Mobile IoT Solutions to Air Quality Monitoring

Prof. You-Chiun Wang

Department of Computer Science and Engineering
National Sun Yat-sen University, Taiwan

Dec. 28, 2021

Outline

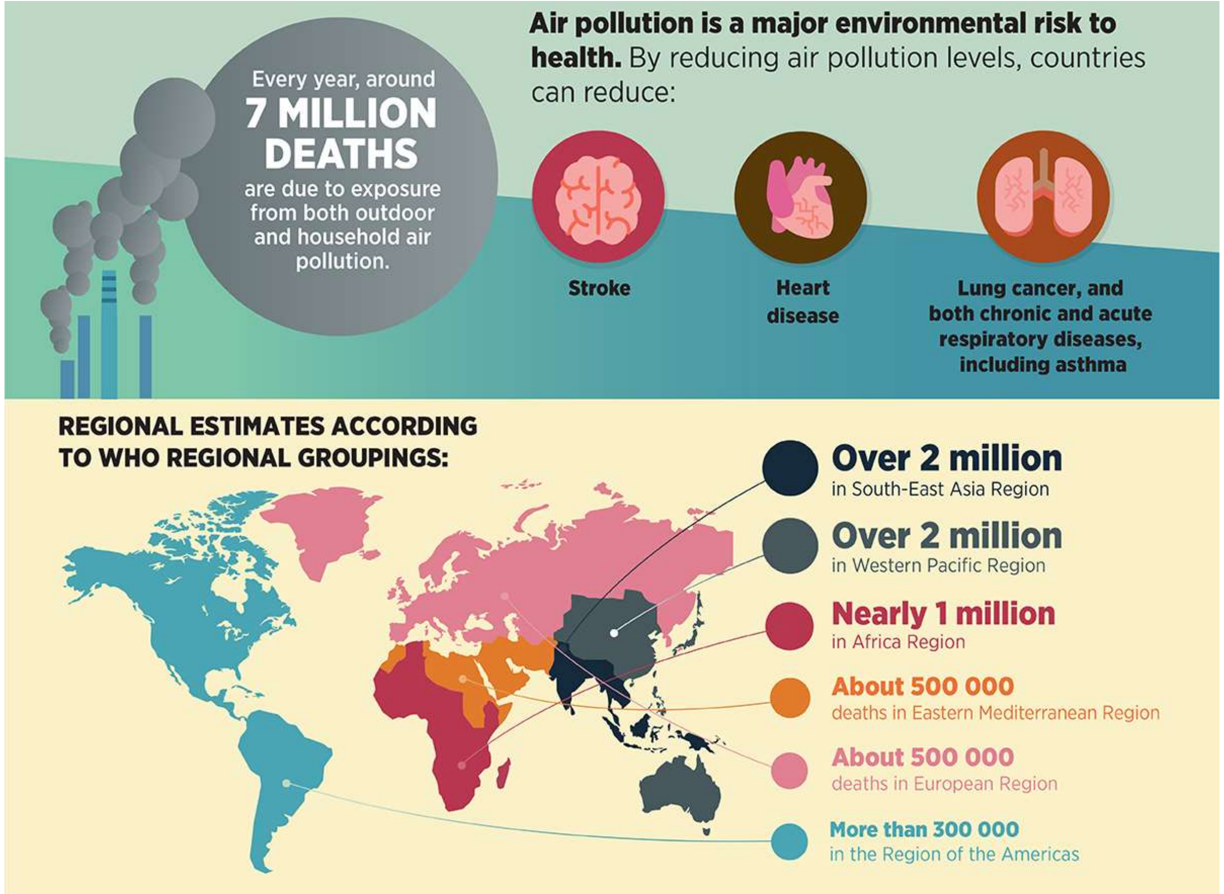
National
Sun Yat-sen
University

- Introduction
- Evaluation of air quality
- Collection of air quality
- Analysis of raw data
- Adjustment of reporting rates
- Research challenges
- Conclusion

Introduction

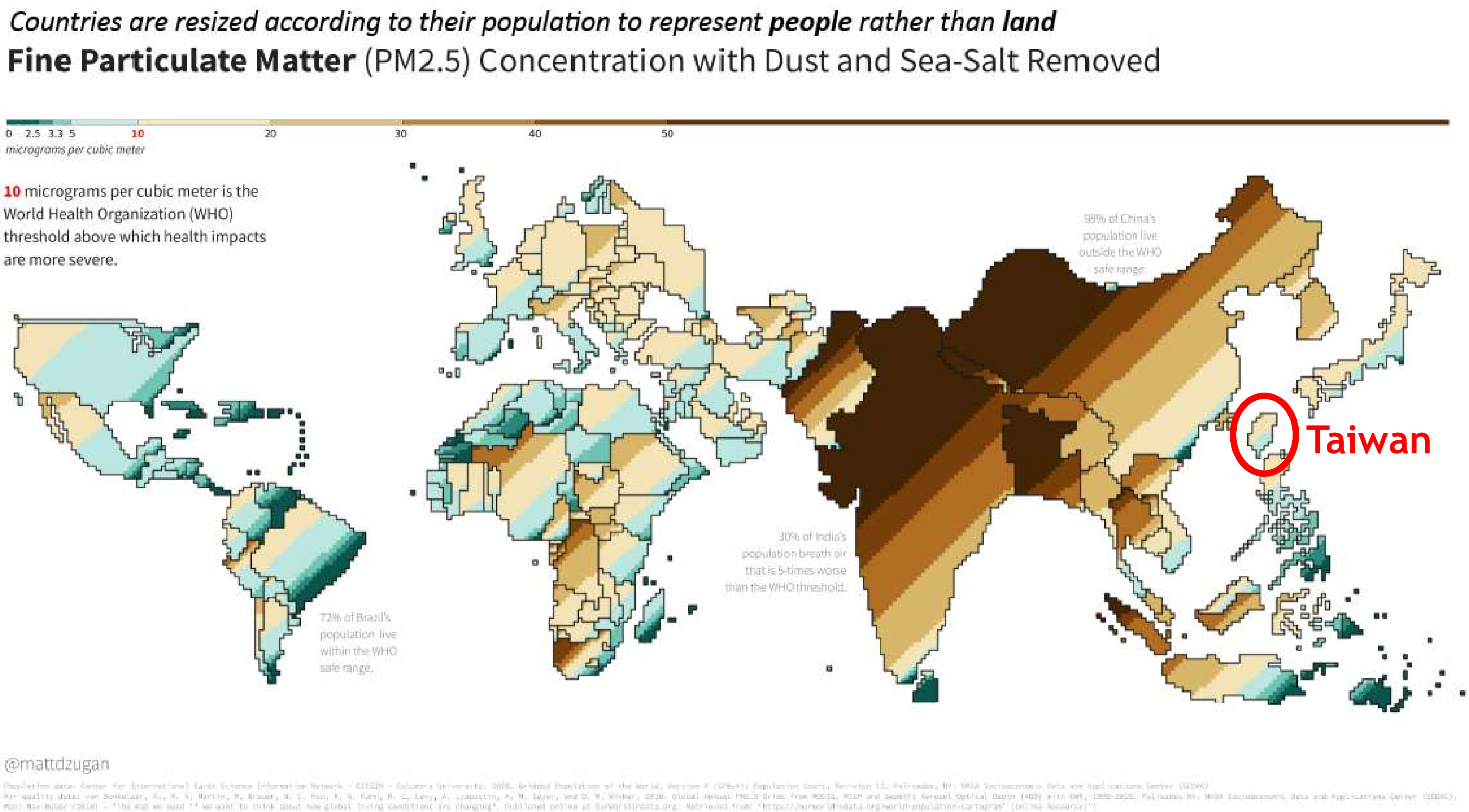
National
Sun Yat-sen
University

- WHO warns that air pollution has been one of the most serious *environmental health risks* in the world.

