**Liteature Survey**

**Group 5 LI LINGYU LIU YICHAO WU JINGYAN YANG QINGSHAN YE FANGDA**

**Summary of Existing Systems and Their Effectiveness**  
Currently, three predictive policing systems stand out for their widespread application and research value: PredPol and HunchLab in US, and Crime Anticipation System (CAS) in Netherlands.

PredPol, initially introduced in 2012, originated as a research project at the University of California. It has since evolved into the most widely implemented predictive policing algorithm in the United States. The system leverages three simple primary types of historical data: crime type, crime location, and crime time. By analysing these data, PredPol generates crime risk maps that indicates areas with high crime probability. The algorithm has proven effective in optimizing police resource allocation and reducing crime rates in certain areas.

HunchLab is another popular predictive policing system in US but more complicated. It aims to predict how likely a particular crime type is to occur at various locations across a time period. HunchLab is based on machine learning and requires non-crime data. Same as PredPol, HunchLab will first collect crime related event data including location and time. The system can also process complicated geographic data contains points, lines and polygons, temporal data sets like weather data and school schedule, census data like house vacancy status, and natural terrain data.

Crime Anticipation System (CAS) is a similar system used in Netherlands, and it relies on historical crime data to generate a heatmap to identify short-term crime risks. CAS operates on the principle that crimes are not randomly distributed but are concentrated in certain geographic and temporal clusters. The system integrates geographic clustering methods with temporal crime trends, recognizing that recent crimes often predict future crimes in nearby areas within a short time window. Same as PredPol, CAS will also collect basic crime data including type, location and time. Besides, CAS also process geospatial data like neighborhood layouts and temporal data like holidays and festivals.

**Review of Modeling Approaches**  
PredPol uses a straightforward statistical approach assuming that crime event follows a statistical distribution with two parameters: space and time. It believes past crimes are indicative of future crimes trends in nearby areas.

**Feature Engineering Techniques**

**Evaluation Metrics and Their Appropriateness**