million data samples within 14 days. The results show that our system achieves up to $4.0\times$ reduction on average error compared to state-of-the-art approaches.

ACM Classification Keywords

H.4.2. [Information Systems Applications]: Data Analytics

Author Keywords

Air Pollution Estimation, Adaptive Approach, Urban Sensing

INTRODUCTION

Air pollution has become a major public health crisis worldwide [1, 2]. According to the World Health

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.
UbiComp/ISWC'18 Adjunct, October 8-12, 2018, Singapore, Singapore. 2018 Association for Computing Machinery.
ACM ISBN 978-1-4503-5966-5/18/10...\$15.00
<https://doi.org/10.1145/3267305.3274179>.

Organization (WHO), every year more than 3 million deaths are linked to exposure to outdoor air pollution and 92% of the world's population lives in places where air pollution levels exceed WHO limits [3]. It is essential to provide fine-grained (e.g. such as block level at every hour) air pollution information to both residents and city managers [4].

Prior researches have investigated using mobile sensing platforms to improve coverage and resolution over static approaches [5, 6]. Sensors are installed on vehicle fleets (especially taxis) to collect air pollution data with reduced cost and improved mobility. Two major approaches are used to derive fine-grained air pollution estimation with mobile: physics guided and data driven. The first methods adopt physical principle, expressed in equation, to describe how air evolves over time and space [7, 8, 9]. These methods can calculate very high resolution pollution maps. However, these approaches generally have low accuracy due to 1) difficulty to capture all possible factors that influence air pollution, such as geographical information like buildings, pollution sources etc., and 2) inaccurate empirical parameter assumption or estimation for dynamic urban environment. The second methods, instead of utilizing empirical physical knowledge, only rely on the sensed data (air pollution and other information such as traffic, weather etc.) to derive air pollution estimation [10, 5, 6, 11, 12]. Such methods are highly biased by the data collection. The accuracy is high when sufficient data is collected, but deteriorates when sensing coverage is sparse.

Major challenges: sparse and time-variant sensing coverage. (i.) *Sparse Sensing Coverage:* Since most mobile sensing nodes (vehicles or taxis) gather around busy areas, it leaves large parts of the city not sensed at any given time. Consequently, insufficient information is provided to achieve fine-grained estimation. (ii.) *Time-variant Sensing Coverage:* The sensed and

non-sensed areas change significantly over time due to uncontrollable movements of these mobile nodes. As a result, the estimation accuracy varies over time.

This paper presents a physics guided and adaptive (PGA) approach to derive air pollution estimation with high spatial (block level) and temporal (hourly) resolution. We design an adaptive system structure, which dynamically combines the advantages of a physics guided model and a data driven model using a particle filter structure. To address the challenge of sparse sensing coverage, the system utilizes the temporal and spatial relations of air dispersion from the physics guided model to introduce extra information. The physics guided model, based on such relations and collected data, infers information at the uncovered areas based on the temporal and spatial relations and the collected data. To address the challenge of time-variant sensing coverage, PGA adaptively selects combination information from either the physics guided model or data driven model at different time and locations. The intuition here is the system weighs more on the physics guided model when sensing coverage is sparse and weighs more on the data driven model when sensing coverage is dense.

The main contributions of this paper are:

- Reformulating a physics guided model for air dispersion state evolution estimation and combining it with a data driven model.
- Proposing an adaptive scheme to improve estimation from a physics guided model with "observation" generated from a data driven model.
- Deploying and evaluating the system on a large-scale vehicular sensing platform for evaluation.

To evaluate the system, we deploy our low cost air pollution sensing hardware (~ \$1000 each) on 29 taxis in the city of Shenzhen, and collected around 26.3 million data samples within 14 days.

Figure 1: The figure shows our PGA system architecture.

The remainder of the paper is organized as follows: Section presents the system overview. Section describes how we designs the system algorithm in detail. Section presents the experiments for system evaluation and section concludes the paper.

SYSTEM OVERVIEW

To achieve accuracy for fine-grained estimation, we have to conquer the key challenges of the *sparse sensing coverage* & *time-variant sensing coverage*. Fig. 1 shows the adaptive architecture of PGA, which is composed of the mobile vehicular platform and the cloud. When the vehicles move around the city, the sensing hardware on the vehicles collects the air pollution data and transmits them to the cloud to for air pollution estimation. The algorithm run on the cloud consists of four major operational phases: pre-processing, physics guided state evolution estimation, adaptive estimation correction and hypothesis integration.