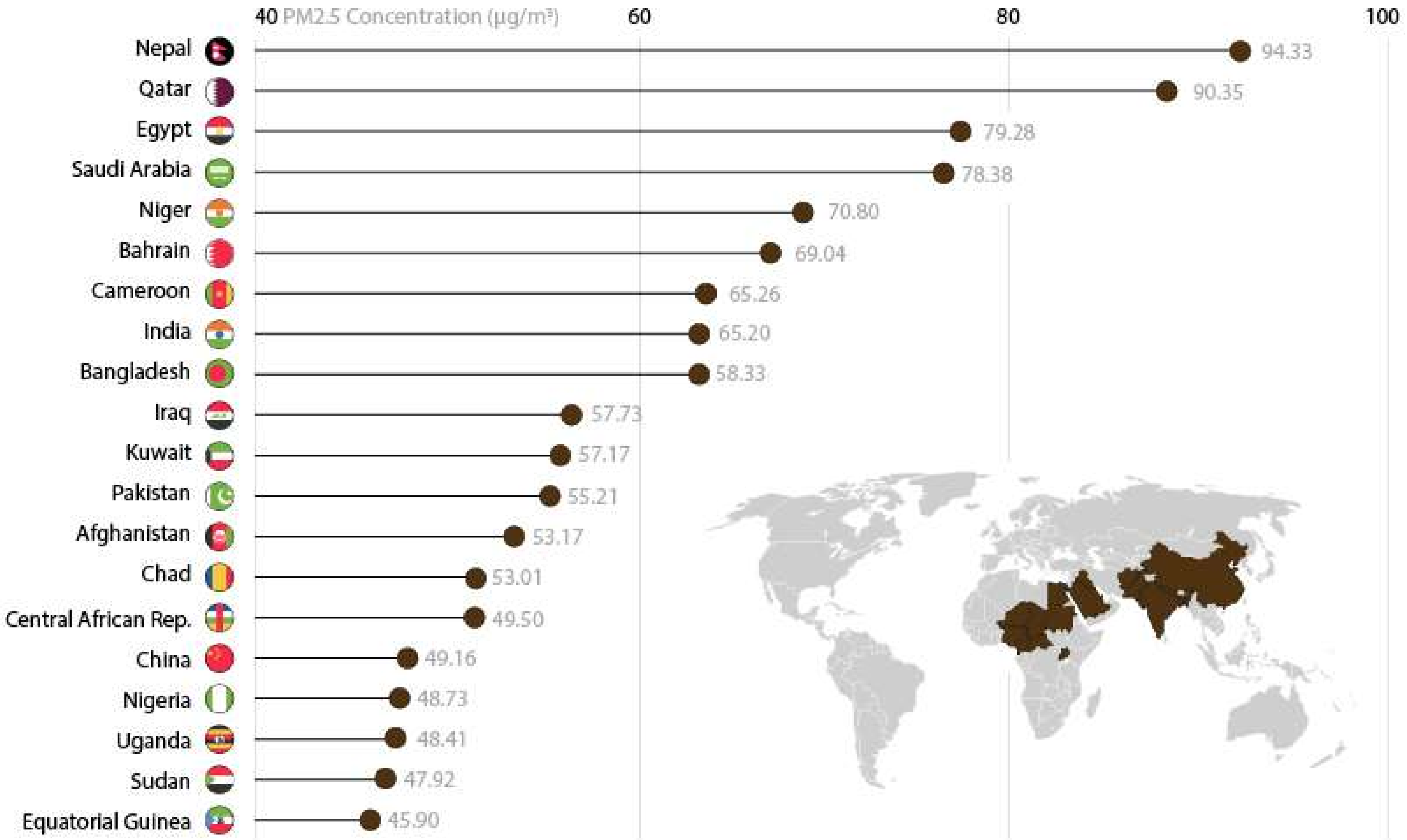


PM 2.5 in the world

National
Sun Yat-sen
University



Countries with the most polluted air

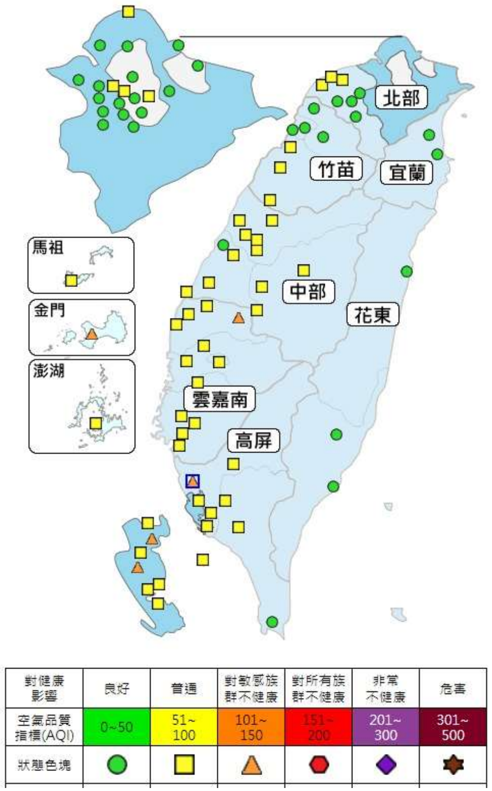


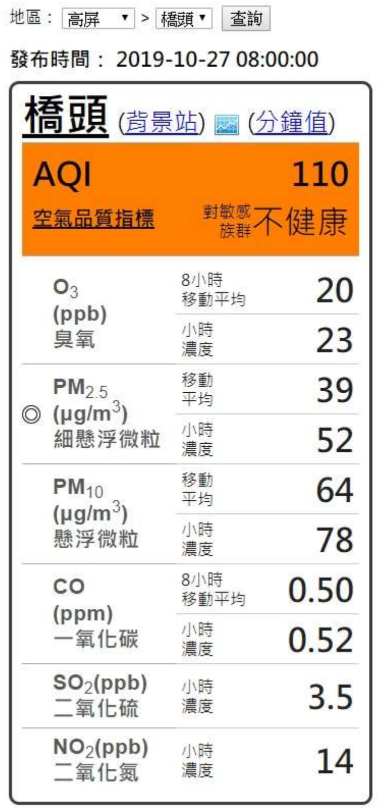
Air quality concentration is measured in terms of micrograms per cubic metre (µg/m³).
Source: World Health Organization, 2016 (Latest data available)

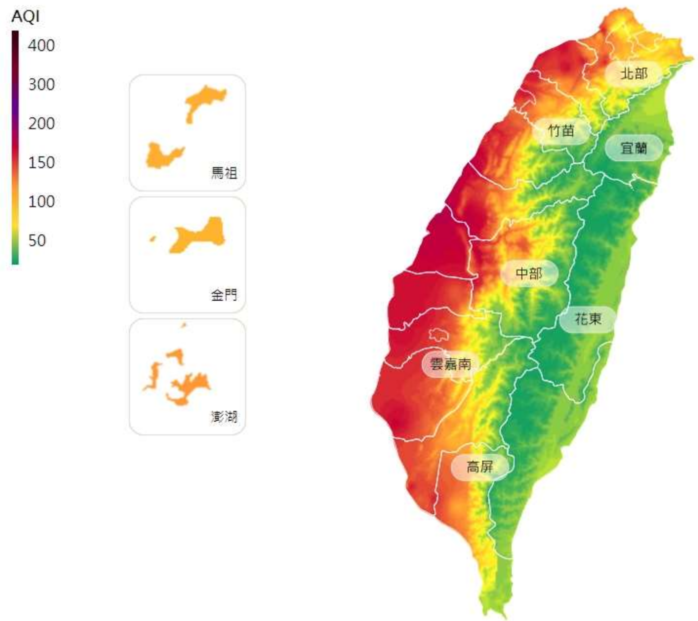
Traditional solutions to air quality monitoring

National
Sun Yat-sen
University

- Governments install some large, expensive monitoring stations on the **dedicated** locations in cities.
  - These stations offer *large-scale* monitoring of air quality.



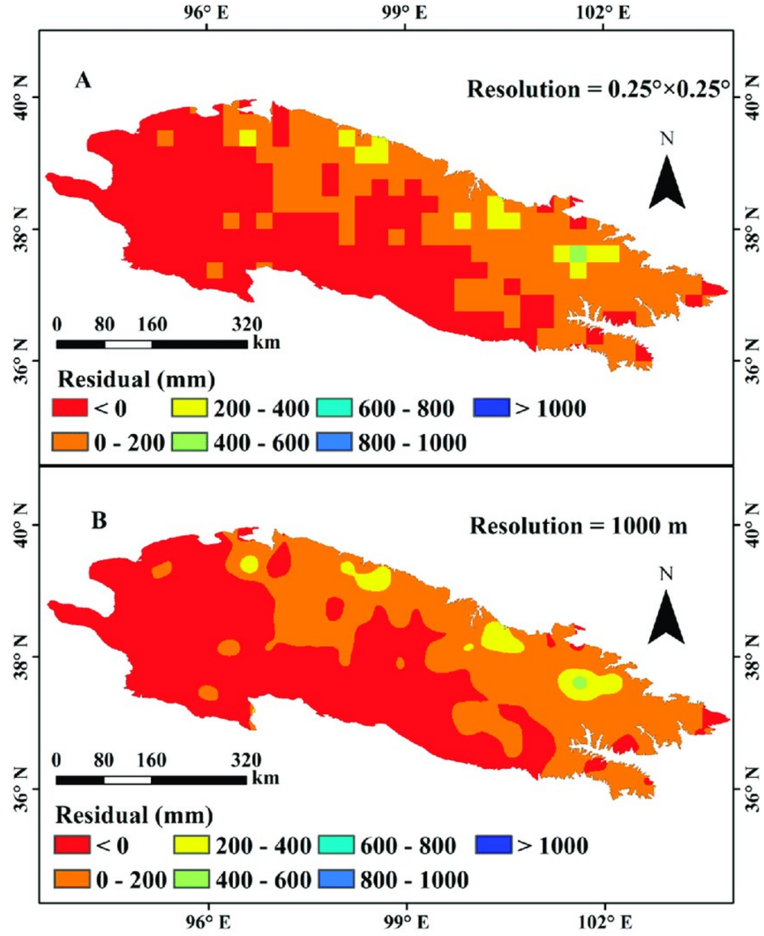




What is the problem?

National Sun Yat-sen University

- Static monitoring stations may provide relatively few samples of air-quality, results in a **coarse** resolution.
- It lacks **flexibility** to use static monitoring stations, as some sites chosen to install stations (e.g., *displaced plants* or *new parks*) may become redundant due to the development of a city.



Mobile IoT solutions

National
Sun Yat-sen
University

- Collect air quality by citizens
  - Pedestrian-based solutions
  - Bike-based solutions
  - Car-based solutions
- Analyze raw sensing data
- Adjust reporting rates

Outline

National
Sun Yat-sen
University

- Introduction
- Evaluation of air quality
  - Air quality index (AQI)
  - Air pollution dispersion models
- Collection of air quality
- Analysis of raw data
- Adjustment of reporting rates
- Research challenges
- Conclusion

AQI (1/2)

AIR QUALITY INDEX

National Sun Yat-sen University

- AQI gives an intelligible **index** to measure and report to the public *how clean or polluted the air is* during one day.

US AQI Level		PM2.5 (μg/m³)	Health Recommendation (for 24 hour exposure)
Good	0-50	0-12.0	Air quality is satisfactory and poses little or no risk.
Moderate	51-100	12.1-35.4	Sensitive individuals should avoid outdoor activity as they may experience respiratory symptoms.
Unhealthy for Sensitive Groups	101-150	35.5-55.4	General public and sensitive individuals in particular are at risk to experience irritation and respiratory problems.
Unhealthy	151-200	55.5-150.4	Increased likelihood of adverse effects and aggravation to the heart and lungs among general public.
Very Unhealthy	201-300	150.5-250.4	General public will be noticeably affected. Sensitive groups should restrict outdoor activities.
Hazardous	301+	250.5+	General public at high risk of experiencing strong irritations and adverse health effects. Should avoid outdoor activities.

AQI (2/2)

National Sun Yat-sen University

- AQI is measured based on the following pollutants:

Air Quality Index(AQI)							
AQI	O_3 (ppm) 8hr	O_3 (ppm) 1-hr(1)	$PM_{2.5}$ ($\mu g/m^3$) 24-hr	PM_{10} ($\mu g/m^3$) 24-hr	CO (ppm) 8-hr	SO_2 (ppb) 1-hr	NO_2 (ppb) 1-hr
Good 0～50	0.000 - 0.054	-	0.0 - 15.4	0 - 54	0 - 4.4	0 - 35	0 - 53
Moderate 51～100	0.055 - 0.070	-	15.5 - 35.4	55 - 125	4.5 - 9.4	36 - 75	54 - 100
Unhealthy for Sensitive Groups 101～150	0.071 - 0.085	0.125 - 0.164	35.5 - 54.4	126 - 254	9.5 - 12.4	76 - 185	101 - 360
Unhealthy 151～200	0.086 - 0.105	0.165 - 0.204	54.5 - 150.4	255 - 354	12.5 - 15.4	186 - 304(3)	361 - 649
Very Unhealthy 201～300	0.106 - 0.200	0.205 - 0.404	150.5 - 250.4	355 - 424	15.5 - 30.4	305 - 604(3)	650 - 1249
Hazardous 301～400	(2)	0.405 - 0.504	250.5 - 350.4	425 - 504	30.5 - 40.4	605 - 804(3)	1250 - 1649
Hazardous 401～500	(2)	0.505 - 0.604	350.5 - 500.4	505 - 604	40.5 - 50.4	805 -1004(3)	1650 - 2049

Outline

National
Sun Yat-sen
University

- Introduction
- Evaluation of air quality
  - Air quality index (AQI)
  - Air pollution dispersion models
- Collection of air quality
- Analysis of raw data
- Adjustment of reporting rates
- Research challenges
- Conclusion

Air pollution dispersion

National
Sun Yat-sen
University

- Air pollution is caused by the emission of pollutants such as *particulates* or *harmful gases* to the atmosphere from some **sources**.

Point source

Volume source

Inputs to a dispersion model

National
Sun Yat-sen
University

Meteorology
- Wind speed
- Wind direction
- Temperature

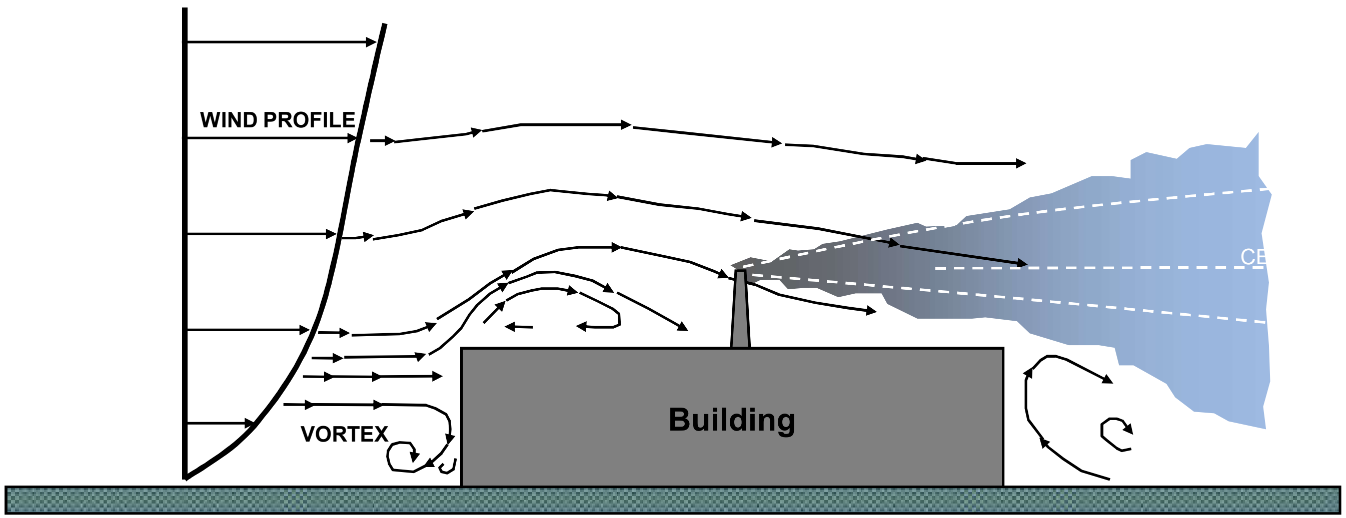
Terrain & Land Use
- Rural or urban
- Elevations

Emissions
- Rates
- Hrs of operation

Sources
- Source type
- Parameters

Downwash
- Buildings

Gravitational Settling
- Particle Size
- Density

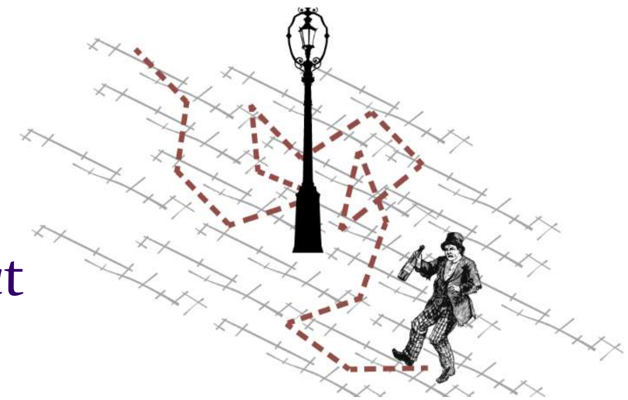


Dispersion models (1/3)

National
Sun Yat-sen
University

- **Box model:** (for a simplified environment)
  - Pollutants are *homogeneously distributed* in a **box-shaped** space. It finds the *average concentration* of pollutants.

