Optimizing Police District Allocation in Chicago

Pontus Fredstam
M.Sc. Sociotechnical systems
Email: Pfredst@upv.edu.es

Mateusz Król
M.Sc. Computer Science
Email: mkrol@upv.edu.es

Arne Vanheule
M.Sc. Computer Science
Email: avanheu@upv.edu.es

Email:

***Abstract*—In our work we present a critical challenge of optimizing police district allocation in Chicago using Mixed Integer Linear Programming (MILP). By developing a novel optimization model, we propose a data-driven approach to redistribute 23 police district assignments across community areas while minimizing response times and balancing crime workloads. Our model introduces a sophisticated framework that considers crime density, geographical proximity, and operational constraints, offering a transformative methodology for urban public safety resource management. In our approach we concentrate on minimizing the total response time needed for the police officers to get to the crime-weighted places throughout the areas the particular police stations are responsible for. Our goal is to maintain an equal workload balance across all the police districts, so that none of the police stations is overloaded with the amount of work, what currently is the fundamental problem existing in the city. The system is implemented using the OR-Tools optimization library with the SCIP solver. The implementation of our model and its execution let us establish new areas of responsibility for particular police stations providing more workload balance based on the density and amount of crimes appearing in such areas.**

I. Introduction

Urban crime management requires sophisticated resource allocation strategies. Chicago, with its complex urban landscape of 77 community areas and 23 police districts, encounters a challenging optimization problem. Traditional district assignments often fail to account for dynamic crime patterns and workload distribution. Those are the two major factors that procure the professional burn-out among the Chicago police officers, that is why it is needed to observe consistently the changes that appear throughout the city and adjust the distribution of police districts appropriately. After applying different policies in order to tackle crime within the city, we can come to the conclusion that perhaps the structure and organization of the police departments is outdated according to how crime in the city has evolved (in type and quantity) in recent years. It is for this reason that it is needed to reorganize the distribution of police districts and areas in an effective and efficient manner. Therefore, it is necessary to determine the new structure of the 23 districts and the corresponding 5 areas on the basis of objective criteria.However, there exists some requirements to be fulfilled. The first of them is to maintain exactly 23 police districts whose base of operations will be each of the 23 police stations. Next, we should consider that the complexity of coordinating these areas is related to the number of associated police districts as well as their characteristics. And then, the definition of a police district determines which areas of the city the available police officers will patrol, as well as which areas of the city the available officers will go to when incidents occur. Figure 1 demonstrates the number of crimes in respective community areas in the city. Darker colour represents highest number of offences.

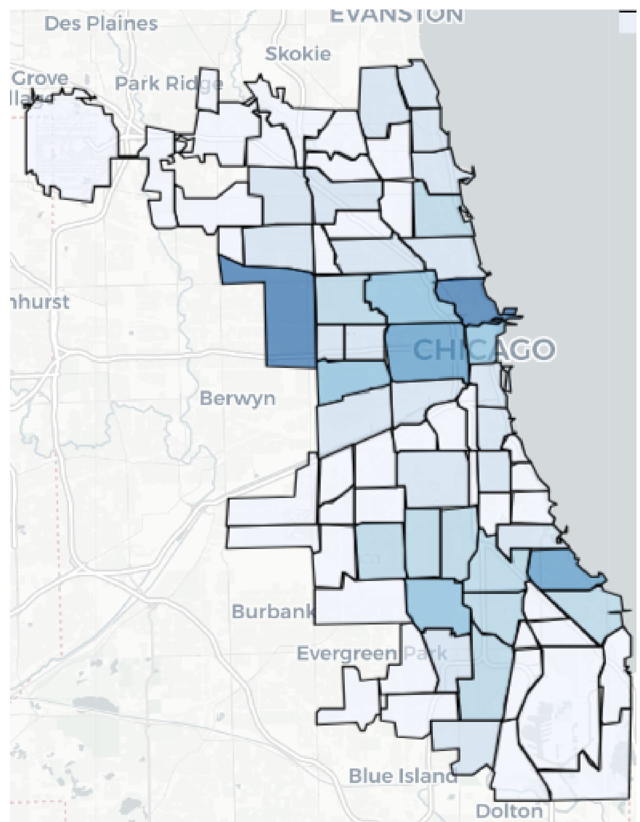

Fig. 1. Number of crimes in Chicago depending on the community area

To carry out the project we have been given a series of datasets, which contain information about the current crime structure in the city:

- The stations.json file contains information for each of the 23 police stations that the city of Chicaco has. It is a collection of JSON objects, where each object contains information for one of the stations in the city.
- The community-areas.json file contains information for the 77 community areas in the city of Chicago. It is a collection of JSON objects, where each object contains information for one of the community areas.