Pre-Processing Module: According to resolution requirement, the pre-processing module discretizes the temporal and spatial domain into n_x by n_y congruent *sub-areas* (x_i, y_j) and *time slices* t_k . x_i, y_j, t_k correspond

to longitude, latitude and time index respectively. The size of each sub-areas is $\Delta x \times \Delta y$ and $\Delta x = \Delta y = l$. The *sub-areas* and *time slices* combine to form the smallest of the system *unit* (x_i, y_j, t_k) , which defines our pollution map resolution. When the cloud receives data from mobile nodes, the pre-processing module puts them into different units. If multiple data exists in the same unit, the system adopts the average value for further processing. In order to reduce the computational complexity, rather than using a centralized Particle Filter jointly estimating the air pollution of all the *sub-areas*, we associate a Particle Filter to each *sub-area*. It is noticed that the particle filters at different *sub-areas* are not independently calculated. Instead they interact with each other through the physics guided state evolution estimation, since the temporal evolution of air pollution involves air exchange over spatial domain, as shown in Eq. 1.

Physics Guided State Evolution Estimation Module for Sparse Sensing Coverage:

To address the sparse sensing coverage problem, the system takes advantage of the spatial and temporal relations of air dispersion shown in Eq. 1. The physics guided model, based on such relations, allows the system to maintain air pollution field information when few data is collected (sparse sensing coverage). Since it is difficult to predict the next state from the current estimate with this differential equation, we first reformulate Eq. 1 to a predictable form ($C(t) = f(C(t-1))$). Then we integrate the reformulated equation in the Particle Filter structure for state evolution estimation. This module takes the estimation output at *time slice* t_{k-1} as input and applies the reformulated form of the physics guided model to estimate the air pollution at *time slice* t_k . In addition, before the state evolution estimation is conducted at each *time slice*, the environmental factor parameters S are estimated from the collected data.

Adaptive Estimation Correction Module for

Time-Variant Sensing Coverage: Since the areas of sparsity can change over time due to the mobility of the vehicles, the system should use the data model only if it can improve the estimation. Therefore, PGA incorporates an adaptive scheme for estimate correction. There are three major steps. First, due to high required resolution and limited number of mobile nodes, most *sub-areas* do not have real data "observation". Therefore, we must generate data "observation" for estimate correction using the data model. Second, whether the generated "observation" have similar value as ground truth depends on whether there are real collected data nearby (close in proximity and time). When the generated "observation" are near the real collected data, the generated "observation" would be similar to ground truth values, and vice versa. Correcting the estimation with these generated "observation" likely improve the accuracy. Third, since the vehicles are mobile, these sparse locations change and thus PGA must make this adaptation on-the-fly. Therefore, at different *sub-areas* and *time slices*, the system adaptively decides to conduct or skip the estimation correction according to whether or not the correction with generated "observation" improves the performance.

Hypothesis Integration Module: After the adaptive estimate correction, the system gives each *sub-area* a distribution representing the possible values and their weights. Higher weights mean higher possibilities. The distribution not only represents the possible real air pollution value hypothesis but also maintains an assessment of the likely accuracy of the estimate [13]. The final pollution value of each *sub-area* is the weighted sum of the distribution. In addition, the historical estimation results will be conveyed into the physics guided state evolution estimation module for the next state estimation.

ALGORITHM DESCRIPTION

This section provides the algorithm description in the key modules in Fig. 1: (physics guided) state evolution estimation and (adaptive) estimation correction. These two modules are designed to address the two main challenges in this paper: the sparse sensing coverage and time-variant sensing coverage problem respectively.

Reformulating the Physics Guided Model for State Evolution Estimation

The physics guided model we utilize in this paper is convection diffusion equation, which is widely used to describe the transmission of particles energy and other physical quantities including air dispersion [14]. The equation is a good fit for deriving a fine-grained air pollution estimation since it reveals the dynamics of the air evolution and describes the relations of quantities at different times and locations [15]. These relations help to infer air pollution information at the uncovered areas to derive fine-grained air pollution estimation when the sensing coverage is sparse.

