



# Optimization Methods

## Using Optimization to Mitigate Crime in Cambridge

by

Anders Steiness: [steiness@mit.edu](mailto:steiness@mit.edu)  
Aziz Malouche: [azizm571@mit.edu](mailto:azizm571@mit.edu)  
Azfal Peermohammed : [azpeer@mit.edu](mailto:azpeer@mit.edu)

Date of submission  
December 12, 2024

---

# 1 Introduction

Public safety is a foundational pillar of any thriving community. Effective police resource allocation is a critical challenge for public safety agencies, especially in urban environments where crime dynamics vary spatially and temporally. Police departments often face constraints in staffing, limiting their ability to respond promptly to incidents, and maintain public trust. Optimally restructuring where officers are stationed and how they are assigned could help mitigate the challenges caused by reduced personnel, ensuring that the limited resources available are used where they can make the greatest impact.

Using crime data from Cambridge, Massachusetts, this analysis seeks to determine the optimal locations for police stations and the allocation of officers to address crime patterns efficiently. By aligning resource placement with areas of greatest need, this approach aims to improve outcomes despite staffing constraints.

Although focused on Cambridge, the methodology offers insights that are relevant to cities nationwide. Any municipality with access to similar data can apply this approach to reimagine its police station and officer allocation strategies, enhancing its ability to respond to and solve crimes. In a time of increased scrutiny and constrained resources, optimizing police resources can be a vital step toward building safer, more resilient communities. By improving the efficiency and equity of resource distribution, this data-driven approach has the potential to maintain public trust in law enforcement and strengthen the relationship between police departments and the communities they serve.

## 2 Data Processing & Preparation For The Model

### 2.1 Constructing Structure of Cambridge

The road network data for the city of Cambridge was extracted using the `osmnx` and `networkx` libraries. Each intersection within the city was modeled as a node in a graph, with edges representing the roads connecting these intersections. The geographic coordinates (longitude and latitude) for each node were obtained through these tools, enabling a spatial representation of the road network. Refer to Figure 1 to see the structure of the map. To support the optimization model, we computed the shortest path distances between every pair of nodes in the network. This step resulted in the creation of a distance matrix, which serves as a key parameter in the final optimization model. It was assumed that police vehicles would traverse the road network using the shortest paths between nodes.

### 2.2 Extracting the Crime Data

The crime data used in this study was sourced from Kaggle that uses official data from the City of Cambridge <sup>1</sup>. The dataset detailing various attributes of reported crimes including information on the type of crime, the time of occurrence, and the associated address. To enhance the usability of this data, the geographic coordinates (longitude and latitude) of each crime's reported address were extracted using the Google Maps API. The dataset covers all reported crimes in Cambridge from January 2019 to May 2024, totaling approximately 30,000 crime records. The crime is then mapped to the closest node in the Cambridge road structure network. Figure 2 shows a heatmap of the constructed crime distribution over the road network.

---

<sup>1</sup><https://www.kaggle.com/datasets/melissamonfared/cambridge-crime-data-2009-2024>



**Figure 1:** Cambridge Road Structure



**Figure 2:** Cambridge Crime Heatmap

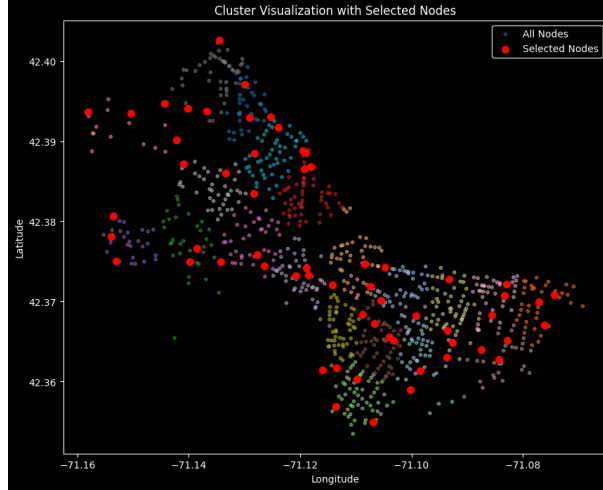
As seen in Figure 2, there are clear patterns in where crime is committed. We observe that Massachusetts Avenue has a high prevalence of crime especially around Central Square and Harvard Square. One potential explanation is that these areas are also where the most people are.