- **Lagrangian model:**
  - Motion of particles follows a **random-walk** model.
  - Use a *mobile reference system* to predict trajectories of particles.



  - Dispersion of air pollution is estimated based on **statistics** of moving trajectories caused by a great deal of particles.

Dispersion models (2/3)

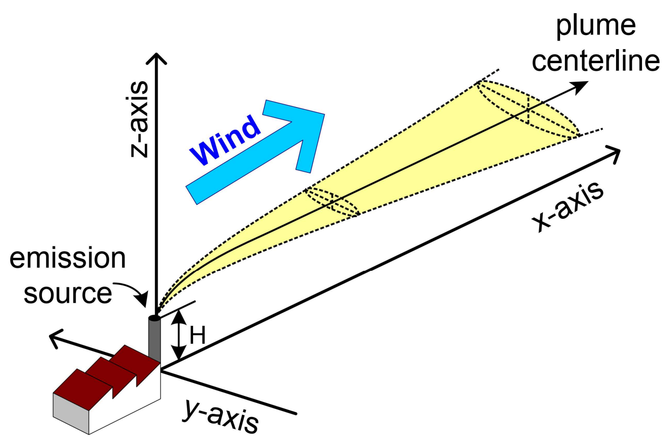
- **Eulerian model:** (variation of the Lagrangian model)
  - Use a *3D Cartesian coordinate system* to track trajectories of particles when they leave from their initial positions.

- **Dense-gas model:**
  - Simulate diffusion of gas pollution plumes which are **heavier** than the general air (usually *toxic gases*).
  - When a stream of dense gas is injected into the flowing air, it may produce a *wide and flat plume* at the ground level.

Dispersion models (3/3)

National Sun Yat-sen University

- **Gaussian model:**
  - Dissemination of pollutants has a **normal** probability distribution. It is used to estimate the diffusion of *continuous, floating* air pollution plumes from ground-level or elevated emission sources.
  - Industrial source complex (**ISC3**):
    Developed from the Gaussian model to evaluate both *diffusion* and *sedimentation* of pollutants in the air



Outline

National
Sun Yat-sen
University

- Introduction
- Evaluation of air quality
- Collection of air quality
  - Pedestrian-based solutions
  - Bike-based solutions
  - Car-based solutions
- Analysis of raw data
- Adjustment of reporting rates
- Research challenges
- Conclusion

CitiSense (1/3)

National Sun Yat-sen University

- In CitiSense, pedestrians carry *smart phones* and *wearable sensor boards* to monitor air quality in a day, especially during times when exposure to air pollutants will be the highest (e.g., a **rush-hour** commute)
  - Sensor board contains CO, NO_2, and O_3 detectors, and also *weather-related sensing devices* to monitor temperature, humidity, and barometric pressure.
  - Sensor board can talk with a smart phone via **Bluetooth**.

N. Nikzad et al., "CitiSense: Improving geospatial environmental assessment of air quality using a wireless personal exposure monitoring system," *Proc. ACM Conf. Wireless Health*, 2012, pp. 1–8.

CitiSense (2/3)

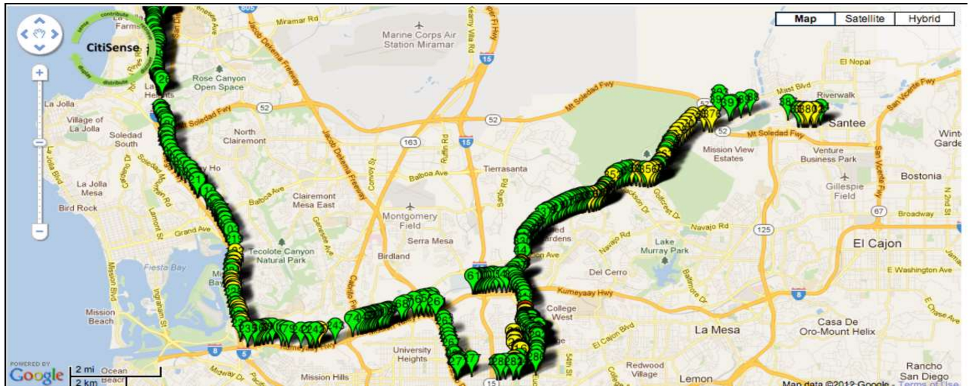
National
Sun Yat-sen
University

- **Android**-based APP is developed to *display air quality* on the phone and help the user *post data to social networks* such as facebook and twitter.

CitiSense (1/2)

National
Sun Yat-sen
University

- By **Google Maps**, users can browse through the readings of air quality they collected on any given day.
  - Each sensor reading is shown as a *color-coded and numbered marker*, ordered by its monitoring time.



  - CitiSense was deployed in San Diego, USA.
    - 16 participators each had a commute journey for at least 20mins, and they were users of the same social network.

Yang's work (1/3)

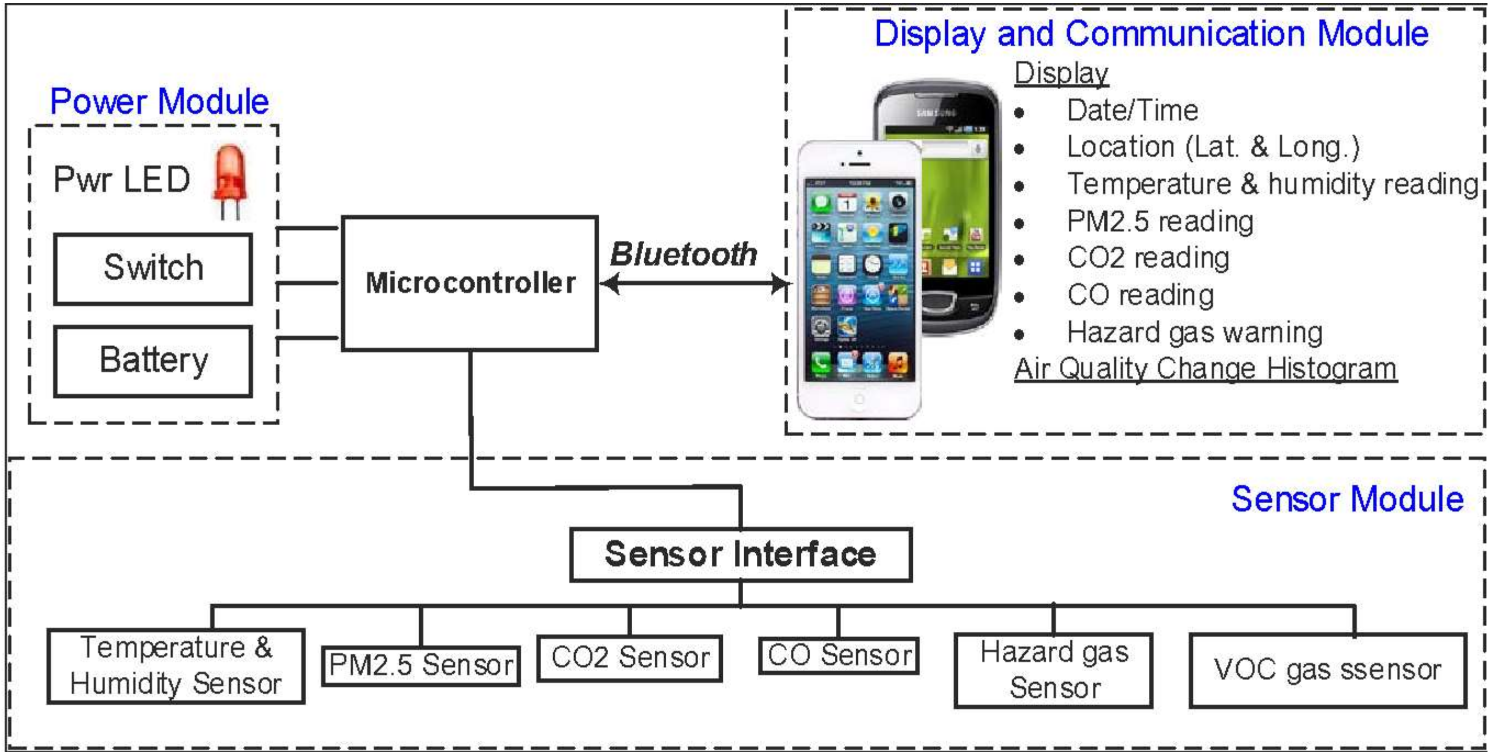
National
Sun Yat-sen
University

- Each pedestrian carries a *smart phone* and a box of sensors (called a *sensor unit*) to detect air pollutants.
  - Smart phone serves as the **middleware** between the sensor unit and the server.
  - When a user wants to measure air quality, he/she can execute a specific APP installed on the smart phone, which *triggers the sensor unit* to detect air pollutants and report sensing data.
  - Server then manages the collected data and present the monitoring result of air quality through *a map-based interface*.

Y. Yang and L. Li, "A smart sensor system for air quality monitoring and massive data collection," *Proc. Int'l Conf. Information and Comm. Technology Convergence*, 2015, pp. 147–152.

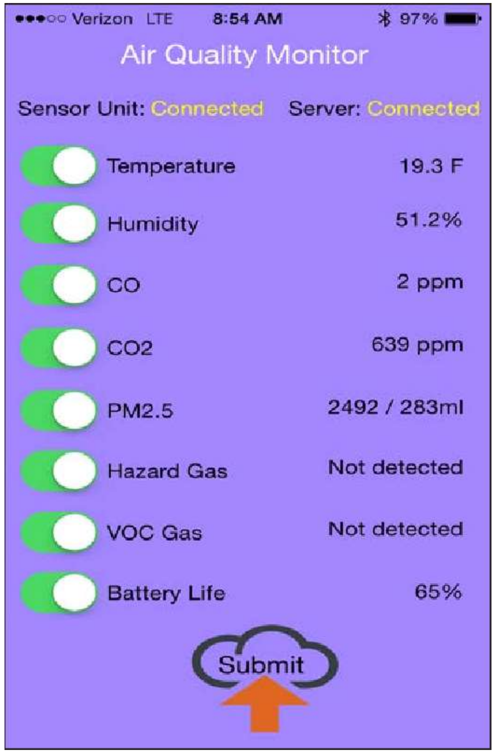
Yang's work (2/3)

- System architecture:
  - VOC (volatile organic compound) sensor: Detect acetone, alcohol, formaldehyde, methanol, nitrogen, styrene, sulfur, and toluene in the air.