The form of the convection diffusion equation is shown in Equation 1. In the equation, $C[\text{kg}/\text{m}^3] = C(x, y, z, t)$ is the gas (air pollutant) concentration at location (x, y, z) and time t . $S[\text{kg}/\text{m}^3\text{s}]$ represents the environmental factor parameters. There are two parameters in the equation: the wind velocity vector $\mathbf{v} = (v_x, v_y)$ $[\text{m}/\text{s}]$ and the diffusion coefficient $\mathbf{K} = \text{diag}(K_x, K_y, K_z)$ $[\text{m}^2/\text{s}]$, which is a diagonal matrix with entries representing the turbulent eddy diffusivities [14].

$$\frac{\partial C}{\partial t} = \nabla \cdot (\mathbf{K} \nabla C) - \nabla \cdot (\mathbf{v} C) + S. \quad (1)$$

In equation 1, the temporal changing rate of air pollution $\frac{\partial C}{\partial t}$ consists of three contributions. The first term $\nabla \cdot (\mathbf{K} \nabla C)$ describes the pollutant flux caused by diffusion. The second term $-\nabla \cdot (\mathbf{v} C)$ describes the pollutant flux caused by the wind. The third contribution S describes the local creation or destruction of the air pollution.

In order to address the challenge of sparse sensing coverage, we introduce extra information from the temporal and spatial relations of air dispersion principle [14, 15]. To achieve this, we reformulate the physics guided model (shown in Eq. 1), to provide a **state evolution estimation**. This state evolution estimation allows the system to infer air pollution values when (and where) the physical sensors is absent. Compared to a data-based approach, the estimation is based on physical laws and limits potential errors, when the data samples are not a fair representation of the underlying air pollution distribution [16]. However, since the original physics guided model Eq. 1 is expressed as a differential form, it must be transferred into a predictable form ($C(t) = f(C(t-1))$) for state evolution estimation:

$$\mathbf{C}[k+1] = \mathbf{A}\mathbf{C}[k] + \mathbf{B}\mathbf{U}[k], \quad (2)$$

where k is the *time slice* index, and $\mathbf{C}[k]$ (n_x by n_y matrix) is the discretized form of $C'(t)$. $\mathbf{U}[k]$ is the discretized form of S . The coefficient matrix $\mathbf{A}$ describes the temporal and spatial correlation of the air pollution concentration while coefficient matrix $\mathbf{B}$ reveals the effect of environmental factor $\mathbf{U}[k]$. Given a *time slice* k , the key parameters to be estimated are the time-variant environmental factor, which is the $\mathbf{U}[k]$ in Eq. 2. In our system, we use the air pollution observations at and before *time slice* k to estimate the $\mathbf{U}[k]$ through convex optimization [17, 18]. After $\mathbf{U}[k]$ is estimated, Eq. 2 is totally decided and can be used to estimate the air pollution values at all *sub-areas*.

It is noticed that the physics guided model only captures the major factors that air pollution evolves over time and space. More detailed factors, such as buildings in the area, small pollution sources like human beings, are difficult to be included. **Therefore, a data-driven model is needed to work with the physics guided model to capture the**

Figure 2: This figure shows detailed architecture of the adaptive correction scheme.

detailed information, especially when sufficient data is collected.

Adaptive Estimation Correction

In order to address the challenge of time-variant sensing coverage, we design an adaptive correction on state evolution estimation. Fig. 2 shows the architecture of the adaptive estimate correction. The results of the state evolution estimation are fed into the *estimate correction* unit for performance improvement. However, at any given time, with many more *sub-areas* than mobile sensing nodes, many *sub-areas* are not sensed. We generate "observation" at these "empty" *sub-areas* with (data driven) "observation" generation unit. However, estimation correction with these generated "observation" does not guarantee improved estimation performances. This is because when the *sub-area* is physically located far from *sub-areas* with real collected data, the generated "observations" at the "empty" *sub-area* can have large difference from true value. As a result, conducting estimation correction at these *sub-areas* and *time slices* is likely to cause performance deterioration.

Therefore, at each *sub-area* and *time slice*, the system first computes and compares the quality of outputs from *physics guided state evolution estimation* module and *data driven*

"observation" generation unit. This is done at *quality computation & comparison* unit. After that, the system decides to choose the outputs either before or after *estimate correction*.