- The distances.json file contains information on the distances from each of the 23 police stations to each of the centroids of the 77 community areas. This distance is measured in seconds of travel by car. It is a collection of JSON objects, where each object contains information about the distances from one of the stations to the different community areas.
- The last one is the crimes.json file, which contains information about the crimes that occurred in the city of Chicago in the last year. It is a collection of JSON objects, where each object represents a crime.

To solve the problem we used the Mixed Integer Linear Programming (MILP) strategies with the SCIP solver implemented in the OR-Tools optimization library. In order to apply this solution and visualize the results we used the Python environment in the Google Colaboratory notebook.

Our research aims to:

- Develop a mathematical model for optimal police district allocation
- Minimize crime-weighted response times
- Ensure balanced workload across districts
- Provide a reproducible methodology for urban public safety optimization

We begin with a description of the proposed mathematical model to solve the optimization problem. Then we focus on explaining the implementation details as well as the key strategies we used to create a concise model. Next section includes all the experiments we conducted in order to obtain the data needed for further analysis of the behaviour of the model. Later in the conclusions section we can find the summary of the performance of the model and its results. Finally in the last section there are described the possibilities of future work on the current system and its potential improvements.

II. Proposed Optimization Model

A. Decision Variables

$$X_{sc} = \begin{cases} 1, & \text{if community area } c \text{ is} \\ & \text{assigned to police station } s, \\ 0, & \text{otherwise.} \end{cases}$$

Binary variables $X_{s,c}$ (in the code referred to as $community\_assignment_{s,c}$) are chosen because the problem involves a yes/no decision (i.e., whether a police district covers a community area or not). This simplifies the model and ensures clarity in the optimization objectives.

$$Y_{a,s} = \begin{cases} 1, & \text{if police station } s \text{ is} \\ & \text{assigned to area } a, \\ 0, & \text{otherwise.} \end{cases}$$

Binary variables $Y_{a,s}$ (in the code referred to as $area\_assignment_{a,s}$) are provided to divide all the police stations into 5 different police areas and ensure that we have a clear division of the whole city.

$$W_a, \text{where } W_a \in \mathbb{R}$$

is the workload assigned to area a. This decision variable is in the code called $area\_unbalanced\_workload_a$.

$$Z_{a,s,c} = \begin{cases} 1, & \text{if total crime of community area } c \\ & \text{ is assigned to station } s \text{ in area } a \\ 0, & \text{otherwise.} \end{cases}$$

Binary variables $Z_{a,s,c}$ (in the code referred to as $area\_station\_workload\_assignment$) are provided to divide the crime of each community to a station in an area.

B. Constants

- $d_{s,c}$: Distance in seconds between police station s and community area c, referred to as $distance\_station\_communitys, c$ in the code.
- $crimes\_per\_area_c$: Crime frequency in community area c.
- $\overline{crimes}$: Average amount of crimes in all the community areas, referred to as $avg\_crimes\_per\_community$ in the code.
- α: Maximum percentage above the average crime rate that each police district is allowed to handle, referred to as $max\_police\_station\_overload$ in the code.
- $max\_allowed\_distance$: This is the maximum distance between two police stations that can be in the same police area.

C. Constraints

1) Single Assignment Constraint:

$$\sum_{s=1}^{23} X_{s,c} = 1 \quad \forall c \in \{1, \ldots, 77\}$$

This constraint ensures that each community area is assigned to exactly one police district. It prevents a community area from being left without coverage or being assigned to multiple districts.

2) Workload Balance Constraint:

$$\sum_{c=1}^{77} crimes\_per\_area_c \cdot X_{s,c} \leq \alpha \cdot \overline{crimes} \quad \forall s \in \{1, \ldots, 23\}$$

This constraint ensures that no single police station is responsible for too many high-crime areas. It works by comparing the total crime rate assigned to a police district with the average crime rate of the community areas. This helps distribute the workload more fairly and prevents any district from being overwhelmed.