Yang's work (3/3)

- APP interface:



- A prototype of the proposed system was demonstrated in Prairie View, Texas, USA.

AirSense (1/3)

National Sun Yat-sen University

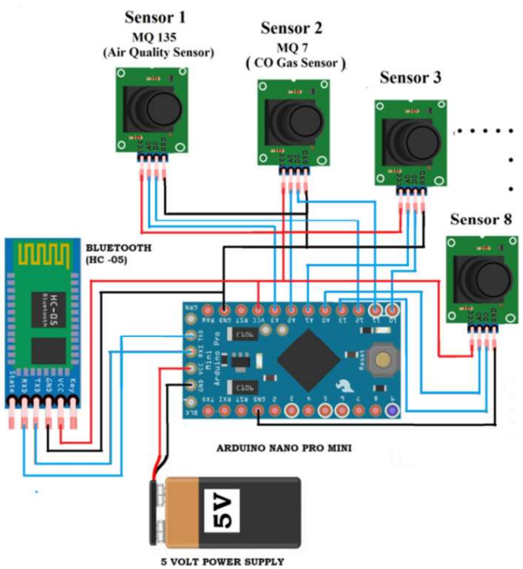
- AirSense contains four tiers:
  - *Crowd sensing tier:* Users carry sensor suites and smart phones to collect data. They can get the analyzed data from smart phones.
  - *Air quality sensing tier:* Each sensor suite sends data to a nearby phone by **Bluetooth**.
    - Sensor suite reports its data only when there is a *significant change* in the readings to save the Bluetooth bandwidth.
  - *Data forwarding tier:* Sensing data collected by smart phones will be sent to a cloud server via 4G or Wi-Fi.
  - *Data analysis tier:* The cloud server finally calculates the AQI value based on the collected data.
    - It can also construct a *pollution footprint* for each participator, which is shown on the smart phone in the form of a mobile APP.

J. Dutta et al., "Towards smart city: sensing air quality in city based on opportunistic crowdsensing," *Proc. ACM Int'l Conf. Distributed Computing and Networking*, 2017, pp. 1–6.

National
Sun Yat-sen
University

AirSense (2/3)

- AirSense implementation:
  - Sensor suite is implemented by using an **Arduino** board to integrate with *PM, O_3, NO_2, SO_2 sensors* and a *Bluetooth module*.
  - **OpenShift** is used as a cloud server, which adopts MySQL as the database to store data, where *ID of each sensor suite* is selected as a primary key to query data.



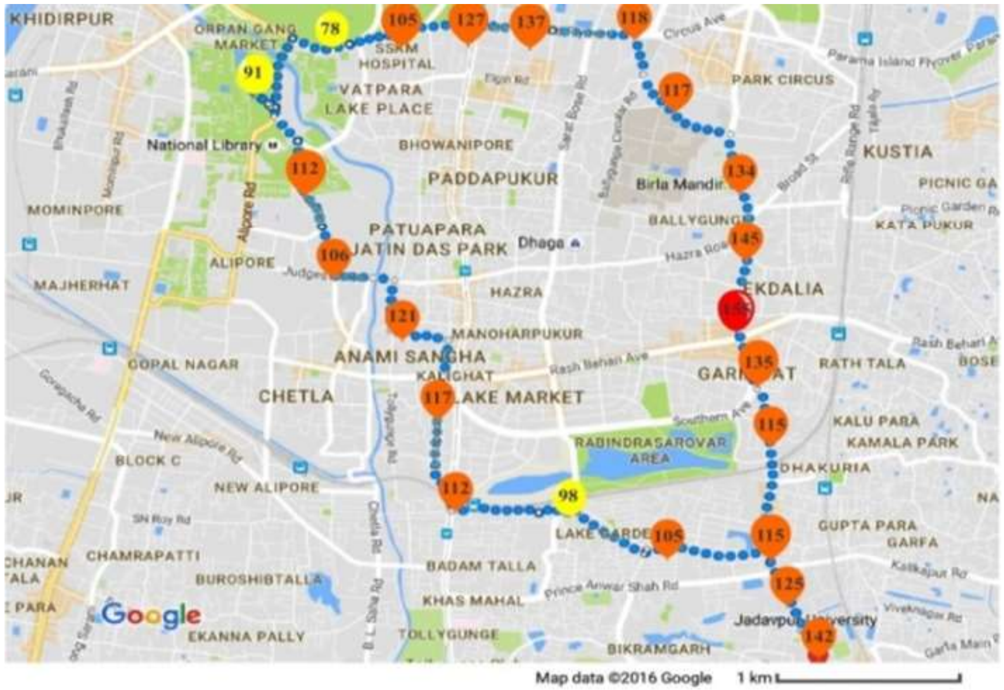
Sensor suite

OpenShift cloud server

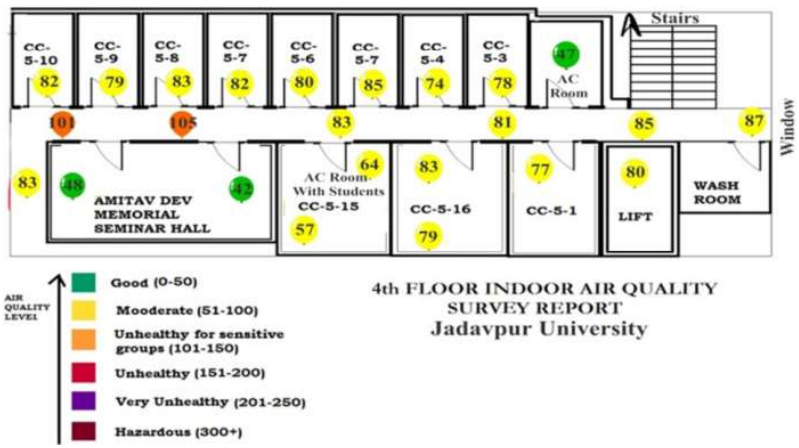
AirSense (3/3)

National
Sun Yat-sen
University

- AirSense was deployed in Kolkata, India for demonstration.



Pollution footprint



AQI map (outdoor/indoor)

Outline

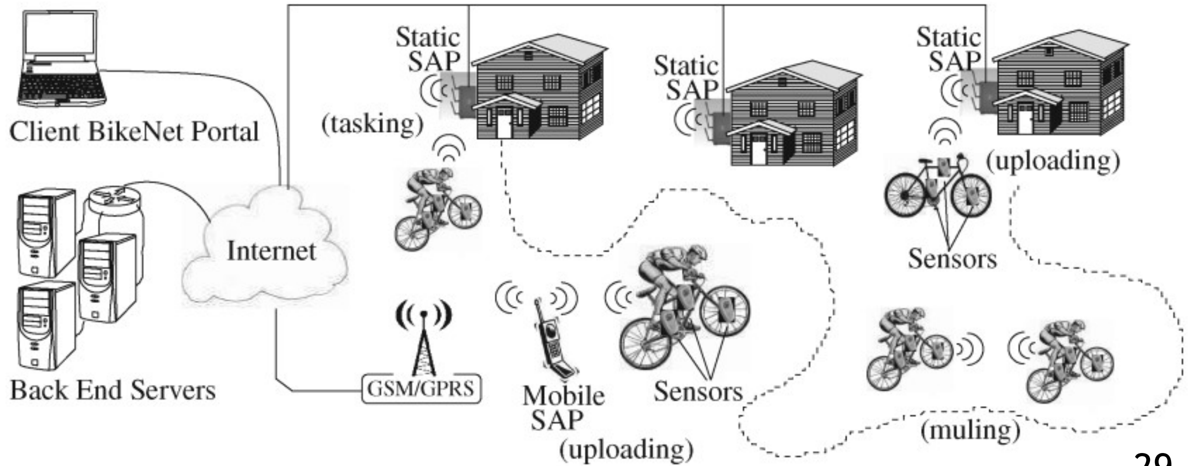
National
Sun Yat-sen
University

- Introduction
- Evaluation of air quality
- Collection of air quality
  - Pedestrian-based solutions
  - Bike-based solutions
  - Car-based solutions
- Analysis of raw data
- Adjustment of reporting rates
- Research challenges
- Conclusion

BikeNet (1/2)

National
Sun Yat-sen
University

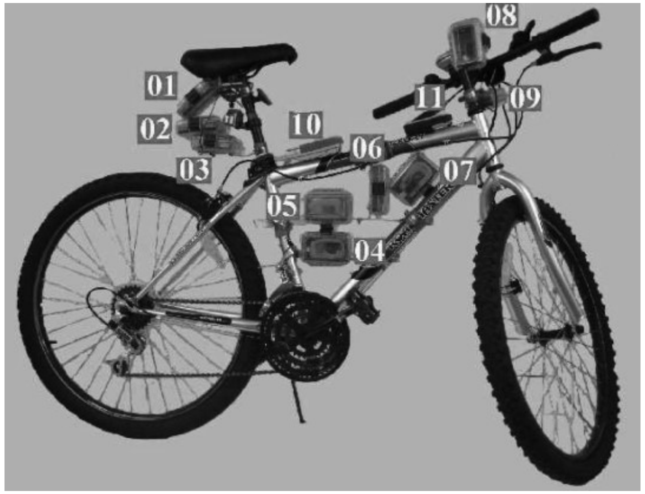
- BikeNet uses multiple sensors installed on a bike to gather *quantitative data* related to the ride of a cyclist.
  - Cyclist's **performance**: Riding speed, distance traveled, and calories burned by the cyclist.
  - Environmental conditions: **Air pollution**, allergen, noise, and terrain roughness of the given route.
  - *CO_2 map of streets* where cyclists ever rode through is shown on the web portal.

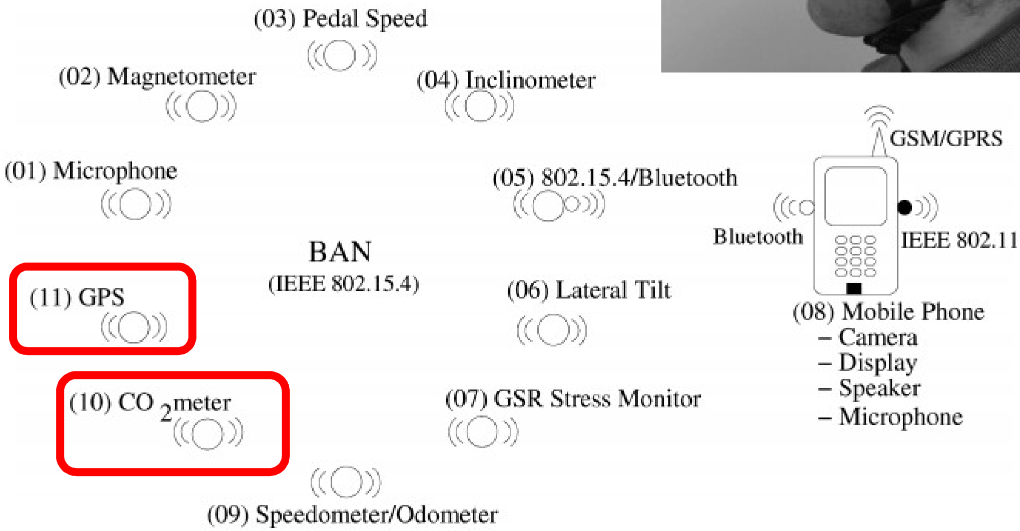


S.B. Eisenman et al., "BikeNet: A mobile sensing system for cyclist experience mapping," *ACM Trans. Sensor Networks*, vol. 6, no. 1, pp. 6–39, 2009.