Data Driven "Observation" Generation: Since most *sub-areas* are not sensed at a given *time slice*, "observation" for estimation correction at these *sub-areas* are unavailable. Therefore, in order to make estimation correction possible at all *sub-areas*, we use a data-driven method to generate the "observation" at these "empty" *sub-areas*. Our system adopts an Artificial Neural Network(ANN) for "observation" generation due to its high capability to learn complex non-linear relationship [12, 19]. The details of implementation of the ANN could be found in [5, 12]. After the "observation" are generated, it is possible to correct the estimation at each *sub-area* to get the new weight assignment by [20]

$$p(z_t|y_t) = f_{N,\sigma_s^2}(z_t - y_t^{[i]}), \quad (3)$$

where z_t and y_t represent generated "observation" and state evolution estimation output respectively. f_{N,σ_s^2} is the density probability of data "observation" noise, which is obtained from experimentation. The Eq. 3 needs to be discretized on both temporal and spatial domain. After all weights have been computed, resampling step [20] can take place to reassign the importance of particles according to their similarities to new "observation".

Quality Computation & Comparison: A key problem from sparse and time-variant sensing coverage is that consistently correcting estimations at all *sub-areas* and *time slices* does not necessarily improve performance. This is because at some *sub-areas* and *time slices*, the generated "observations" have very large difference from ground truth values. Using these low quality generated "observation" to correct the estimations only lead to performance deterioration.

As a result, our adaptive correction scheme chooses to correct the estimation or not based on quality comparison to achieve the least error at each *sub-area* and *time slice*.

With adaptive correction, the system has two choices: 1) conducting the correction with the generated "observation" from the data driven model if the quality of the generated "observation" is higher than the quality of physics guided state evolution estimation or 2) skipping the correction after estimation if the quality of the generated data is lower than the quality of physics guided state evolution estimation.

We use entropy of the state evolution estimation distribution to calculate its quality [21, 22]. This is because entropy represents the number of bits required for optimal encoding, which is used as an indication of estimation uncertainty. The quality of generated "observation" can be inferred from the distance between generated "observation" and real sensed data. Usually, "observation" generated by sensed data from farther *sub-areas* cause higher errors. We utilize $E[k]_{i,j}$ and $D[k]_{i,j}$ to denote the above two quality metrics at *sub-area* (i, j) and *time slice* k . After calculating these two quality metrics, the system normalizes them with α to keep them at same order of magnitude for quality comparison. Based on the comparison result, the system chooses to conduct estimation correction if $E[k]_{i,j} - \alpha * D[k]_{i,j} < 0$ or skip the estimate correction if $E[k]_{i,j} - \alpha * D[k]_{i,j} > 0$.

EVALUATION

Experiment Setup

In order to achieve large scale deployment for system evaluation, we deploy the sensing hardware on 29 electric taxis in the city of Shenzhen for 14 days. The sensed area is very sparse within each *time slice* of 1 hour (1% – 15% coverage), although around 100,000 data samples are collected. This is because taxi routes cover overlapping areas and collect redundant information.

Although we do not restrict the motion of the taxis, we

define the sensed area to the central business area, which covers 256 km², from (113.8668°E to 114.1538°E and 22.5150°N to 22.58868°N). This corresponds to the three most important districts of the city: Luo Hu, Fu Tian and Nan Shan. During the deployment, around 26.3 million data samples are collected and transferred to the cloud. We adopt the data from all 5 national air pollution monitoring stations within the experiment area as testing set [23]. These national air pollution monitoring stations publish hourly air pollution data, which is used as testing set of our experiment. The 5 air monitoring stations include: Xi Xiang, Hong Hu, Hua Qiao Cheng, Nan You, Li Yuan, which are located at the north-west, north-east, center, south-west and south-east of the experiment area. For easy representation, we assign id 1 ~ 5 to them. However, these 5 monitoring stations are not enough to evaluate our system which claims high resolution and large scale. To address the problem of lacking official air pollution monitoring station data, at each *time slice*, we also adopt data from 40% of the sensed area data as testing set.

We apply spatial resolution of 500m by 500m, which correlates to block level scale. We set temporal resolution as 1 hour, which is the same as most government run monitoring stations [10]. In order to measure the performance of our method, we defined estimation absolute error $E[k]_{i,j}$ at *sub-area* (i, j) and *time slice* k as

$$E[k]_{i,j} = \left| C_{est}[k]_{i,j} - C_{gt}[k]_{i,j} \right|, \quad (4)$$

where $C_{est}[k]_{i,j}$ and $C_{gt}[k]_{i,j}$ are the derived value and ground truth at *sub-area* (i, j) and *time slice* k respectively. This metric represents deviation between the derived value and ground truth value.

