**Project Plan: Predicting Parking Availability Using Spatio-Temporal Data**

Qiongyan Wang / Weilin Ruan

**Background**

Urban traffic congestion is a growing concern in densely populated areas, significantly contributing to air pollution and social issues. With an increasing number of vehicles, the demand for efficient parking space management has become critical. In Singapore, the challenge of predicting real-time parking availability (PA) is compounded by complex spatial and temporal factors, including weather conditions, land use, and road density. Existing models often fail to capture the intricate relationships among these variables, leading to suboptimal predictions and exacerbating traffic congestion.

**Objective**

The primary objective of this project is to develop a robust spatio-temporal prediction model for parking availability in Singapore. This model aims to integrate a diverse range of external factors, enhancing the accuracy of PA forecasts. By leveraging advanced deep learning techniques, we seek to provide real-time predictions that can aid drivers in locating available parking spaces and assist urban planners in making informed decisions regarding parking management.

**Methodology**

To achieve this objective, we will adopt the following methodology:

**Data Collection and Preprocessing:** We will utilize the SINPA dataset, which comprises a year’s worth of parking availability data from 1,687 parking lots in Singapore, enriched with various spatial and temporal factors. Data preprocessing will include cleaning, normalization, and the integration of external features such as meteorological data and land use information.

**Model Development:** We will implement a deep learning based framework. This will involve:

1. Hierarchical Network: Using hierarchical information to assist in improving the prediction accuracy of each parking lots.
2. Heterophilic GCN: We will use a heterophilic GCN to cosider each parking lot’s
3. Multi-Head Self Attention (MSA): We will introduce a multi-head self attention network to caputre the temporal relationship between each time steps .

**Model Training and Evaluation:** The model will be trained using a split dataset (training, validation, and testing) to ensure generalizability. We will employ performance metrics such as Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) to evaluate the model’s accuracy against baseline models.