POLICE ALLOCATION IN CAMBRIDGE

BY ANDERS STEINESS, AZIZ MALOUCHE, AND AZFAL PEERMOHAMMED

MOTIVATION

Current Day Station

Topics Under Question

- **Challenge:** Urban crime patterns and staffing constraints require optimized police resource allocation to maintain safety and trust.
- **Approach**: Analyze Cambridge crime data to determine optimal station locations and officer staffing.
- **Impact**: Improve response times, solve crimes efficiently, and strengthen community trust in law enforcement.
- **Scalability**: Framework applicable to other cities, promoting equitable and effective resource use nationwide.

MIT MANAGEMENT SLOAN SCHOOL

HEATMAP OF CRIME IN CAMBRIDGE

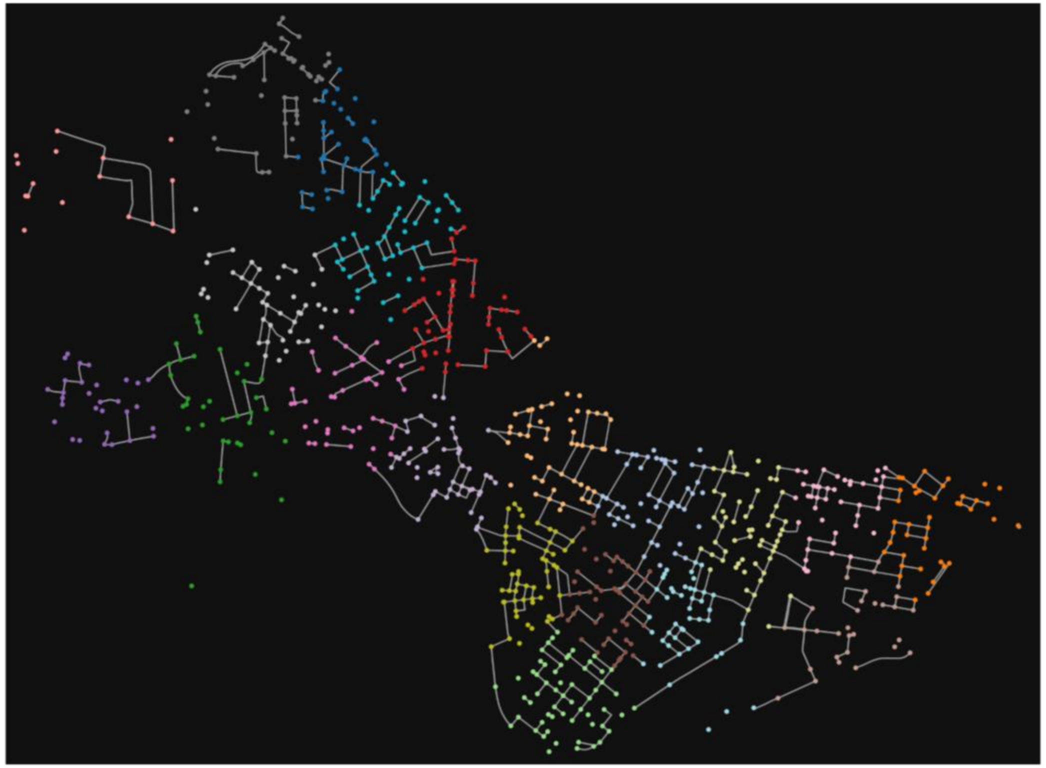
Mass Avenue leads the way in crime

Data Generation

- Road Structure sourced using osmnx
- Distance Matrix computed
- Crime Report incl. crime type and address originally from the City of Cambridge Police Department
- Crime Longitude and Latitude gathered using Google Maps API
- Mapped crime to nearest node
- Crime Data Used from 2020 to 2024
- Out-of-sample data: 2024