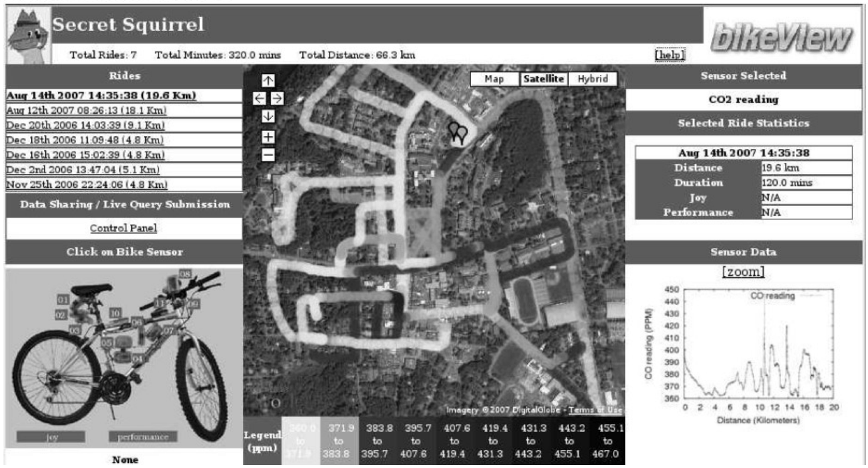
BikeNet (2/2)

- Each bike carries the following devices:



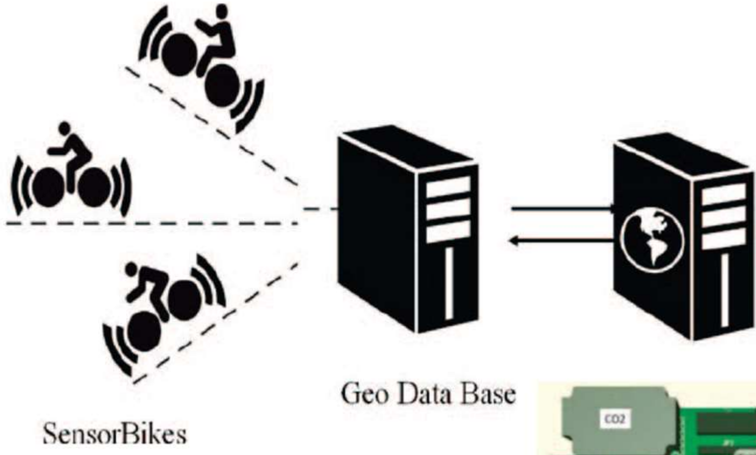


- CO_2 map of streets: Handover, New Hampshire, USA

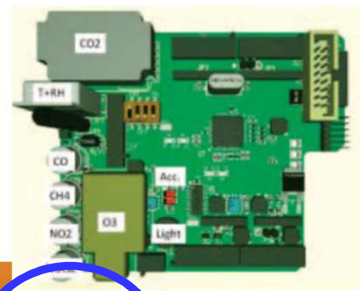


SensorWebBike (1/2)

- SensorWebBike contains three major components:
  - SensorBikes
  - Geo database
  - Web server



- SensorBike:
  - Use **Arduino** board to integrate sensors.
  - Air pollutants: CO, CO_2, O_3, NO_2, CH_4
  - Weather-related data: Noise, humidity, temperature

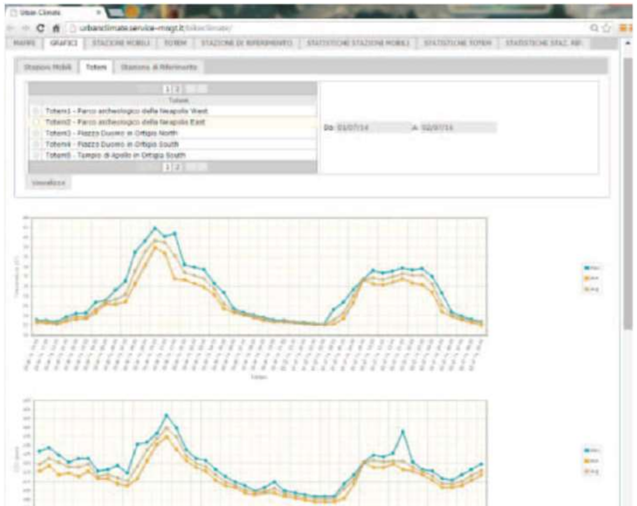


C. Vagnoli et al., “The SensorWebBike for air quality monitoring in a smart city,” *Proc. IET Conf. Future Intelligent Cities*, 2014, pp. 1–4.

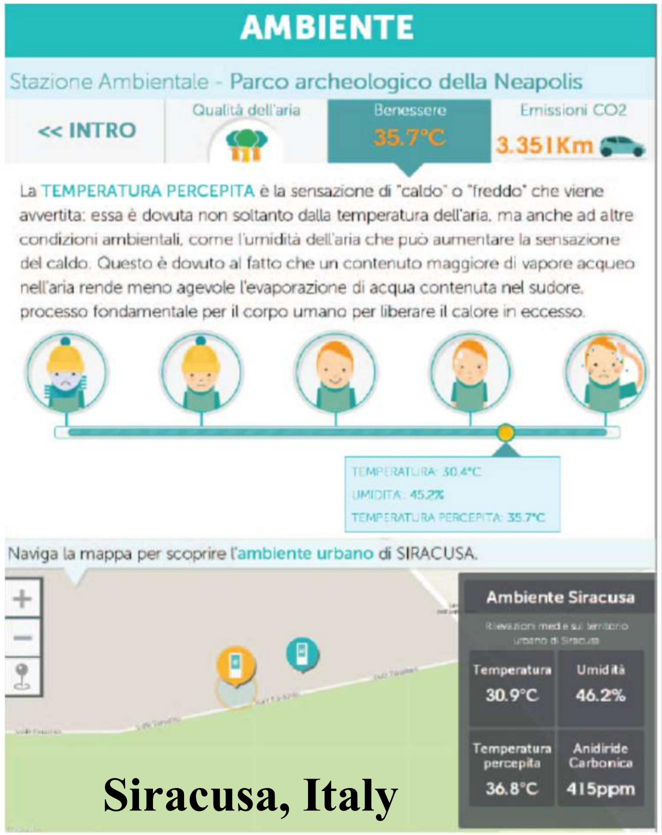
SensorWebBike (2/2)

National Sun Yat-sen University

- Geo database:
  - Store and manage sensing data reported by SensorBikes.
  - Follow the **OGC** (open geospatial consortium) data format for *inter-operability*.
- Web server: Let users view, query, and analyze air quality.



Visualization of data (by Java2EE)



Siracusa, Italy

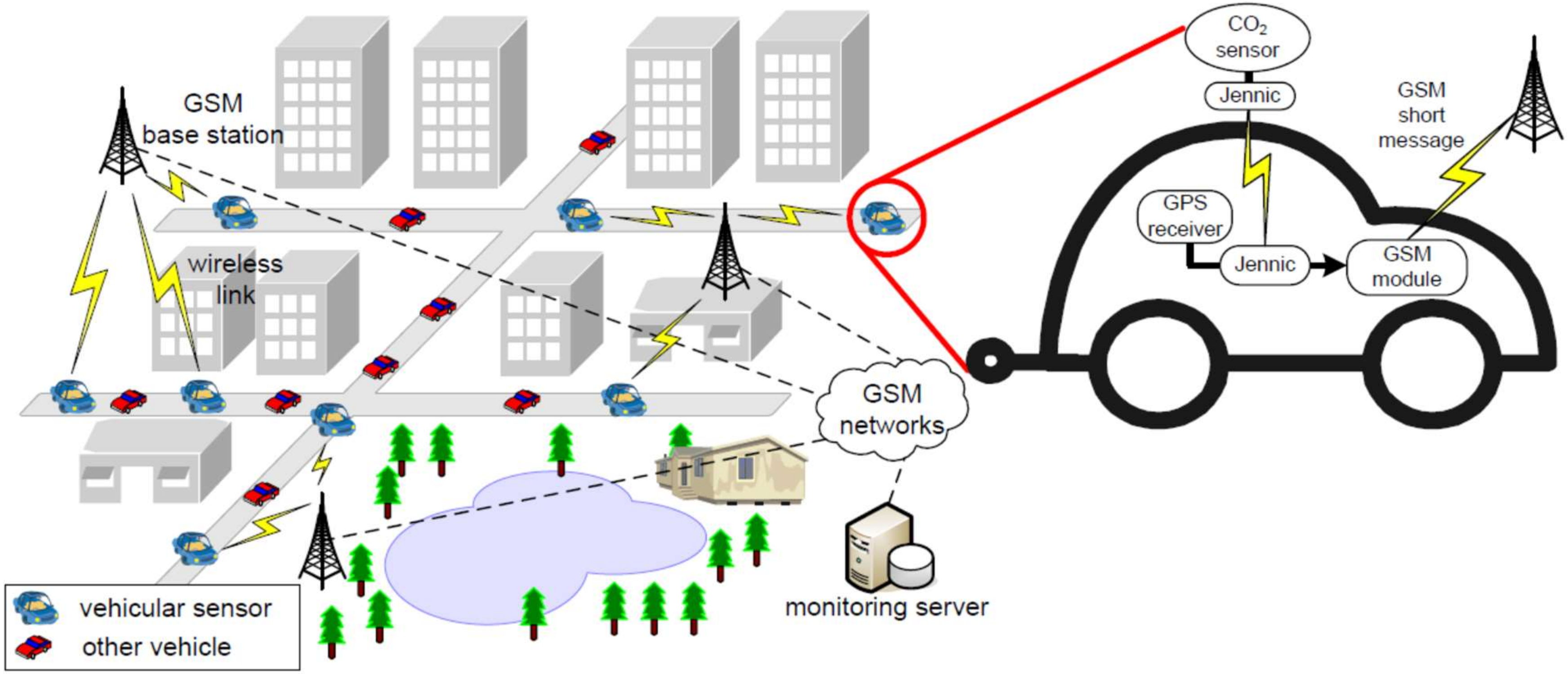
Outline

National
Sun Yat-sen
University

- Introduction
- Evaluation of air quality
- Collection of air quality
  - Pedestrian-based solutions
  - Bike-based solutions
  - Car-based solutions
- Analysis of raw data
- Adjustment of reporting rates
- Research challenges
- Conclusion

Hu's work (1/2)

- Hu et al. develop a mobile IoT solution by cars to monitor *CO_2 concentration* in urban areas.

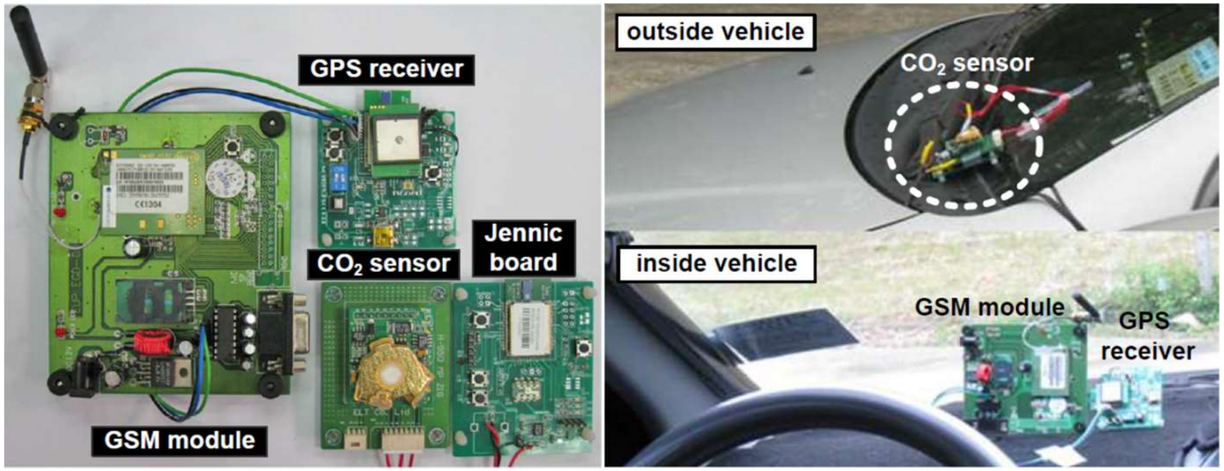


S.C. Hu, Y.C. Wang, C.Y. Huang, and Y.C. Tseng, "A vehicular wireless sensor network for CO2 monitoring," *Proc. IEEE Conf. Sensors*, 2009, pp. 1498–1501.