### 2.3 Downsampling

It became evident that keeping the whole road structure, consisting of 1869 nodes, made the optimization model computationally very expensive. Further, a smaller representation of Cambridge would likely provide similar optimization problem since including every road intersection of Cambridge might not be necessary. However, preserving the structure of the map is key to achieving a smaller approximate representation of Cambridge. Firstly, we omitted nodes with zero crimes occurred since those are irrelevant. Secondly, to preserve the structure of graph, we conducted K-means clustering to partition the graph into 20 clusters based on longitude and latitude (with 100 random initializations to ensure robustness). We sampled 45 nodes from each cluster such that we achieved a subgraph of 900 nodes (see Appendix 6 for different subgraph sizes). This approach ensured a balanced and representative downsampling of the road network while preserving its spatial distribution. Appendix 1 shows downsampled version with 300, 500, 700, and 900 nodes where the colors denote different clusters. The rest of the project continues with a 900 node representation of Cambridge as it preserves the structure of Cambridge, keeps a lot of nodes, and is computationally durable. It should be noted that all crimes were remapped to their closest node in the downsampled version.

### 2.4 Candidate Selection

Even though the downsampled road network consists of hundreds of nodes, not all nodes will be candidates for the Cambridge Police Department. Firstly, not all nodes are available to be a police station for various reasons. Secondly, it was also too computationally demanding. Thus, a candidate set of potential nodes to host a police station needs to be determined. The requirements for the candidate set for this project is that it should be computationally feasible, the candidate set should be spread out, and there should be enough candidate nodes to make an interesting optimization problem. To achieve this, the downsampled road network was clustered again into 20 clusters using kmeans cluster, and then the 3 nodes with the highest frequency of crimes in each cluster was included as candidate nodes. This results in 60 candidate nodes that are computationally durable, making the optimization problem interesting, and spread throughout Cambridge as seen in the Figure below.



**Figure 3:** Candidate Nodes

## 2.5 Extracting Crime Types, Distance Matrix, and Other Parameters

Our dataset consists of approximately 53 crime types, which were grouped into high, medium, and low severity categories based on fines and potential jail time outlined by Cambridge. High-severity crimes, such as homicides, required 8 officers for an immediate response. Medium-severity crimes needed 4 officers, and low-severity crimes, like noise complaints, required 2 officers. Delays in responding to high-severity crimes incur a cost of \$150,000, medium-severity crimes \$10,000, and low-severity crimes \$500. This prioritization ensures our model emphasizes timely responses to high-severity crimes first, followed by medium and low-severity ones. We also factored in a gasoline cost of \$0.10 per meter traveled. The total number of officers available corresponds to the current staffing level for Cambridge Police Department <sup>2</sup>.

## 3 Methodology

This section will cover how the problem was tackled where it first will cover the formulation for a deterministic and adaptive stochastic optimization problem. Scenario generation will then briefly be outlined and what time periods that will be covered. At last, the baseline models that our optimization models will be compared against will be established.

### 3.1 Optimization Models

#### 3.1.1 Optimization Models Formulation

This project both employs an adaptive stochastic optimization model and a deterministic model. The motivation for a two-stage stochastic optimization model for police station location is to account for the inherent variability in crime dynamics over time to provide robust/generalizable solutions, which will be evaluated on an out-of-sample test set (first 5 months of 2024). The deterministic formulation is almost equivalent, the main difference is that the adaptive stochastic optimization model includes scenarios for second stage decision variables. The deterministic model could essentially be seen as the same formulation with only one scenario and one stage decision variables. Thus, only the adaptive stochastic model formulation will be covered.

<sup>2</sup>Police Data Initiative, “Cambridge, Massachusetts Police - Police Data Initiative,” Police Data Initiative, 12 June 2020. Accessed 5 Dec. 2024.