DOWNSAMPLED GRAPH

900 Nodes Subgraph for Computational Purposes

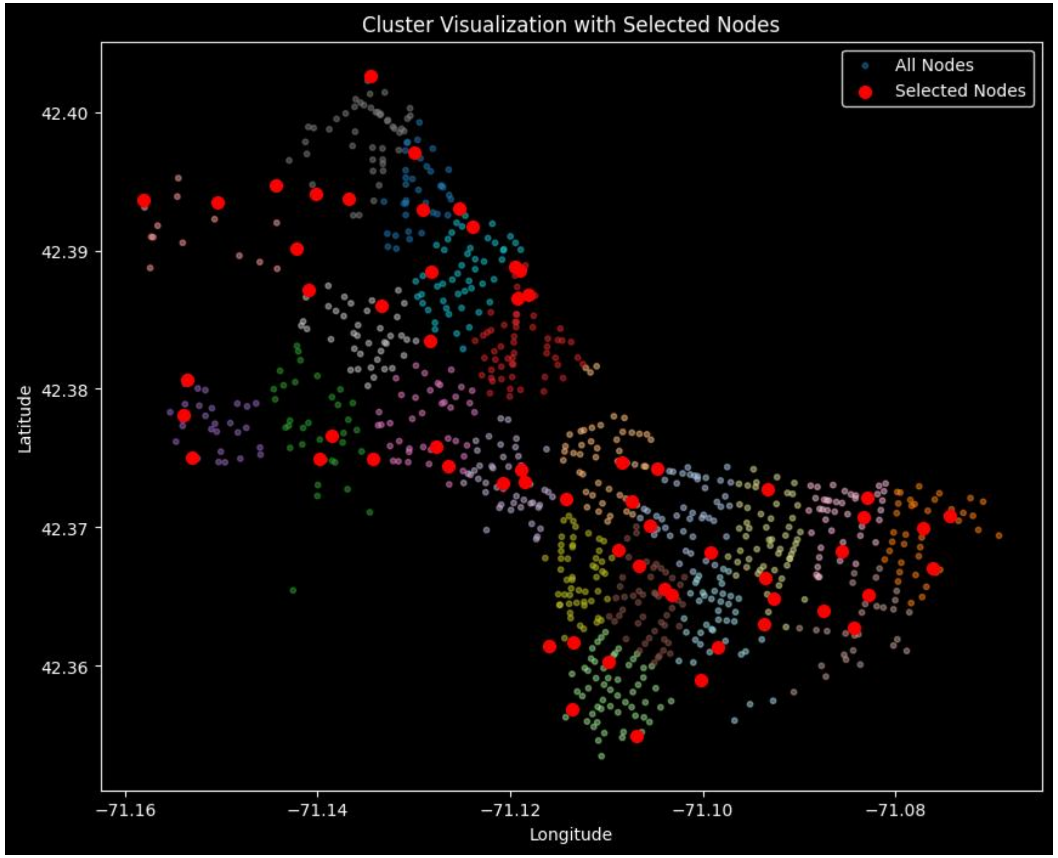


Downsample Procedure

- K-means clustering with 20 clusters (100 random initialization) based on longitude and latitude data
- Randomly sampled 45 nodes from each cluster
- Remapped Crime to closest node in new subgraph
- Computationally more feasible problem

POLICE STATION CANDIDATE NODES

60 Selected Candidate Nodes



Candidate Nodes for Police Station Procedure

- K-means clustering with 20 clusters (100 random initialization) based on longitude and latitude data
- Selected the two nodes with highest crime in each cluster
- Computationally practical problem

MIT MANAGEMENT SLOAN SCHOOL

MODEL FORMULATION

Decision Variables and Parameters

First Stage Decision Variable

$x_j \in \{0,1\}$: Whether a police station is built at candidate node $j \in J$

$y_j \in \mathcal{Z}_+$: The number of officers assigned to station $j \in J$.

$s_j \in \{0,1\}$: A binary variable representing whether a station has an extended jurisdiction.

$a_{ji} \in \{0,1\}$: A binary variable to enforce the coverage area for the police stations.

