City Issue Work Warning System

# **1. Introduction**

## **Motivation**

Infrastructure management and maintenance is always a major concern for community improvement and city governing. The aim of *city issue work warning system* is to apply data-driven strategies to deal with the city Infrastructure issues by forecasting the possible occurrences of different infrastructure issues. Based on the forecasting information, authorities are able to prepare for the possible city infrastructure issues to avoid possible personal and property damage, plan infostructure upgrade and finally improve the lives of its residents.

## **Solution**

The forecasting system is based on the correlation of environment factors (weather, traffic etc) and the city issue work. By modeling the environment data and city issue work data for last ten years, we build the forecasting system. Based on the model and reliable extrapolation of future environment data, we can make the prediction of future city Infrastructure issues.

# **2. Data**

* **The water and precipitation data:**

Community Collaborative Rain, Hail & Snow Network - Total Precipitation Summary <https://www.cocorahs.org/ViewData/TotalPrecipSummary.aspx>

**Dataset Overview:**

The water and precipitation data in csv format with around 120,000 observations. It is collected by the water monitoring stations in Champaign county and environment analysis department.

**Variables reference:**

*Observation Date* - Weather observation date; In this dataset, the observation date is from November 1st, 2006 to April 1st, 2019

*Observation Time* - The exact time to record the weather

*Entry Data Time* - The date to write down the record of weather data

*Station Number* - Number of each weather station in Champaign county

*Station Name* - Name of each weather station in Champaign county

*Latitude* - Latitude of each weather station in Champaign county

*Longitude* - Longitude of each weather station in Champaign county

*TotalPrecipAmt* - Total Precipitation Amount around each station; Precipitation is represented in inches per 24-hour period

*NewSnowDepth* - The depth of new snow that has fallen since the previous observation, usually 24 hours

*NewSnowSWE* - Snow Water Equivalent (SWE) is the amount of water contained within the snowpack in 24 hours

*TotalSnowDepth* - The total depth of snow, ice pellets, or ice on the ground

*TotalSnowSWE* - The total amount of water contained within the snowpack on the ground

*DateTimeStamp* - The date and exact time to record the weather

* **The city issue data**

Provided by the City government.

**Dataset Overview:**

The data is in csv format, with around 54,000 observations. It is published by the city council regarding the city work from 2009 with its type, location and working status.

**Variables reference:**

*# of SR* - Each cityworks’ number

*DESCRIPTION* - Type of cityworks(Clean inlet, Pick up abandoned bicycle(s), etc.)

*DATETIMEINIT* - The start date  to work on each cityworks

*DATETIMECLOSED* - The finished date on each cityworks

*STATUS* - The status about each cityworks, it is either ‘Closed’ or ‘Open’

*PROBADDRESS*- The address of each construction site

*PROBLOCATION* - The original address of each construction site ; It would be NA, if the construction address is not changed

*SRX* - Coordinate of the issue(x-axis) *SRY* - Coordinate of the issue(y-axis)

* [**Vehicle Turning Movement Counts**](https://data.ccrpc.org/dataset/traffic_counts/resource/3da71f28-2f11-49f9-bb6e-451ecdce1994) **data**

<https://data.ccrpc.org/dataset/traffic_counts>

Raw Data Access is in progress

* [**Pedestrian Crossing Counts**](https://data.ccrpc.org/dataset/traffic_counts/resource/eb541724-572b-406a-ac19-ecd283bfff97) **data**

<https://data.ccrpc.org/dataset/traffic_counts>

Raw Data Access is in progress

* [**Historical Average Daily Traffic**](https://data.ccrpc.org/dataset/traffic_counts/resource/a632505b-bd5e-4aad-bae7-a84f66b3f7ed) **data**

<https://data.ccrpc.org/dataset/traffic_counts>

Raw Data Access is in progress

* **Daily and Monthly Temperature data**

<https://stateclimatologist.web.illinois.edu/data/champaign-urbana/>

Raw Data Access is in progress

# **3. Methods**

## **Overview**

Based on R and shiny application, to visualize the insights of infrastructure issues and procced the extrapolation analysis to predict the future infrastructure issues.

## **Prototypes**

A close up of text on a white background

Description automatically generatedThe sketched interface of our Shiny App is shown below:

* **Factor Analysis**

Visualization 1: Precipitation/Snow Analysis

Line plot to reflect the annual precipitation/snow trend with bar button to select year from 2009 to 2019; Heat map to reflect the geographic distribution of annual precipitation/snow with bar button to select year from 2009 to 2019

Visualization 2: Traffic Analysis

Line plot to reflect the traffic trend with bar button to select year from 2009 to 2019; Heat map to reflect the geographic distribution of annual traffic variation with bar button to select year from 2009 to 2019

Visualization 3: Temperature Analysis

Line plot to reflect the temperature trend with bar button to select year from 2009 to 2019; Bar plot to identify the months with extreme temperature each year with a bar button to select year from 2009 to 2019.

Visualization 4: People Analysis

Line plot to reflect the annual human traffic trend with bar button to select year from 2009 to 2019; Heat map to reflect the geographic distribution of annual human traffic with bar button to select year from 2009 to 2019

* **Infrastructure issue Analysis**

Visualization 1: Type Analysis

Bar plot to identify the popular types of city issue work (different labels) of each year with bar button to select year from 2009 to 2019

Visualization 2: Location Analysis

Heat map to display the distribution of total number of city issue in Champaign county (visualize number of city issue work in different district); Heat map to display the distribution of different types of city issue work with bar button to select different city issue work type.

Visualization 3: Trend Analysis

Line plot to reflect the annul trend of total number of city issue work variation with bar button to select year from 2009 to 2019; Line plot to reflect the annul trend of total number of different type of city issue work variation (lines in different colors and label) with bar button to select year from 2009 to 2019.

* **Correlation Analysis**

Visualization 1: Correlation on Overall Trend

Line plot including line of total number of city issue variation and lines of different factors (Precipitation/Snow, Traffic, Temperature and People) variation to detect the possible correlation.

Visualization 2: Correlation on geographic distribution

Compare different heat maps including

1. Distribution of total number of city issue in Champaign county in a year with button to select year from 2009 to 2019

2. Geographic distribution of annual precipitation/snow variation with bar button to select year from 2009 to 2019

3. Geographic distribution of annual traffic variation with bar button to select year from 2009 to 2019

4. Geographic distribution of annual human traffic variation with bar button to select year from 2009 to 2019

Visualization 3: Model evaluation

1. Scatter plots and correlation coefficients calculated to test relevance

2. Modeling the past city issue work records and different factors with different models, including random forest, ridge regression, lasso and elastic net. Compare different model with relative criteria including 1: MSE 2.MSR 3. Test and Training data etc. List the comparison result and choose the final best model.

* **Prediction Model**

**Part 1:**

Input: Date

Output:

1. Total number of city issue work might happen on that day and number of city issue work in different districts, both in table

2. Number of major types of city issue work might happen on that day with location information in table

**Part 2:**

Input: Location

Output:

1. Future trend of overall city issue work in this district in line plot

2. Future trend of different types of city issue work in this district in line plot with different line corresponded to different types