### 3.2 Formulation

The adaptive stochastic optimization formulation includes first-stage decision variables for:

1. Allocating police stations.
2. Assigning police officers to police stations.

There are two possible sizes for the police stations: a large station, determined by the number of officers assigned, which has a coverage area of 2000 meters, and a small station, which has a coverage area of 1000 meters. Only one station must be small, while the remaining stations may be large. The coverage distance is based on road distance. A crime is considered to be responded to late if it occurs outside the coverage area of the station sending officers. It is considered on time if it falls within the coverage radius. Let  $I$  denote the set of all the 900 nodes that constitute the graph. Let  $J$  denote the set of candidate nodes. Let  $S$  denote the set of scenarios. Let  $C$  denote the set of crime types.

The second-stage decision variables are:

1. Assigning officers from station  $j$  to respond to crimes at node  $i$  for crime type  $c$  in each scenario  $s$ .
2. The number of crimes responded to too late at node  $i$  for crime type  $c$  in each scenario  $s$ .

The scenarios represent distinct outcomes for each node within the graph, capturing variations in crime frequency.

#### *First Stage Decision Variable*

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

$y_j \in \mathbb{Z}_+$ : 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*

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

$u_{ics} \in \mathbb{Z}_+$ : 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.

$z'_{jics} \in \mathbb{Z}_+$ : 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  $c$ .

$\pi_c$ : 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.

---

## Objective Function

The cost of our optimization problem is to minimize the number of crimes that police arrive to late within a specified time period and the transportation cost of when dispatching police officers. We assume police officers travels in pairs.

$$\min \sum_{s \in S} \lambda_s \left( \sum_{i \in I} \sum_{c \in C} \pi_c u_{i,c,s} + t \sum_{j \in J} \sum_{i \in I} \sum_{c \in C} d_{ij} \frac{(z_{j,i,c,s} + z'_{j,i,c,s})}{2} \right)$$

$\lambda_s$  is weight put on each scenario, which is  $\frac{1}{\text{Number of Scenarios}}$ .

## Constraints

Restrict the number of Police stations that are built to be exactly  $k$  (Note that we tested for  $k = 2$  and  $k = 3$ ) by enumerating over the candidate node set,  $J$ .

$$\sum_{j \in J} x_j = k$$

Ensure that the number of assigned police is 281, which reflect the current staffing conditions of Cambridge Police Department <sup>3</sup>.

$$\sum_j y_j = 281$$

Only assign officers to stations that were actually built.

$$y_j \leq 281x_j, \quad \forall j \in J$$

Create a minimum threshold for the number of police officers that should be assigned to a police station. This is to make sure we have enough police officers to dispatch. Here 40 was chosen as an appropriate level.

$$y_j \geq 40 - M(1 - x_j), \quad \forall j \in J$$

If station  $j$  has an extended coverage, it must have at least  $P$  officers assigned. When  $k = 3$  we set  $P = 100$ , which allows two police stations with an extended radius. When  $k = 2$ , we set  $P = 141$ , which allows one police station with an extended radius.

$$y_j \geq Ps_j \quad \forall j \in J$$

Create an upper bound for how many officers can be assigned to a station with an unextended coverage to ensure that only police stations with extended coverage have more than  $P$  police officers.

$$y_j \leq (P - 1)(1 - s_j) + 281s_j \quad \forall j \in J$$

At most  $k - 1$  of the stations can have an extended coverage.

$$\sum_j s_j = k - 1$$

Defining an expression for the coverage radius for each police station, depending on whether it has extended coverage area:

$$D_j = 1000 + 1000 \cdot s_j \quad \forall j \in J$$

If the distance from a police station  $j$  to node  $i$  is beyond the coverage area for the police station, then  $a_{j,i}$  will be 1:

$$\frac{d_{ij} - D_j}{M} \leq a_{ji} \quad \forall i \in I, \quad \forall j \in J$$

If  $a_{j,i} = 1$  then police officers from police station on node  $j$  will not be able to assign police officers on time for crimes happening in node  $i$ :

---

<sup>3</sup>Police Data Initiative, "Cambridge, Massachusetts Police - Police Data Initiative," Police Data Initiative, 12 June 2020. Accessed 5 Dec. 2024.