3) Station Home Community Constraint:

$$X_{s,crimes\_per\_area_c} = 1, \quad \forall s \in \{1, \ldots, 23\}$$

Where $crimes\_per\_area_c$ is defined as the mapping function that gives the community area associated with police district s. This constraint ensures that the community area where the police station is located must be assigned to this station as one of its areas of surveillance. It prevents the situation where the police officers that work in one particular station and most probably know the problems of this area the best would not be in charge of handling the crime offences in this area.

4) Single Assignment of Areas to Police Stations:

$$\sum_{a=1}^{5} \mathrm{Y}_{a,s} = 1 \quad \forall s \in \{1, \ldots, 23\}$$

This constraint ensures that each police station is assigned to exactly one police area. It prevents a police station from being assigned to multiple areas and guarantees that every station is responsible for one community area.

5) Logical "AND" Constraint for Crime Rate Assignment:

$$Y_{a,s} + X_{s,c} \geq 2 \cdot Z_{a,s,c},$$
$$\forall a \in \{1, \ldots, 5\}, \quad s \in \{1, \ldots, 23\}, \quad c \in \{1, \ldots, 77\}$$

$$Y_{a,s} + X_{s,c} \geq 2 \cdot Z_{a,s,c},$$
$$\forall a \in \{1, \ldots, 5\}, \quad s \in \{1, \ldots, 23\}, \quad c \in \{1, \ldots, 77\}$$

These constraints create an assignment variable for crime rate, which is a logical "and" between $area\_assignment$‘ and $community\_assignment$. This ensures that a community area and a police station are only linked to a crime rate workload if both the area and community assignments are active. It enforces the condition that both assignments must be true for the crime rate workload assignment to be valid.

6) Maximum Distance Constraint:

$$\mathrm{d}_{s_1,s_2} \leq \text{max\_allowed\_distance},$$
$$\forall s \in \{1, \ldots, 23\}, \quad s_1, s_2 \in \{1, \ldots, 23\}, \quad s_1 \neq s_2$$

This constraint ensures that the distance between any two assigned police stations (s_1 and s_2) does not exceed a defined maximum allowed distance. If the distance between two stations exceeds this limit, the area cannot be assigned to both stations at the same time.

7) Unbalanced Workload Constraint:

$$\sum_{s=1}^{23}\sum_{c=1}^{77} \mathrm{Z}_{a,s,c} \cdot \text{crimes\_per\_area}_c - \overline{\text{crimes\_per\_area}} \leq \mathrm{W}_a, \quad \forall a \in \{1, \ldots, 77\}$$

$$\overline{\text{crimes\_per\_area}} - \sum_{s=1}^{23}\sum_{c=1}^{77} \mathrm{Z}_{a,s,c} \cdot \text{crimes\_per\_area}_c \leq \mathrm{W}_a, \quad \forall a \in \{1, \ldots, 77\}$$

These two constraints ensure that the total workload assigned to each area does not deviate too far from the average workload. The first equation ensures the workload does not exceed the average by more than a specified threshold, and the second one ensures the workload is not less than the average by more than the same threshold.

D. Objective Function

Objective Function =

$$\min\Big(+ \sum_{s,c} (X_{sc} \cdot \text{distance\_station\_community}[s,c] \cdot \text{crimes}$$

$$\min Z : \sum_{s=1}^{23}\sum_{c=1}^{77} d_{sc} \cdot crimes\_per\_area_c \cdot X_{sc}$$

The objective function incorporates three key components: the travel time in seconds between police station s and community area c , the crime rate in community area c , and the decision variable X_{sc}. The function is designed to minimize the total weighted response time across all community areas. By including the crime rate as a weighting factor, the model prioritizes areas with higher crime rates, ensuring a more efficient allocation of resources to locations where timely responses are most critical. On the other hand it is necessary to put an emphasis on places, which have more crime occurrences, because there the police needs to more often, so that their weight should be adjusted accordingly.