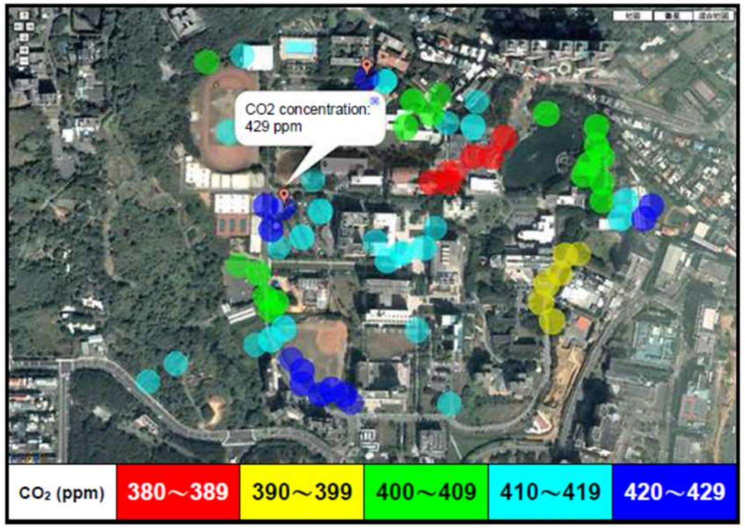
Hu's work (2/2)

National Sun Yat-sen University

- Hardware deployment:



- Monitoring result is shown on **Google Maps**:
Hsinchu city, Taiwan



HazeWatch (1/2)

- In HazeWatch, cars carry *metal oxide* sensors or *electrochemical* sensors.
  - Monitor CO, NO_2, and O_3 pollutants.

(a) Metal Oxide Sensor

(b) Electrochemical Sensor

- *Commercial monitors* offering more accurate detection of pollutants are installed on the roadside to **calibrate** the readings of sensors on cars.

(a) Commercial Monitor

(b) Calibration Chamber

V. Sivaraman, J. Carrapetta, K. Hu, and B.G. Luxan, "HazeWatch: a participatory sensor system for monitoring air pollution in Sydney," *Proc. IEEE Conf. Local Computer Networks*, 2013, pp. 56–64.

HazeWatch (2/2)

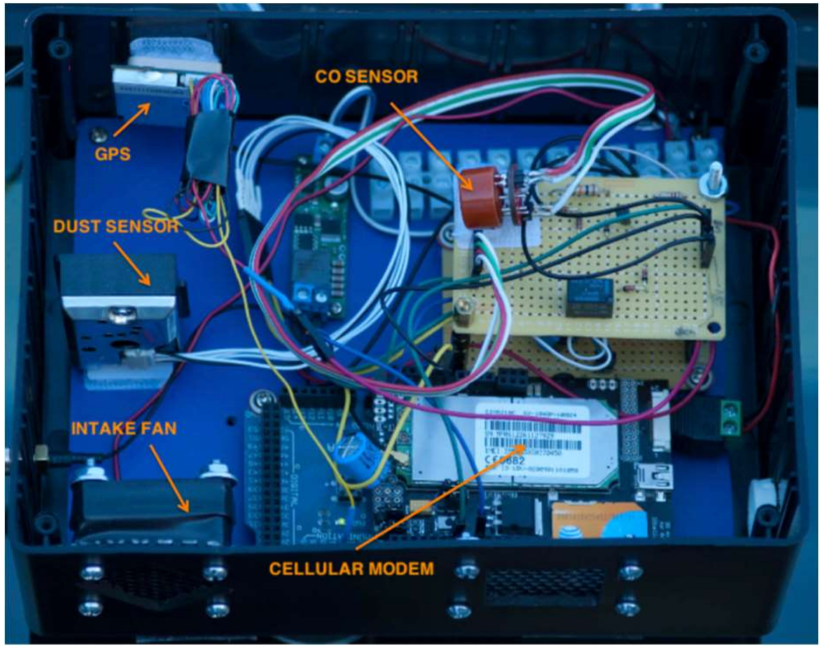
National Sun Yat-sen University

- Two **interpolation** methods can estimate values of air quality on the points without actual sensing data.
  - *Inverse-distance weighting method* estimates the pollutant's concentration on a point by assigning weights to all neighboring points, where a point farther away from the interpolation point has a smaller weight.
  - *Kriging method* is based on the calculation of the empirical semi-variogram over the data, where variogram describes the degree of spatial dependence of a spatial random field or stochastic process.
- HazeWatch was deployed in Sydney, Australia.

Devarakonda's work (1/3)

National Sun Yat-sen University

- Two mobile sensing models for air quality－
  Public transportation infrastructure:
  - Buses act as a mobile platform to collect air quality, where they move along fixed routes (usually **high volume** roads).
  - Each bus is installed with a *mobile sensing box*, containing
    - Arduino board
    - Communication module
    - GPS receiver
    - CO sensor
    - Dust sensor (to monitor PM)



S. Devarakonda et al., "Real-time air quality monitoring through mobile sensing in metropolitan areas," *Proc. ACM SIGKDD Int'l Workshop on Urban Computing*, 2013, pp. 1–8.

Devarakonda's work (2/3)

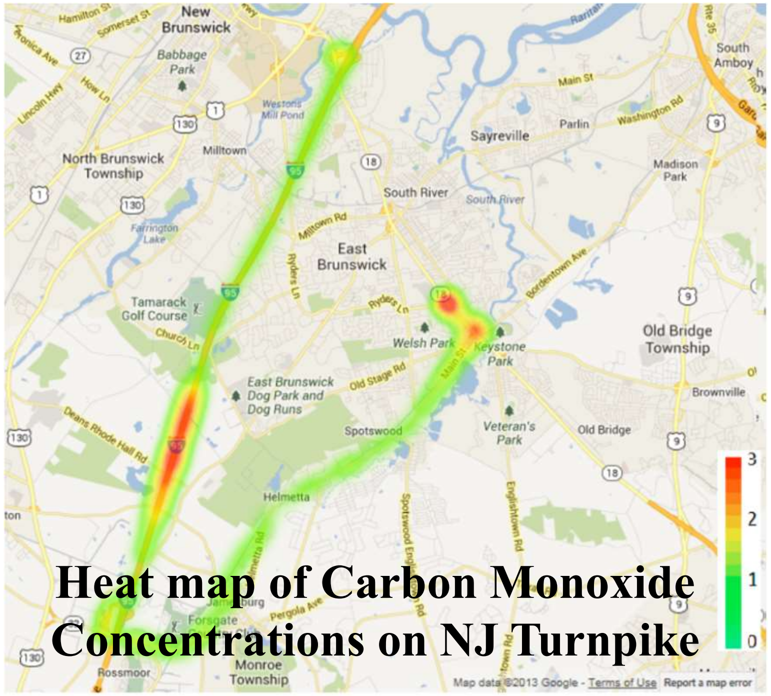
National
Sun Yat-sen
University

- Two mobile sensing models for air quality－ *Social community-based sensing*:
  - Drivers can install a *personal sensing device* on their cars and register to participate in collecting air quality.
  - Personal sensing device has a CO sensor and can talk with the driver's smart phone through a **Bluetooth** link.

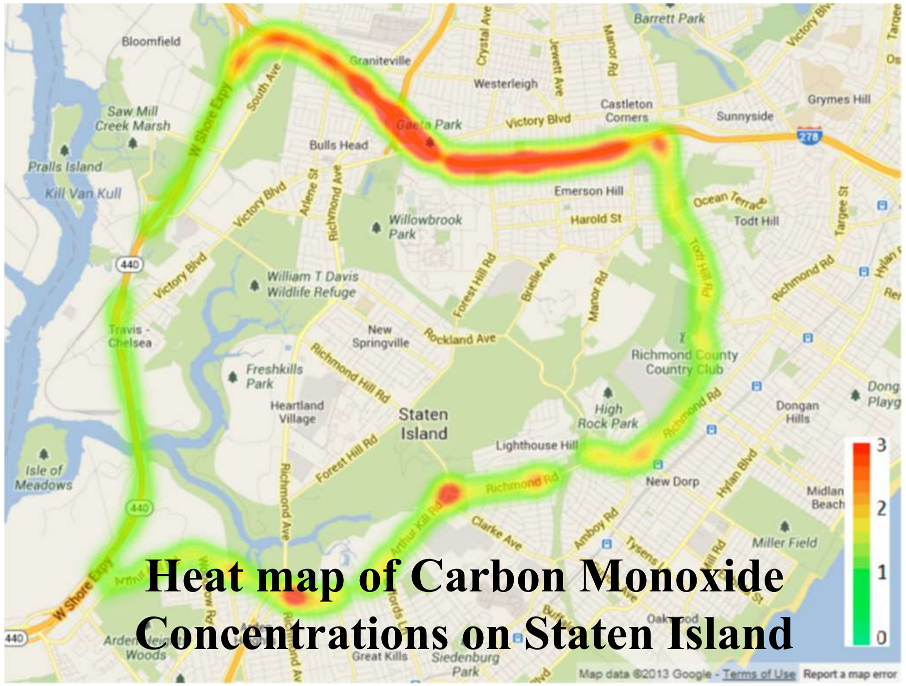
Devarakonda's work (3/3)

National
Sun Yat-sen
University

- Collected data are *geo-tagged* and sent to a server. Then, the server translates them to *AQI values* and display the degree of air pollution by a **heat map**.



Heat map of Carbon Monoxide Concentrations on NJ Turnpike



Heat map of Carbon Monoxide Concentrations on Staten Island

Outline

National
Sun Yat-sen
University

- Introduction
- Evaluation of air quality
- Collection of air quality
- Analysis of raw data
- Adjustment of reporting rates
- Research challenges
- Conclusion

O_3 analysis (1/3)

- Hasenfratz et al. connect a smart phone with an O_3 sensor through its **USB** interface.
  - To detect ground-level O_3 concentration, they measure the resistance of the sensor's **SnO_2** layer.
  - Smart phone polls the O_3 sensor to acquire its raw sensor readings every 100ms, which contains the SnO_2 layer's *resistance R* and the *on-board temperature T*.
  - *Temperature-compensated resistance* can be estimated by

$$R_t = R \times \exp[K(T - T_0)]$$

    - T_0 is the *reference temperature* and K is a coefficient used to adjust the *difference of temperature*.
    - In general, T_0 is set to 25°C and K is set to 0.025.

D. Hasenfratz, O. Saukh, S. Sturzenegger, and L. Thiele, "Participatory air pollution monitoring using smartphones," *Proc. Int'l Workshop on Mobile Sensing*, 2012, pp. 1–5.