Second Stage Decision Variables

$z_{jics} \in \mathcal{Z}_+$: The number of officers assigned from station j to node i for crime type c under scenario s.

$u_{ics} \in \mathcal{Z}_+$: The number of late responses to a crime at node i for crime type c under scenario s at a given a given time period.

$z'_{jics} \in \mathcal{Z}_+$: The number of officers dispatched from police station j to node i for crime type c under scenario s, where the response time was not quick enough (i.e., outside of station j's coverage area).

Parameters

d_{ij}: The distance from node i to node j in meters.

N_{ics}: The total number of crimes at node i of crime type c for scenario s.

r_c: The total number police required to be dispatched to be equipped for crime type c.

π_c: The penalty for not having the capacity to attend a crime of type c within an acceptable time frame.

t: The transportation costs for The Cambridge Police Department when dispatching police officers to crimes throughout Cambridge.

Optimization Methods

- Adaptive Stochastic Optimization
- Deterministic Optimization
- Baseline Non-Optimization Model

First Stage Decision Variables

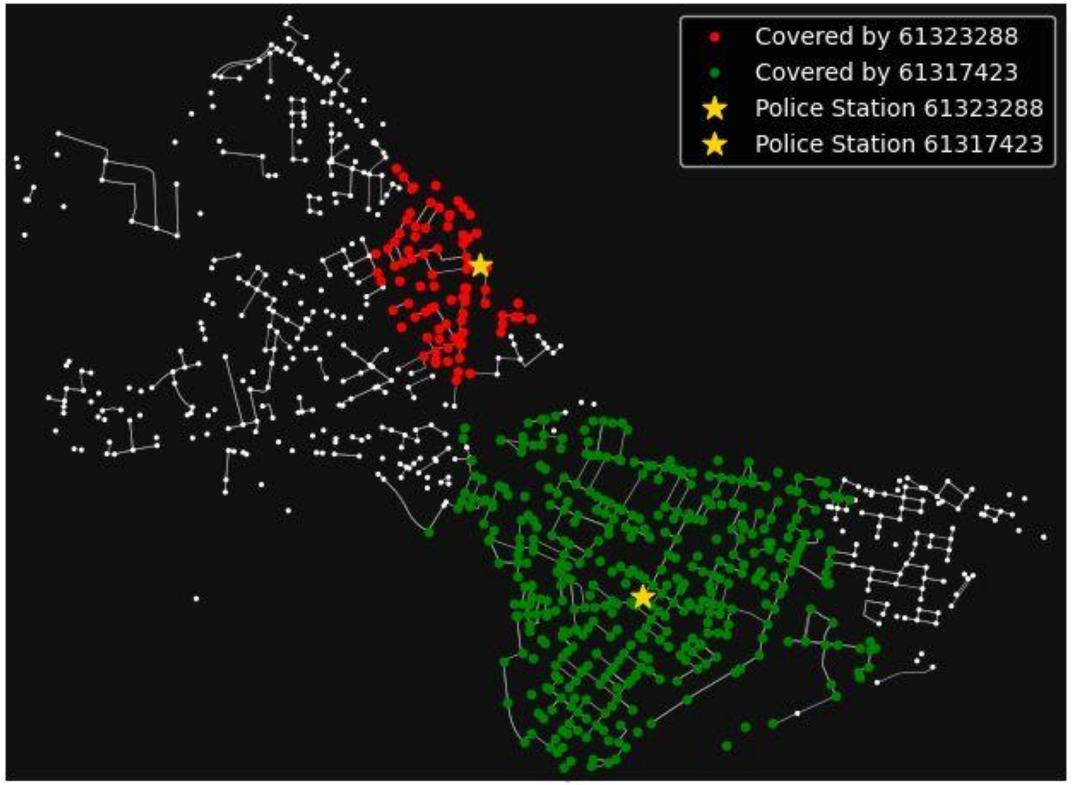
- Police Station Facility Location
- Police Officers Assignment to Stations

