**The standard air quality level with the index of AQI (IAQI)**

|  |  |
| --- | --- |
| AQI | Air quality level |
| 0-50 | Good (G) |
| 51-100 | Moderate (M) |
| 101-150 | Unhealthy for sensitive groups (U-S) |
| 151-200 | Unhealthy (U) |
| 201-300 | Very unhealthy (VU) |
| 301-500 | Hazardous (H) |

**Definition 1**:

In predicting smog pollution in the following periods (e.g., the next 24 hours), we divide the pollution into types:

|  |  |
| --- | --- |
| Predicting Types | Definition |
| None-Pollution | AQI <= 100 at all hours |
| Slight-Pollution | AQI <=200 at all hours and AQI >-101 at some hour(s) |
| Severe-Pollution | AQI >=201 at some hour(s) |

We use a vector as the label of a period:



Example: (1, 0, 0) stands for Severe-Pollution

**Definition 2:**

Smog Evolution Graph (***G***): a *fully connected* and *directed* network where each node represent a city







, where vi represents aerosol optical thickness of ith node, n is the number of nodes, eij represents the diffusion factor from ith node to jth node.

**Definition 3:**

Smog pollution predictor:

, where t represents a time period, e.g., the next 24 hours, represents a spatial classifier for the time period t,  represents a temporal classifier for the time period t, f represents an ensemble framework.

**Problem:**

Given 1) a graph G, where node vi represents aerosol optical thickness of ith city and edge eij represents diffusion factor from ith city to jth city, 2) each city’s temporal features including air pollutant concentrations Fa, weather forecast Fw and meteorological elements Fm

We aim to predict each city’s smog pollution level Lt in the period t, e.g., the next 24 hours.

**Sub-Problem 1:**

Given:

1. Historical daily aerosol optical thickness records across China: Ds
2. Historical hourly concentrations of air pollutants in various ground stations in China: Da
3. Physical length between any two cities: Dl
4. Historical hourly records of wind speed and direction in various ground stations in China: Dw

We aim to build instances of graph G: estimate a city’s aerosol optical thickness vi and diffusion factor eij from ith city to jth city

Actually, we will learn a rule (*R1*) to estimate a city’s aerosol optical thickness with Ds and Da  (supervised training with ANN), and learn a rule (*R2*) to estimate the diffusion factor with Dl and Dw (unsupervised learning, e.g., PCA/Clustering, or manually defined formulas)

**Sub-Problem 2:**

Given:

1. Historical instances of smog evolution graph G
2. Historical smog pollution level of each city, e.g., Beijing

We aim to find the evolution pattern in the graph G: train the spatial classifier 

**Sub-Problem 3:**

Given:

1. Each city’s temporal features including air pollutant concentrations Fa, weather forecast Fw and meteorological elements Fm
2. Historical smog pollution level of each city, e.g., Beijing

We aim to train the temporal classifier 

IAAI: (Sep 10)

Emerging application case study

Title 1: Graph Processing for Smog Disaster Prediction with Multiple Sources

Title 2: Smog Disaster Prediction using Ground Sensor and Satellite Sensor Data

WWW: (Nov 3)

Title 1: Graph-based Multi-view Learning: A Case Study for Smog Disaster Prediction using Web-Scale Data