---


$$z_{jics} \leq M \cdot (1 - a_{ji}) \quad \forall j \in J, \quad \forall i \in I, \quad \forall c \in C, \quad \forall s \in S$$

If  $a_{j,i} = 1$  then police officers from police station on node  $j$  can be late to crimes happening in node  $i$ :

$$z'_{jics} \leq M \cdot a_{ji} \quad \forall j \in J, \quad \forall i \in I, \quad \forall c \in C, \quad \forall s \in S$$

Ensure the demand for each crime type at each node is either satisfied on time or counted as a late response:

$$\sum_{j \in J} (z_{jics}) + r_c \cdot u_{ics} = r_c \cdot N_{ics} \quad \forall i \in I, \quad \forall c \in C, \quad \forall s \in S$$

Ensure that the number of late police officers equates the number of cops needed for the crimes that were not met on time:

$$\sum_{j \in J} z'_{jics} = r_c \cdot u_{ics} \quad \forall i \in I, \quad \forall c \in C, \quad \forall s \in S$$

Ensure that only candidate nodes with a police station can dispatch on-time and late officers.

$$\sum_{i \in I} \sum_{c \in C} (z_{jics} + z'_{jics}) \leq M \cdot x_j \quad \forall j \in J, \quad \forall s \in S$$

### 3.3 Scenarios

The scenarios are based upon a sample average approximation, so every scenario is weighted equally. There are many different ways to generate scenarios. In this project, for simplicity, the scenarios are based upon intervals. We use two different interval versions for scenario generation:

- 1) Scenario 1 is crime in 2022 for each node. Scenario 2 is crime in 2023 for each node. Then for the deterministic model the crime will be the crime occurred in 2022 and 2023. This version omits the COVID years since the crime dynamics and frequency are outliers. Thus, it is not expected with high probability that covid scenarios will occur in the near future. This scenario version will both be used for the number of police stations equal to 2 and 3.
- 2) Scenario 1 is crime in 2020 for each node, Scenario 2 is crime in 2021 for each node, Scenario 3 is crime in 2022 for each node, and Scenario 4 is crime in 2023 for each node. Then for the deterministic model the crime will be the crime occurred in 2020 and 2023. This version includes the COVID years to see how the adaptive stochastic performs against the deterministic and baselines when including non-likely future scenarios.

### 3.4 Baseline

We employ four different baselines models that do not utilize optimization in locating police stations. Let  $k$  be the number of police stations we are set to allocate among the candidate nodes. The first is randomly choosing  $k$  candidate nodes to place the police stations. The second benchmark model is simply assigning the stations to be at the  $k$  highest crime occurring nodes. The third benchmark is using K-Means to cluster the map into  $k$  clusters and then assign a police stations to closest node to each centroid. The final benchmark is assigning one station to where the current Cambridge Police Station is and randomly selecting the remaining  $k-1$  candidate stations.

---

## 4 Results

The analysis compared six models developed to optimize police station placements and officer allocations. Each model was assessed based on its effectiveness in minimizing late crimes and late officer responses while maximizing on-time officer deployments. Evaluations were conducted on both the training data (2022-23 or 2020-23) and a test set representing unseen data from 2024 (first 5 months).

Three primary versions were considered:

- **Placement of 2 Police Stations** using 2022-23 as in-sample data.
- **Placement of 3 Police Stations** using 2022-23 as in-sample data.
- **Placement of 3 Police Stations** using 2020-23 as in-sample data (to include COVID years).

These versions aimed to identify the best strategies for enhancing police response times and optimizing officer allocation based on the demands of the city of Cambridge.

For each model and for each version, we will evaluate the in-sample performance and out-sample performance, which includes the total cost incurred (objective value from optimization problem) and the number of crimes it was late to. Further, we will report the number of police officers that were late to a crime, and the number of police officers that were quick enough to respond to a crime. It should be noted that the out sample results were based upon the optimal police station locations and police assignments from the in sample data, and then re-optimizing the dispatchment of police officers to a specific crime. Moreover, the in-sample run for the non adaptive methods had the second stage variables re-optimized like we did in homework 4 with the deterministic model to make the objective comparable with the adaptive model for the in-sample data.