We use the average value of $E[k]_{i,j}$ over *sub-areas* and *time slices* as a measure of system accuracy and its standard deviation as a measure of the system robustness.

Figure 3: The figure shows the hourly air pollution estimation derived by our PGA with the spatial resolution of 500m by 500m.

Smaller average values and standard deviations represent better accuracy and higher robustness respectively.

In order to illustrate how PGA improves performance compared to other existing methods, we adopt the following methods as baselines: **Pure Physics Guided Model (PHY)**: the physics guided model we use in our method, which only relies on the temporal and spatial relations from Eq. 1. **Pure Data Driven Model (ANN)**: the data driven model we use to generate "observations" for the adaptive estimate correction in our method. **AirCloud**: a state-of-the-art method AirCloud [5], which is a combination of two data-driven models.

System Performance

Fig. 3 shows the hourly air pollution estimation derived by our PGA with the spatial resolution of 500m by 500m. We compare our PGA with three baselines on accuracy and stability to validate system design. To illustrate the overall performance of our PGA and three baselines, we plotted the cumulative distribution function (CDF) of average absolute errors in Fig. 4.

Four methods show consistent performance difference. Our PGA performs best. AirCloud ranks second, which is slighter better than ANN. PHY performs worst. PHY's low accuracy due to its bad online parameter estimation which brings a consistent and accumulating bias. AirCloud is better than ANN since AirCloud is a combination of ANN

Figure 4: This figure shows cumulative distribution function (CDF) of absolute errors, using our PGA and three baselines.

and Gaussian Process, which takes advantage of both methods. Our PGA performs best since our adaptive scheme is able to choose the model with better accuracy at most *sub-areas* and *time slices*, which limits the error. When sensing coverage is sparse, our PGA is biased to the physics guided model and when sensing coverage is dense, PGA is biased to the data driven model.

Given absolute error $< 1\text{PPM}$, PHY, ANN and AirCloud only gets 0%, 56% and 56% respectively, while our method gets 89%. Furthermore, all of our estimations are limited to 1.6PPM. In contrast, all three baselines have absolute error $> 3.5\text{PPM}$, which are very large estimation error and may cause residents' panic or ignorance on air pollution, that threatens their health.

We believe that the improvement comes from adaptively integrating the physics guided model with the data-driven model. AirCloud combines two data-driven models, which means that the limitation of the data-driven model when the data is sparse cannot be relieved. On the other hand, our PGA introduces the physics guided model to handle **sparse sensing coverage**. In addition, the designed adaptive

scheme makes the physics guided model and the data-driven model cooperatively to address the problem of **time-variant sensing coverage**. As a result, PGA achieves $4\times$ reduction in the error $< 1\text{PPM}$ compared to the AirCloud, which is a significant improvement.

CONCLUSION

This paper presents PGA, a physics guided and adaptive approach to derive fine-grained ($500m$ by $500m$ by $1hour$) air pollution estimation. The system is designed based on Particle Filter structure since it offers the ability to combine the advantages of a physics guide model and data driven model. 1) To address challenge of sparse sensing coverage, the system adopts the temporal and spatial relations of air dispersion from a physics guided model to extract more information for estimation. The physics guided model is reformulated as the state evolution estimation model in Particle Filter structure. 2) To address the dynamic sensing coverage challenge, the system adaptively trusts either the result directly from the physics guided model or the result after correction through "observations" generated from the data driven model. The choice is made based on quality comparison of the estimation by the physics guided model and generated "observations" by the data driven model to optimize performance for each *sub-area* and each *time slice*. To evaluate the system, we deploy 29 taxis in the city of Shenzhen for 14 days, which collect around 26.3 million data samples. The evaluation results show our system achieves up to $4.0\times$ reduction on average error compared to state-of-the-art approaches.

ACKNOWLEDGEMENTS

This project is funded in part by National Science Foundation #1149611 and Carnegie Mellon University's Mobility21 National University Transportation Center (grant number 69A3551747111), which is sponsored by the US Department of Transportation.

