**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:**

For a city,

Given: 1) instances of graph G that contains the aerosol optical thickness and diffusion factors of the surrounding cities, 2) records of the city’s temporal factors including air pollutant concentrations, weather forecast and meteorological elements

Purpose: train the smog pollution predictor  with the ensemble of spatial classifiers  and a temporal classifiers 

**Sub-Problem 1:**

Given:

1. Daily aerosol optical thickness records in the past 12 years across China
2. Hourly concentrations of air pollutants in various ground stations in China

Purpose:

Learn a rule (*R1*) to estimate a city’s aerosol optical thickness with air quality observations of the ground sensors

Represent various input with a feature named aerosol optical thickness. It conforms to the insight that aerosol is the subject that actually diffuses from city to city.

(supervised training with ANN)

**Sub-Problem 2:**

Given:

1. Spatial length between any two cities
2. Hourly records of wind speed/direction of all the cities

Purpose:

Learn a rule (*R2*) to estimate the diffusion factor (discrete or continuous) from one city to another.

Find the pattern from different wind situations and use a discrete to represent the diffusion level.

(clustering or some manually defined formula)

**Sub-Problem 3:**

Given:

1. Daily instances of smog evolution graph G
2. Daily smog pollution label of a city, e.g., Beijing

Purpose:

Train the spatial classifier 

**Sub-Problem 4:**

Given:

Temporal feature sets:

1. Current ground sensor observations including meteorological elements and air quality concentrations of ith city
2. Forecasted meteorological elements of ith city

Purpose:

Train the temporal classifier 