	Cost	Late Crimes	Late Officers	On Time Officers
<b>Random Baseline</b>	\$274,390,730	10,128	48,730	15,346
<b>Highest Crime Baseline</b>	\$141,894,483	5,295	25,070	39,006
<b>kMeans Baseline</b>	\$142,500,415	5,160	24,854	39,222
<b>Present Day Baseline</b>	\$134,582,595	5,076	23,990	40,086
<b>Deterministic</b>	\$110,995,394	4,243	19,968	44,108
<b>Adaptive</b>	\$110,995,394	4,243	19,968	44,108

**Table 1:** *2022-23 Model Comparisons (2 stations)*

	Cost	Late Crimes	Late Officers	On Time Officers
<b>Random Baseline</b>	\$112,503,017	2,145	10,378	3,106
<b>Highest Crime Baseline</b>	\$63,212,389	1,064	5,284	8,200
<b>kMeans Baseline</b>	\$63,532,116	1,056	5,320	8,164
<b>Present Day Baseline</b>	\$60,596,770	1,028	5,100	8,384
<b>Deterministic</b>	\$48,603,011	843	4,124	9,360
<b>Adaptive</b>	\$48,603,011	843	4,124	9,360

**Table 2:** *2024 Model Comparisons (2 stations)*



---

	Cost	Late Crimes	Late Officers	On Time Officers
Random Baseline	\$164,231,437	6,189	29,310	34,766
Highest Crime Baseline	\$113,432,700	4,169	19,872	44,204
kMeans Baseline	\$82,962,397	3,030	14,454	49,622
Present Day Baseline	\$105,465,225	4,047	18,998	45,078
Deterministic	\$66,178,291	2,518	11,700	52,376
Adaptive	\$33,945,299	1,169	5,628	58,448

**Table 3:** *2022-23 Model Comparisons (3 Stations)*

	Cost	Late Crimes	Late Officers	On Time Officers
Random Baseline	\$70,476,471	1,242	6,114	7,370
Highest Crime Baseline	\$50,535,542	827	4,152	9,332
kMeans Baseline	\$36,823,531	635	3,138	10,346
Present Day Baseline	\$47,727,523	836	4,094	9,390
Deterministic	\$29,269,493	513	2,482	11,002
Adaptive	\$14,508,239	234	1,152	12,332

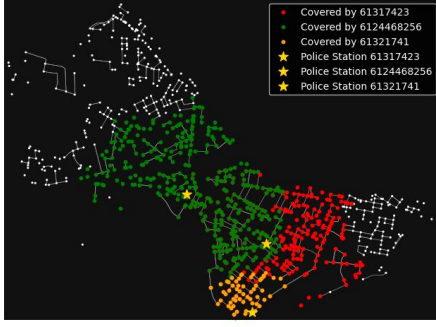
**Table 4:** *2024 Model Comparisons (3 stations)*

	Cost	Late Crimes	Late Officers	On Time Officers
Random Baseline	\$147,398,008	12,080	54,818	60,656
Highest Crime Baseline	\$97,051,298	7,959	35,866	79,608
kMeans Baseline	\$73,917,687	5,928	27,088	88,386
Present Day Baseline	\$64,960,339	5,809	25,334	55,942
Deterministic	\$50,501,326	4,294	18,952	96,522
Adaptive	\$60,716,857	5,023	22,476	92,998

**Table 5:** *2020-23 Model Comparisons (3 Stations with COVID years)*

	Cost	Late Crimes	Late Officers	On Time Officers
Random Baseline	\$70,455,471	1,238	6,102	7,382
Highest Crime Baseline	\$50,355,042	828	4,130	9,354
kMeans Baseline	\$36,512,531	630	3,114	10,370
Present Day Baseline	\$47,257,523	831	4,062	9,422
Deterministic	\$27,118,916	464	2,280	11,204
Adaptive	\$31,682,010	526	2,640	10,844

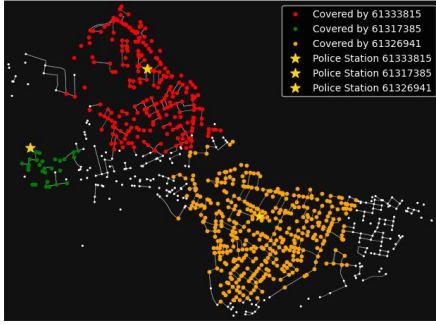
**Table 6:** *2024 Model Comparisons (3 stations with COVID years)*