REFERENCES

1. BBC World News. Air pollution causes 467,000 premature deaths a year in europe. <http://www.bbc.com/news/world-europe-38078488>.
2. The New York Times Energy and Environment. Study links 6.5 million deaths each year to air pollution. <https://www.nytimes.com/2016/06/27/business/energy-environment/study-links-6-5-million-deaths-each-year-to-air-pollution.html>.
3. World Health Organization. Who releases country estimates on air pollution exposure and health impact. <http://www.who.int/mediacentre/news/releases/2016/air-pollution-estimates/en/>.
4. Noel De Nevers. *Air pollution control engineering*. Waveland Press, 2010.
5. Yun Cheng, Xiucheng Li, Zhijun Li, Shouxu Jiang, Yilong Li, Ji Jia, and Xiaofan Jiang. Aircloud: a cloud-based air-quality monitoring system for everyone. In *Proceedings of the 12th ACM Conference on Embedded Network Sensor Systems*, pages 251–265. ACM, 2014.
6. David Hasenfratz, Olga Saukh, Christoph Walser, Christoph Hueglin, Martin Fierz, Tabita Arn, Jan Beutel, and Lothar Thiele. Deriving high-resolution urban air pollution maps using mobile sensor nodes. *Pervasive and Mobile Computing*, 16:268–285, 2015.
7. JC Weil and RP Brower. An updated gaussian plume model for tall stacks. *Journal of the Air Pollution Control Association*, 34(8):818–827, 1984.
8. Zahari Zlatev and Ivan Dimov. *Computational and numerical challenges in environmental modelling*, volume 13. Elsevier, 2006.
9. Paolo Zannetti. *Air Pollution Modeling: Theories, computational methods and available software*. Springer Science & Business Media, 2013.
10. Yu Zheng, Furui Liu, and Hsun-Ping Hsieh. U-air: When urban air quality inference meets big data. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1436–1444. ACM, 2013.
11. Yu Zheng, Xuxu Chen, Qiwei Jin, Yubiao Chen, Xiangyun Qu, Xin Liu, Eric Chang, Wei-Ying Ma, Yong Rui, and Weiwei Sun. A cloud-based knowledge discovery system for monitoring fine-grained air quality. *MSR-TR-2014-40, Tech. Rep.*, 2014.
12. Yidan Hu, Guojun Dai, Jin Fan, Yifan Wu, and Hua Zhang. Blueaer: A fine-grained urban pm_{2.5} 3d monitoring system using mobile sensing. In *Computer Communications, IEEE INFOCOM 2016-The 35th Annual IEEE International Conference on*, pages 1–9. IEEE, 2016.
13. Pei Zhang and Margaret Martonosi. Locale: Collaborative localization estimation for sparse mobile sensor networks. In *Information Processing in Sensor Networks, 2008. IPSN'08. International Conference on*, pages 195–206. IEEE, 2008.
14. S Pal Arya. *Air pollution meteorology and dispersion*. Oxford University Press New York, 1999.
15. John M Stockie. The mathematics of atmospheric dispersion modeling. *Siam Review*, 53(2):349–372, 2011.
16. Chao Zhang, Keyang Zhang, Quan Yuan, Haoruo Peng, Yu Zheng, Tim Hanratty, Shaowen Wang, and Jiawei Han. Regions, periods, activities: Uncovering urban dynamics via cross-modal representation learning. In *Proceedings of the 26th International Conference on*

- World Wide Web*, pages 361–370. International World Wide Web Conferences Steering Committee, 2017.
17. Michael Grant, Stephen Boyd, and Yinyu Ye. Cvx: Matlab software for disciplined convex programming, 2008.
 18. Michael C Grant and Stephen P Boyd. Graph implementations for nonsmooth convex programs. In *Recent advances in learning and control*, pages 95–110. Springer, 2008.
 19. Sun-Chong Wang. Artificial neural network. In *Interdisciplinary computing in java programming*, pages 81–100. Springer, 2003.
 20. S. Thrun, W. Burgard, and D. Fox. *Probabilistic Robotics*. MIT Press, 2006.
 21. Claude Elwood Shannon. A mathematical theory of communication. *ACM SIGMOBILE Mobile Computing and Communications Review*, 5(1):3–55, 2001.
 22. Xinlei Chen, Aweek Purohit, Carlos Ruiz Dominguez, Stefano Carpin, and Pei Zhang. Drunkwalk: Collaborative and adaptive planning for navigation of micro-aerial sensor swarms. In *Proceedings of the 13th ACM Conference on Embedded Networked Sensor Systems*, pages 295–308. ACM, 2015.
 23. China national air pollution monitoring station history data publishing. <http://beijingair.sinaapp.com>.