Second Stage Decision Variables

- Police Officer Dispatchment from Station j to Crime at Node i

MODEL CONSIDERATIONS

Large vs. Small Coverage Area



First Stage Decision Variables

- Police Station Facility Location
- Police Officers Assignment to Stations

Second Stage Decision Variables

- Police Officer Dispatchment from Station j to Crime at Node i

Model Consideration

- Police Station Can Have Large Coverage Area (2000 meters) or Small Coverage Area (1000 meters)
- If Crime Outside Coverage Area, then the Crime will not be met on Time and Cost is Incurred
- If Overlapping Coverage Area, Police Will be Dispatched from Nearest Station

MIT MANAGEMENT SLOAN SCHOOL

REALISTIC SCENARIO GENERATION

2022 Heatmap of Crime in Cambridge

2023 Heatmap of Crime in Cambridge

COVID SCENARIO GENERATION WITH UNUSUAL CRIME DYNAMICS – TO TEST LIMITS OF ADAPTIVE STOCHASTIC MODEL'S PERFORMANCE

2020 Heatmap of Crime in Cambridge

2021 Heatmap of Crime in Cambridge

2024 RESULTS (2 STATIONS)

2024	Cost	Late Crimes	Late Officers
Random Baseline	$113M	2,145	10,378
Highest Crime Baseline	$63M	1,064	5,284
K-Means Baseline	$64M	1,056	5,320
Present Day Baseline	$60M	1,028	5,100
Deterministic Optimization	$49M	843	4,124
Adaptive Optimization	$49M	843	4,124

2024 RESULTS (3 STATIONS)

2024	Cost	Late Crimes	Late Officers
Random Baseline	$70M	1,242	6,114
Highest Crime Baseline	$51M	827	4,152
K-Means Baseline	$37M	635	3,138
Present Day Baseline	$48M	836	4,094
Deterministic Optimization	$29M	513	2,482
Adaptive Optimization	$15M	234	1,152

MIT MANAGEMENT SLOAN SCHOOL

RESULTS (3 STATIONS W/ COVID SCENARIOS)

2024	Cost	Late Crimes	Late Officers
Random Baseline	$70M	1,238	6,102
Highest Crime Baseline	$50M	828	4,130
K-Means Baseline	$37M	630	3,114
Present Day Baseline	$47M	831	4,062
Deterministic Optimization	$27M	464	2,280
Adaptive Optimization	$31M	526	2,640