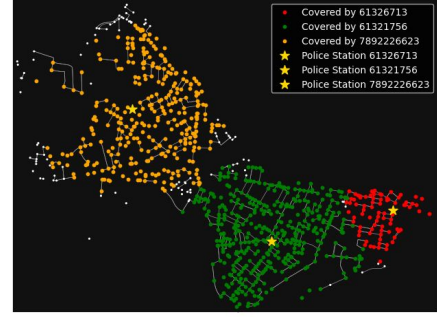
**Figure 4:** Current Baseline (3 Stations)



**Figure 5:** K-Means (3 Stations)



**Figure 6:** Deterministic (3 Stations)



**Figure 7:** Adaptive (3 Stations)

**Case with Two Stations:** The Adaptive and Deterministic models performed identically, achieving the same number of late crimes, late officers, and on-time officer responses for both 2022-23 and 2024. This indicates that both methods identified the same optimal locations for the two stations and allocated officers in a similar manner. These models outperformed the other baselines in this scenario. As expected, the overall performance with two stations was worse compared to the three-station case, as fewer stations limited the coverage area and response capabilities.

**Case with Three Stations:** With three stations, the adaptive and deterministic yielded different solutions. The Adaptive model was the top performer in 2022-23, achieving only 1,169 late crimes. The Deterministic model ranked second, followed by the k-means baseline, present day baseline, the crime hotspot baseline, and finally the random placement baseline. For 2024, the Adaptive model continued to lead, reducing late crimes to 234 and outperforming all other approaches in both datasets. This shows that including multiple potential scenarios in your optimization model can lead to better generalizable solutions.

**Case with Three Stations Using COVID Years:** Including 2020 and 2021 as scenarios, along with 2022 and 2023, actually made the Adaptive Model's out of sample performance worse than the Deterministic Model compared to using only 2022 and 2023. This is likely because the COVID years were not representative of current or future crime patterns. This shows that uncalibrated scenarios can hinder the performance of the adaptive stochastic method's solution.

**Visualizations:** Figure 4 illustrates the scenario of randomly adding 2 stations to the existing one. This approach results in the least coverage and the highest number of late crimes among the four strategies shown. The K-means method (Figure 5) achieves relatively good coverage of nodes

---

on time but does not prioritize station placement based on node-specific crime hotspots, leading to suboptimal results. In comparison, the deterministic approach (Figure 6) places stations more effectively in terms of covering crime hotspots. Interestingly, the adaptive method (Figure 7) provides nearly complete coverage and achieves the best performance overall, highlighting its superior station placement strategy.

**Key Takeaway:** In most cases, adding 1-2 police stations randomly in addition to the already existing Cambridge police station performed better than the random and highest crime baselines, though not as well as the deterministic or adaptive optimization methods, reinforcing the strength of optimization. The benefits of the Adaptive model are particularly evident with three stations, showing nearly a 50% reduction in late crimes in 2022-23 and continued superior performance in 2024. Notice that the expected percentage value of the stochastic solution (2022-2023 data with 3 police stations) insample was 48.7% which approximately aligns with the actual value of the stochastic solution out of sample of 50.4%. This highlights the benefit of including accurate scenarios in optimizing response times and resource allocation. However, COVID years made the Adaptive’s solution suboptimal, showing the risk of including uncalibrated scenarios.

## 5 Conclusion

The analysis conducted provides valuable insights into optimizing police station placements and officer allocation in Cambridge to ensure rapid responses to crimes. Prioritizing high-priority crimes, followed by medium and low-priority incidents, is crucial for maximizing response effectiveness and enhancing public safety.

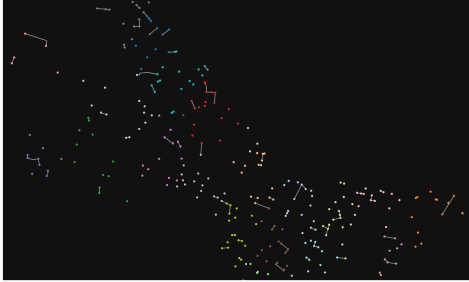
The analysis revealed that adaptive optimization significantly outperformed other strategies, demonstrating performance that was twice as effective as deterministic optimization and nearly five times more efficient than random station placement in Cambridge. It was also 2.5 times better than the K-means baseline and 3.5 times more effective than the highest crime baseline. This was true for both the 2022-23 historical data and the unseen 2024 data. In 2022-23, adaptive optimization resulted in a cost of \$33,945,299 and 1,169 late crimes. In 2024, it resulted in costs of \$14,508,239 and 234 late crimes. Compared to other models, these improvements would mean significant savings in expenditures and a substantial reduction in late crimes, highlighting adaptive optimization as a strategic approach that enhances response times and resource efficiency. To achieve this value from adaptive stochastic optimization, the scenarios also need to be not too inaccurate as it can deter the generalizability of the adaptive’s solution.