O_3 analysis (2/3)

National Sun Yat-sen University

- As the *response curve* of the O_3 sensor is **quasi-linear** with respect to O_3's concentration, the concentration can be approximated by a *first-order polynomial*:

$$f(R_t, \alpha, \beta) = \alpha + \beta R_t$$

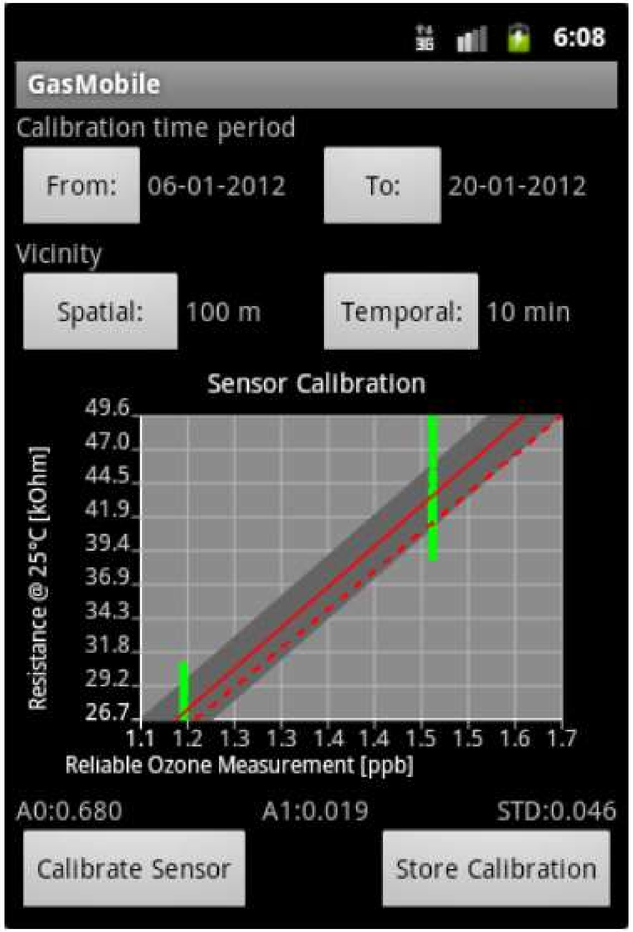
  - α and β are used to *calibrate sensor readings*.
  - Given a data set S with calibration tuples (R_t, M), where M is the reference measurement (e.g., data obtained from the nearby monitoring station), we can adopt *the least-squares method* to find both parameters α and β such that the *sum of squared differences* between $f(R_t, \alpha, \beta)$ and M is minimized:

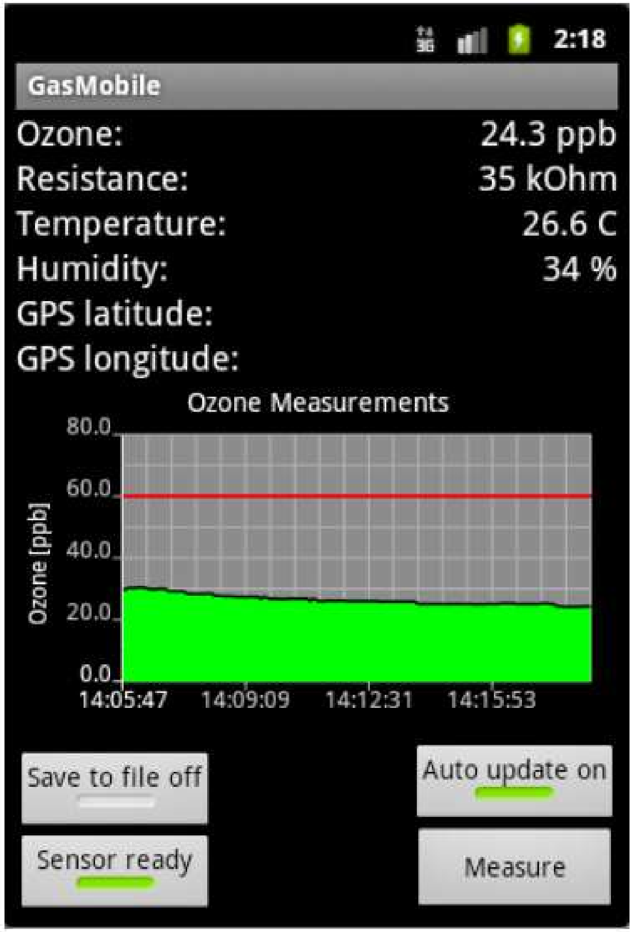
$$\arg \min_{\alpha,\beta} \sum_{\forall (R_t, M) \in S} [f(R_t, \alpha, \beta) - M]^2$$

O_3 analysis (3/3)

National
Sun Yat-sen
University

- Android-based APP:





PM2.5 analysis (1/3)

National Sun Yat-sen University

- Liu et al. use *in-built cameras* on phones to estimate the PM2.5 concentration, which is based on the relationship between the **haze** model and images.
  - Haze is an atmospheric phenomenon caused by dust, smoke, and PM2.5 to *obscure the clarification of sky*, making the image look *brownish* and *blurry*.
- Given a pixel *x* on the image, we can estimate its observed image **irradiance** by

$$O(x) = I(x) \times t(x) + L(1 - t(x))$$

  - $I(x)$ is the *scene irradiance* and L is *global atmospheric light*.
  - $t(x)$ is a meteorological parameter called **transmission**, which reveals the amount of light that can pass through atmosphere.

X. Liu, Z. Song, E. Ngai, J. Ma, and W. Wang, "PM2.5 monitoring using images from smartphones in participatory sensing," *Proc. IEEE INFOCOM Workshop*, 2015, pp. 630–635.

PM2.5 analysis (2/3)

National Sun Yat-sen University

- Three image features are used to improve the accuracy of concentration estimation:
  - *Spatial contrast:* Use the **decrease** of spatial contrast to observe the degradation of image caused by haze, where distant objects in an image with haze will lose its **acuity**.
  - *Dark channel:* The value of dark channel of a pixel x is the *minimum intensity* of the three color (red, green, blue) channels of the image block around x.
    - Dark channel of an image without haze should be **zero**.
  - *HSI (hue, saturation, intensity) color difference:* This difference of sky taken under different weather conditions will change with the *visibility* and *hazy situation*.

PM2.5 analysis (3/3)

- For prediction, one should take *a sequence of photos* $P_{ts} = \{P_{ts}^1, \dots, P_{ts}^m\}$ at a location L for m days, and get PM2.5 data $C_{ts} = \{C_{ts}^1, \dots, C_{ts}^m\}$ from *monitoring stations*.
  - Given a new photo P taken at location L, one can **compare** it with the photos in P_{ts}, and estimate PM2.5 concentration by consulting the data in C_{ts}.

Photos taken in Beijing

Outline

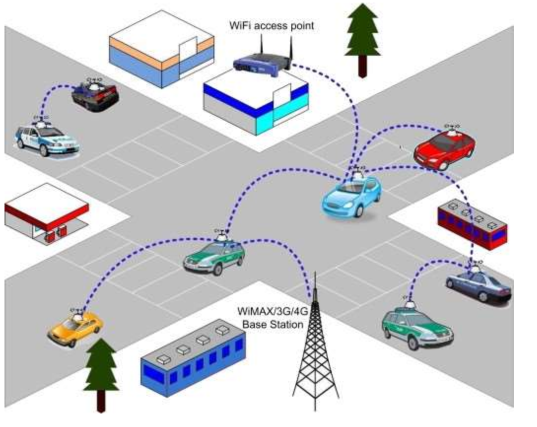
National
Sun Yat-sen
University

- Introduction
- Evaluation of air quality
- Collection of air quality
- Analysis of raw data
- Adjustment of reporting rates
- Research challenges
- Conclusion

Mobile agent (1/2)

National Sun Yat-sen University

- Mitra et al. install mobile agents (MAs) on cars in a **VANET** to perform the task of air-quality monitoring.
  - MAs are *migratory programs* capable of moving from one node to another node and being executed at a destination.



- Initially, server creates a few MAs, each with configuration information including
  - Target *monitoring region*
  - *Time period* to collect sensing data
  - *Type of sensing data* to be collected

G. Mitra, C. Chowdhury, and S. Neogy, "Application of mobile agent in VANET for measuring environmental data," *Proc. Int'l Conf. Applications and Innovations in Mobile Computing*, 2014, pp. 48–53.

Mobile agent (2/2)

National Sun Yat-sen University

- During movement of cars, MAs migrate from one car to another to reach the monitoring region in time.
  - If time **expires**, MAs send their collected data to the server.
  - Each MA checks if its lodging car is moving to the target region, or if the car is stuck in a **traffic jam** by two strategies:
    - *Distance strategy:* MA finds the distance between the lodging car and its destination. If the *distance does not decrease* as time goes by, the lodging car may be stuck in a traffic jam.
    - *Angle strategy:* MA measures the angle between the moving vector of the lodging car and the straight line to the destination. If the *angle increases* as time goes by, it means that the lodging car may move away from the destination
  - If so, MA migrates to another car accordingly.

Homogeneous grid partition (1/5)

National Sun Yat-sen University

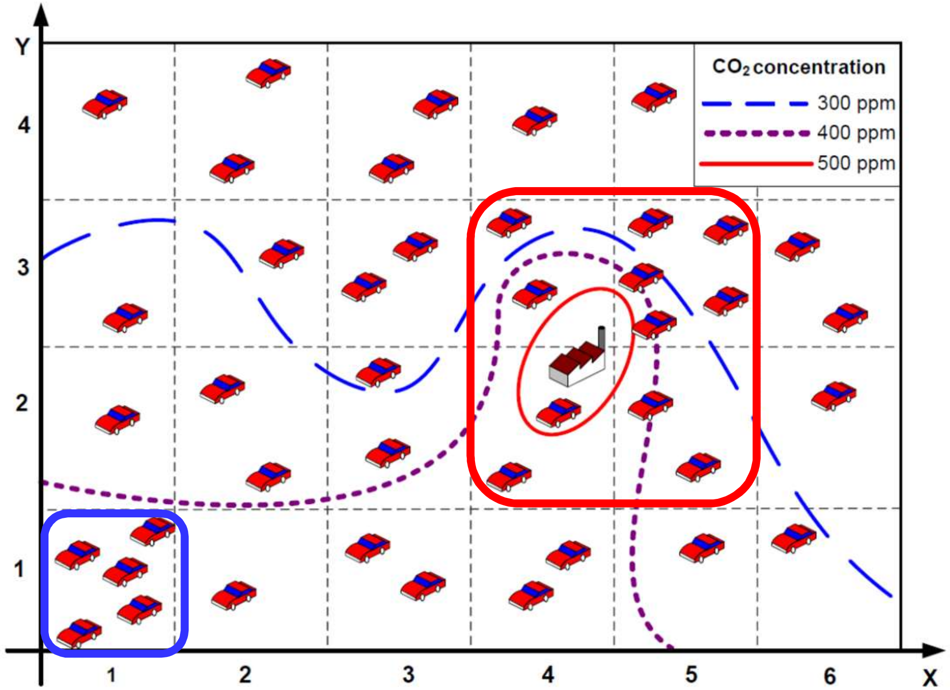
- Hu et al. divide the monitoring region into *homogeneous grids*, and adjust reporting rates of cars in each grid based on *car density* and *pollutant concentration*.
  - It incurs a **cost** for car drivers to transmit sensing data to a remote server (e.g., via short messages).
- The objective is to reduce the overall cost while ensuring the **accuracy** of monitoring result.
  - Obtain *sufficient air-quality samplings* to calculate the distribution of pollutant in the monitoring region.

S.C. Hu, Y.C. Wang, C.Y. Huang, and Y.C. Tseng, "Measuring air quality in city areas by vehicular wireless sensor networks," *J. Systems and Software*, vol. 84, no. 11, pp. 2005–2012, 2011.

Homogeneous grid partition (2/5)

National Sun Yat-sen University

- [Observation] Higher reporting rate is assigned to a grid where the **variation** of pollutant concentration rises.
  - Grids (4, 2), (5, 2), (4, 3), and (5, 3) have larger concentration variations, so *higher reporting rates* are imposed on them to improve the monitoring **accuracy**.
  - Grid (1, 1) has many cars but its pollutant concentration is flat, so it can be assigned a *lower reporting rate* to save the message **cost**.



Homogeneous grid partition (3/5)

National Sun Yat-sen University

- **Variation**-based scheme:
  - Let σ_i be the *standard deviation* of pollutant-concentration values collected from grid G_i in the previous period.
  - Number of air-quality samplings that should be received from G_i in the next period to keep monitoring accuracy is estimated by $S_i^{\text{var}} = \alpha_i^{\text{var}} \times \sigma_i + \beta_i^{\text{var}}$
    - α_i^{var} and β_i^{var} are two constants based on the past experience, and larger values imply *higher monitoring quality* but *larger message overhead*.
  - New reporting rate for G_i is S_i^{var}/n_i, where n_i is the number of cars in G_i that submitted their reports in the previous period.