MIT MANAGEMENT SLOAN SCHOOL

2024 MODELS FOR 3 STATIONS

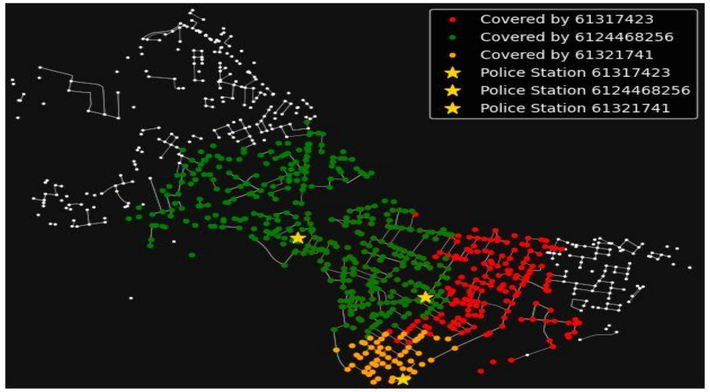


Figure 1: Current Baseline

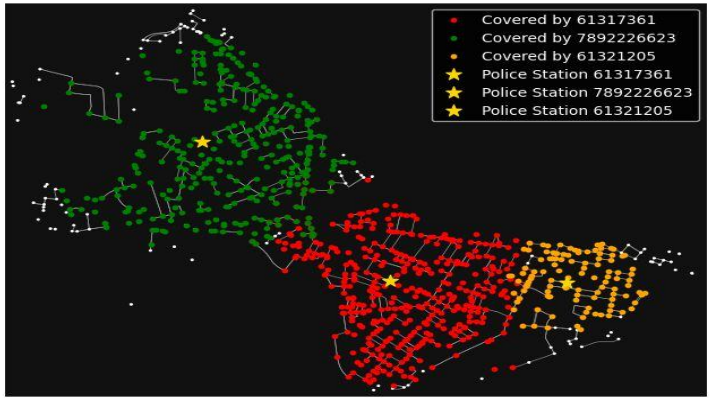


Figure 2: K-Means Baseline

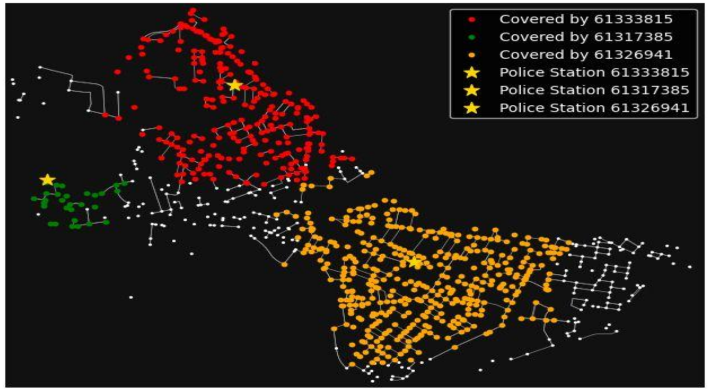


Figure 3: Deterministic

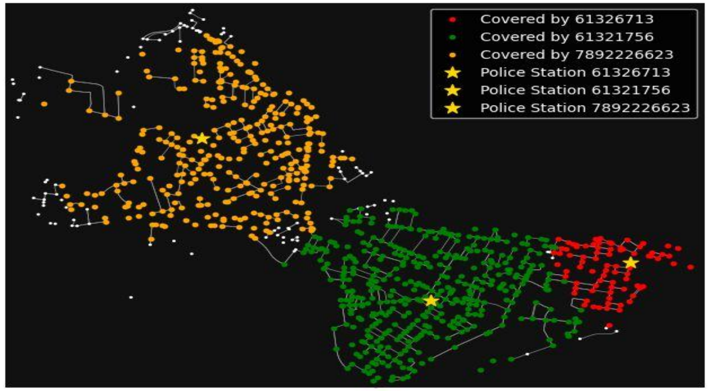


Figure 4: Adaptive

- **Current Baseline:** Lowest on time node coverage and worst performance.
- **K-Means:** Has very high on time node coverage
- **Deterministic:** Focuses on high-crimes, outperforming K-Means and demonstrating mere coverage isn't necessarily optimal
- **Adaptive:** Achieves the best overall performance with high coverage

MIT MANAGEMENT SLOAN SCHOOL

KEY FINDINGS AND TAKEAWAYS

Strength of Adaptive Model and Optimization

Consistent Performance:
Deterministic and adaptive models outperform all four baselines
1. Random
2. Top 3 crimes
3. K-means
4. Existing station

Adaptive Reduces Late Crimes:
•By ~50% compared to deterministic
•2.5x more compared to K-Means
•3.5x more compared to Top Crime Node Baseline

Alignment Between In-Sample and Out-of-Sample Results

2022-23 Results
Adaptive optimization reduced costs to of $33,945,299 and late crimes to 1,169, the lowest among all models

2024 Results
Adaptive optimization reduced costs to $14,508,239 and late crimes to 234, again the lowest among all models.

Limitations of Using COVID Scenarios for Crime Patterns

Impact of Including 2020–2021
Adding scenarios from 2020 and 2021, alongside 2022 and 2023, worsened adaptive performance.

Reason
COVID years were not representative of current or future crime patterns.

Key Insight
Uncalibrated scenarios can hinder the effectiveness of the adaptive stochastic method's solution.