If given an additional week to continue this project, the team would conduct a detailed evaluation of the potential benefits and trade-offs of scaling the model to larger cities. This would include a more precise quantification of the costs associated with expansion compared to the current model. The analysis would use the same metrics—late crimes, late officer responses, and on-time officer responses—focusing on the years 2022-23 and 2024. This deeper analysis would offer a clearer understanding of how strategic expansion could affect crime response times and overall resource allocation on a larger scale.

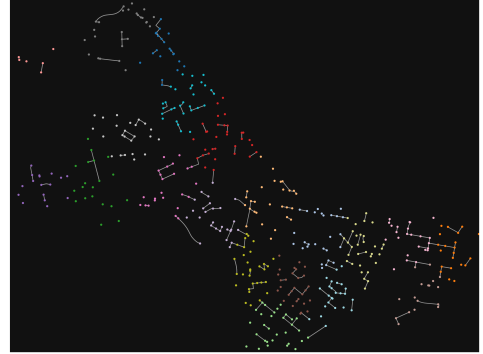
---

## 6 Appendix

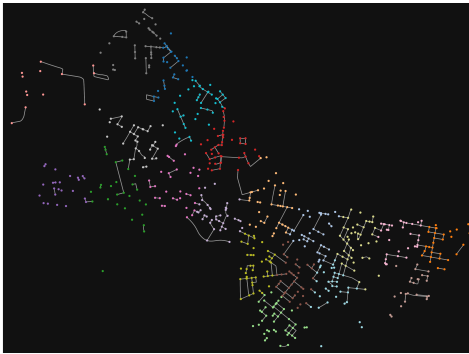
### 6.1 Down Sampling From 1869 Nodes



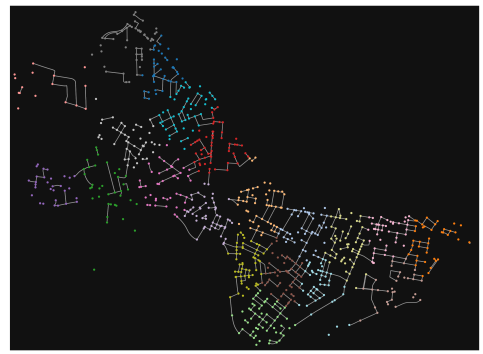
**Figure 8:** 300 Nodes



**Figure 9:** 500 Nodes

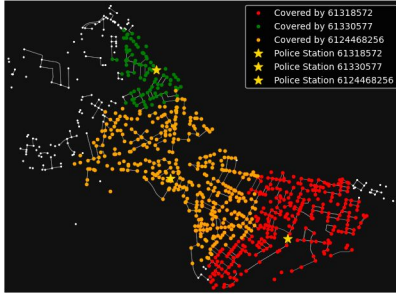


**Figure 10:** 700 Nodes

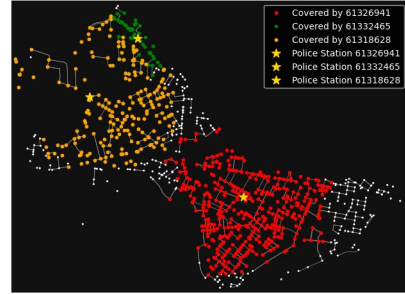


**Figure 11:** 900 Nodes

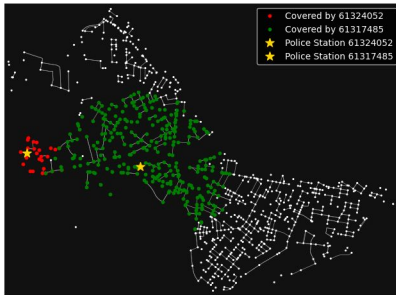
## 6.2 Additional Police Station Locations and Coverage