Homogeneous grid partition (4/5)

National
Sun Yat-sen
University

- **Gradient**-based scheme:
  - Let $V_i^{\max}$ and $V_i^{\min}$ be the sets of the *highest* γ ratio and the *lowest* γ ratio of pollutant-concentration values collected from grid G_i in the previous period, respectively.
  - Gradient of two samplings $x \in V_i^{\max}$ and $y \in V_i^{\min}$ is defined by $\xi(x, y) = (x - y)/D(x, y)$, where $D(x, y)$ is their distance.
  - Average gradient in G_i is measured by

$$\xi_i^{\mathrm{avg}} = \frac{\sum_{x \in V_i^{\max},\ y \in V_i^{\min}} \xi(x, y)}{|V_i^{\max}| \times |V_i^{\min}|}$$

Homogeneous grid partition (5/5)

National Sun Yat-sen University

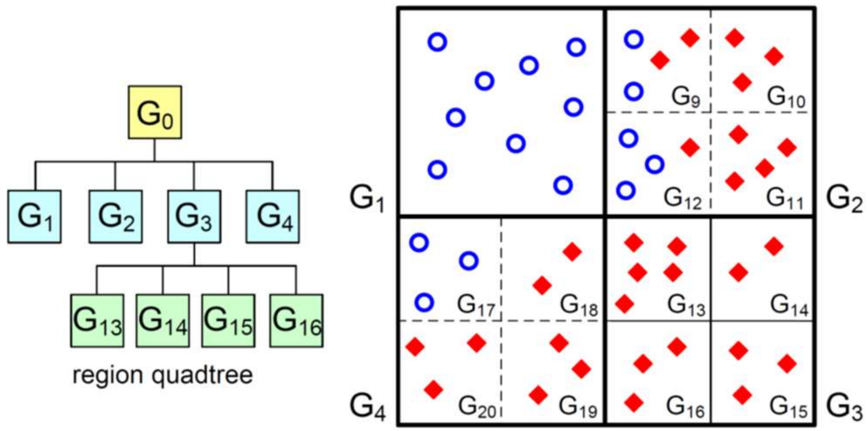
- **Gradient**-based scheme (continued):
  - Necessary number of air-quality samplings expected to be collected from grid G_i in the next period will be
    $$S_i^{\mathrm{gra}} = \alpha_i^{\mathrm{gra}} \times \xi_i^{\mathrm{avg}} + \beta_i^{\mathrm{gra}}$$
    - α_i^{gra} and β_i^{gra} are two constants based on the past experience.
  - New reporting rate for G_i will be S_i^{gra}/n_i.

- [Remark] Gradient-based scheme provides higher monitoring accuracy than variation-based scheme, since it takes the **positions** of air-quality samplings into consideration.

Heterogeneous grid partition (1/5)

National
Sun Yat-sen
University

- Wang et al. propose a heterogeneous grid partition, where the monitoring region is recursively **quartered** and indexed by a *region quadtree*.
  - It is a data structure used to describe a partition of 2D space by iteratively decomposing the space into *four equal quadrants*.
  - Each tree node in the region quadtree has either **four** children (i.e., an internal node) or **no** child (i.e., a leaf node).



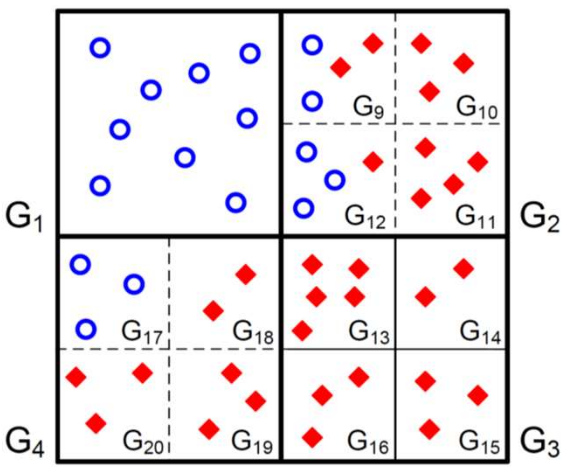
Y.C. Wang and G.W. Chen, "Efficient data gathering and estimation for metropolitan air quality monitoring by using vehicular sensor networks," *IEEE Trans. Vehicular Technology*, vol. 66, no. 8, pp. 7234–7248, 2017.

Heterogeneous grid partition (2/5)

National Sun Yat-sen University

- Let us call a set of air-quality samplings collected in grid G_i **λ-similar** if these sampling *belong to the same AQI class* and the *difference between the largest and the smallest samplings* does not exceed λ.

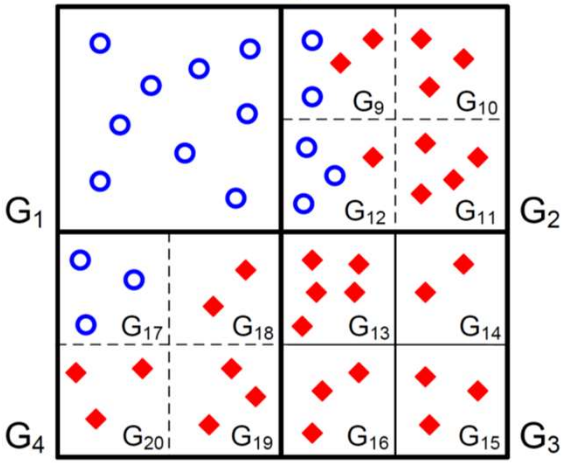
- Four operations are used to maintain grid partition:
  - *No-change operation:*
    If **all** samplings in G_i are $\boldsymbol{\lambda}$-similar (e.g., G_1), there is no need to change the grid, as the pollutant concentration *keeps steady* in the grid.



Heterogeneous grid partition (3/5)

National Sun Yat-sen University

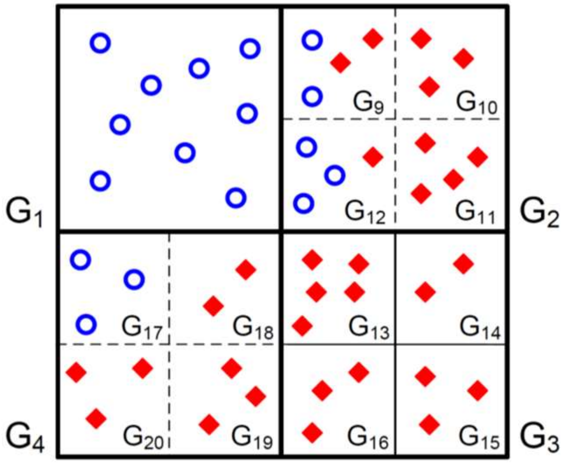
- Four operations are used to maintain grid partition:
  - *Dividing operation:*
    If some **child** grids of G_i have samplings not similar (e.g., G_2), it is better to divide grid G_i to get a more *fine-grained observation*.
  - *Merging operation:*
    If G_i and its 3 sibling grids have only $\boldsymbol{\lambda}$-similar samplings, we can merge them into the same one, as the current grid partition is too **narrow**.
    - [Example] Grids $G_{13}, G_{14}, G_{15}, G_{16}$ are merged to a large grid G_3.



Heterogeneous grid partition (4/5)

National Sun Yat-sen University

- Four operations are used to maintain grid partition:
  - *Marking operation:*
    It is used when a grid has non-$\boldsymbol{\lambda}$-similar samplings, but each of its child grids has **only** $\boldsymbol{\lambda}$-similar samplings.
    - [Example] Grid G_4 has two types of λ-similar samplings, but its child grids G_{17}, G_{18}, G_{19}, G_{20} each has only one type of samplings.
    - For this case, we prefer **not** to divide the grid, as each child grid of grid G_i can share the same reporting rate.



Heterogeneous grid partition (5/5)

National Sun Yat-sen University

- Reporting rate of each grid can be estimated by

$$R_i = \mu \times \frac{\omega(G_i)}{\phi(G_i) \times t_{\mathrm{avg}}(G_i)}$$

  - μ controls the **speed** to sample air quality.

$$\mu = \begin{cases} 0.5 & \text{if all air-quality samplings in grid } G_i \text{ are } \lambda\text{-similar} \\ 2 & \text{otherwise.} \end{cases}$$

  - $\omega(G_i)$ is a baseline for the *number of air-quality samplings* expected to be collected from grid G_i.
  - $\phi(G_i)$ is the *traffic density* in G_i.
  - $t_{avg}(G_i)$ is the average *time that cars stay* in G_i.

Outline

National
Sun Yat-sen
University

- Introduction
- Evaluation of air quality
- Collection of air quality
- Analysis of raw data
- Adjustment of reporting rates
- Research challenges
- Conclusion

Incentive mechanism

National Sun Yat-sen University

- To improve the *accuracy* or *resolution* of monitoring result, we need many sampling data of air quality.
  - One can design an **incentive** mechanism to encourage more people to voluntarily participate in the monitoring mission.

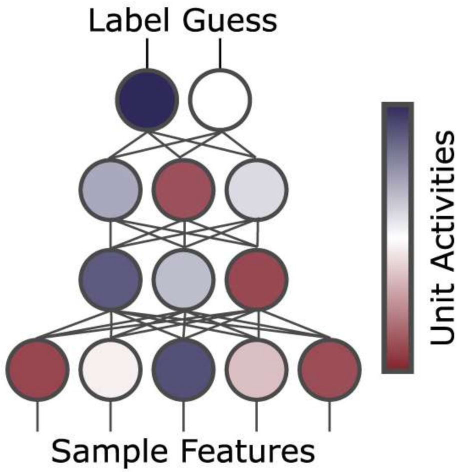
- [Challenge] How can we *adaptively adjust the reward* to guide people to collect air quality in certain regions?
  - For example, increase the sampling frequency when detecting abnormal air pollution.

Data prediction

National
Sun Yat-sen
University

- Collecting air quality on *every position* in the monitoring region is evidently infeasible.
  - It is a challenge to provide accurate **prediction** of air pollutant concentration for those positions with just *little information* or *even without* any sensing data.

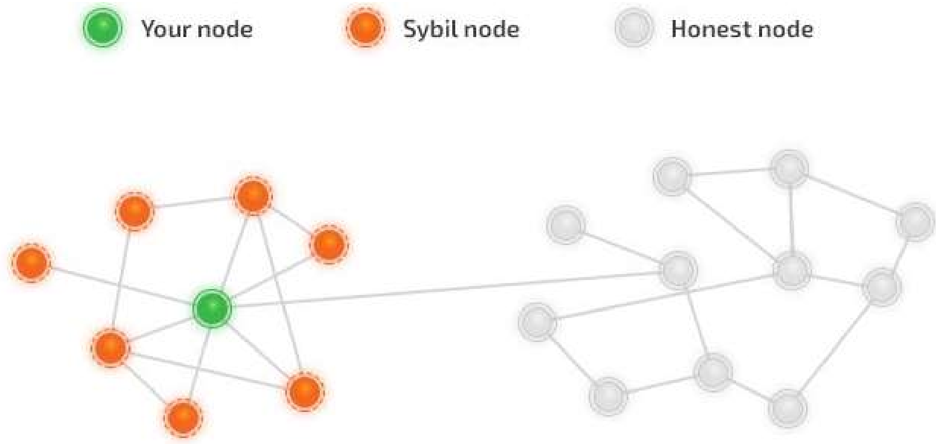
- [Possible Solutions to data prediction]
  - Air pollution dispersion models
  - Data mining
  - Deep reinforcement learning



Security

National
Sun Yat-sen
University

- Mobile IoT solutions rely on users to collect air quality and report their monitoring results.
  - Some sensing data submitted by users may not be correct due to *sensor failure* or even *malicious behavior*.
- [Challenge 1] How can we **distinguish** sensor failure from malicious behavior?
- [Challenge 2] **Sybil attack**



Outline

National
Sun Yat-sen
University

- Introduction
- Evaluation of air quality
- Collection of air quality
- Analysis of raw data
- Adjustment of reporting rates
- Research challenges
- Conclusion

Conclusion

National
Sun Yat-sen
University

- Many mobile IoT solutions are proposed to let citizens monitor air quality and provide the result to public.
  - *Pedestrians*, *cyclists*, and *drivers* can carry gas sensors to collect air quality, which have different mobility models.
- Two related issues:
  - How to *analyze raw data* collected by smart phones to estimate the concentration of pollutants?
  - How to adjust the reporting rate of sensing data by cars to balance between *monitoring resolution* and *cost*?
- Research challenges include *incentive mechanism*, *data prediction*, and *security*.