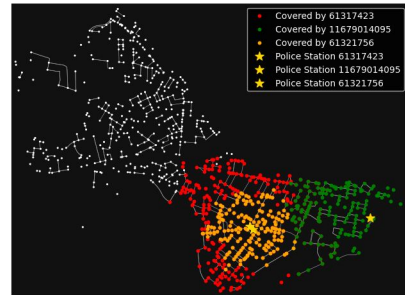
**Figure 12:** Deterministic Using COVID Years (3 Stations)



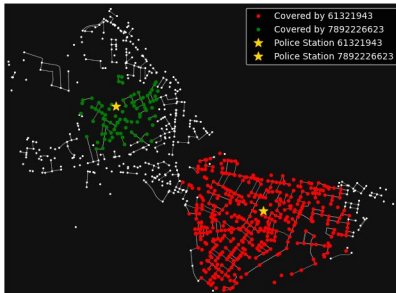
**Figure 13:** Adaptive Using COVID Years (3 Stations)



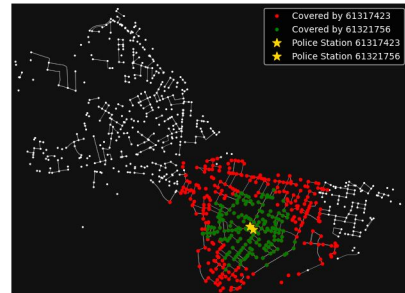
**Figure 14:** Random Baseline (3)



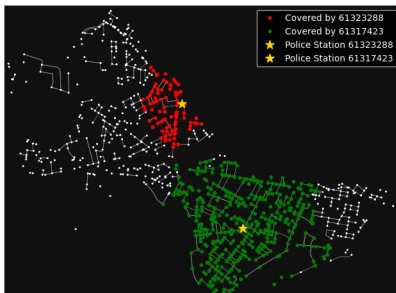
**Figure 15:** Top 3 Crimes Baseline (3)



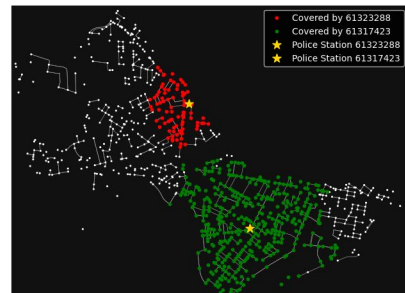
**Figure 16:** K-Means (2)



**Figure 17:** Top 3 Crimes Baseline (2)



**Figure 18:** Deterministic (2)



**Figure 19:** Adaptive (2)

---

**Table 7: *Crime Severity Assignments***

<b>Crime Type</b>	<b>Category</b>	<b>Number of Officers Needed</b>
Hit and Run	H	8
Forgery	L	2
Larceny of Bicycle	L	2
Shoplifting	M	4
Larceny from MV	M	4
Larceny from Residence	H	8
Mal. Dest. Property	M	4
Simple Assault	H	8
Accident	M	4
Flim Flam	L	2
Warrant Arrest	M	4
Threats	M	4
Admin Error	L	2
Larceny from Building	H	8
Aggravated Assault	H	8
Housebreak	H	8
Missing Person	M	4
Harassment	M	4
Auto Theft	M	4
Larceny from Person	H	8
Suspicious Package	M	4
Street Robbery	M	4
Larceny (Misc)	H	8
Commercial Break	L	2
Trespassing	M	4
Drugs	M	4
Violation of H.O.	M	4
OUI	L	2
Indecent Exposure	L	2
Larceny of Plate	M	4
Extortion/Blackmail	L	2
Disorderly	L	2
Rec. Stol. Property	M	4
Drinking in Public	L	2
Commercial Robbery	H	8
Phone Calls	L	2
Larceny of Services	M	4
Counterfeiting	L	2
Noise Complaint	L	2
Weapon Violations	L	2
Arson	H	8
Annoying & Accosting	L	2
Prostitution	M	4
Taxi Violation	L	2
Sex Offender Violation	H	8
Peeping & Spying	L	2
Embezzlement	L	2
Stalking	L	2
Liquor Possession/Sale	L	2
Kidnapping	H	8
Homicide	H	8
Violation of R.O.	M	4