III. Implementation Details

Our implementation leverages OR-Tools with SCIP solver:

The algorithm begins by extracting relevant data, such as crime statistics and geographical distances, from JSON files. This structured format ensures all required information is collected efficiently. Once the data is extracted, it undergoes preprocessing to ensure it is in the correct format for further computation. This step includes computing necessary aggregates like crime rates or average distances between communities and police stations. Then the solver is initialized. The algorithm uses the SCIP optimization solver, a robust framework designed for handling complex mathematical and

Algorithm 1 Police Station Allocation Algorithm

Extract the data from the JSON files
Process the datasets to obtain necessary data
Initialize solver with SCIP
Create decision variables
Calculate average crime rate $\overline{crimes}$
Add single assignment constraints
Add workload balance constraints
Add station-community constraints
Set objective: minimize weighted response time
Solve optimization problem
return Station assignments

integer programming problems. Initializing the solver involves setting up the environment to define variables, constraints, and the optimization objective. Now all the 4 decision variables are created. This step includes creating a structured way to represent all the combinations of possible decision variables in the model. The average crime rate across all communities is calculated. This serves as a baseline metric to help ensure equitable distribution of resources among police stations, accounting for varying crime levels. The next step provides the constraints to the system. It is divided into 3 kinds of them:

- Single Assignment Constraints - each community is assigned to exactly one police station to prevent overlap or unallocated regions.
- Workload Balance Constraints - these ensure that the workload, often influenced by crime rates, is distributed evenly across police stations to prevent overburdening any single station.
- Station-Community Constraints - additional constraints may include maximum distances a station can serve or restrictions based on capacity.

In the next part we focus on defining the objective function. The optimization goal is to minimize the weighted response time across all station-community pairs. This ensures a fair and efficient allocation that prioritizes timely responses, especially for high-crime areas. With the constraints and objective in place, the solver computes the optimal allocation of communities to police stations. This involves addressing the complexities of integer programming, where variables can only take on whole values. At the end the model concludes by outputting the optimized assignments, detailing which communities are served by which police stations.

IV. Experiments

Experimental validation included:

- Sensitivity analysis of workload balance parameter
- Analysis of impact of the maximal distances between stations within one police station on the total response time
- Comparative analysis with current station allocation

To conduct experiments on our model we decided to check its performance based on different parameter values. The two coefficients we were testing were:

- Maximal overload of all the police stations
- Maximal station-station distance factor

First parameter is responsible for approving only those solutions, which provide the overload of particular police area using percentage over the average workload. The other one describes the relationship between two police stations responsible for surveillance on the same police districts. The distance between them cannot surpass the average distance multiplied by this factor.

V. Conclusions

Our MILP model demonstrates:

- Potential for data-driven district optimization
- Significant improvements in response time allocation
- Robust framework for urban safety resource management

Based on the performed experiments on changing the values of $max\_overload$ and $max\_station\_station\_distance$ parameters and observing their impact on total unbalance factor we can conclude that:

- As the maximum overload increases, there seems to be a general decrease in total unbalance, up to a point of the value about 1.3 over the average workload.
- As $max\_station\_station\_distance$ increases, the total unbalance tends to decrease, but after surpassing certain value (about 0.5) our goal coefficient tends to rise significantly
- The combination of $max\_overload = 1.3$ and $max\_station\_station\_distance = 0.48$ gives the lowest total unbalance. This suggests that balancing these two parameters is key to minimizing unbalance.

At the end we managed to achieve a new allocation of the police districts that is visualized on the Figure 2. The map represents how the community areas should be reorganized. One station is responsible for all the community areas highlighted with the same colour.

VI. Future Work

To better reflect real-world scenarios, the model can be expanded to include factors such as the type of crime (e.g., domestic vs non-domestic) and its effect (person arrested or not). Such extensions would enable prioritization of resource allocation based on the criticality of incidents, ensuring that high-severity crimes receive immediate attention. This would require a robust categorization of crimes and potentially integrating qualitative and quantitative metrics for severity assessment. Furthermore, to validate and improve the proposed approaches, a detailed comparative analysis with other existing solvers like Gurobi or CPLEX methodologies can be conducted. This could include benchmarking against traditional optimization solvers, machine learning-based frameworks, or heuristic approaches used in resource allocation problems. Metrics for comparison might include computational efficiency, scalability, solution quality, and robustness under different conditions.

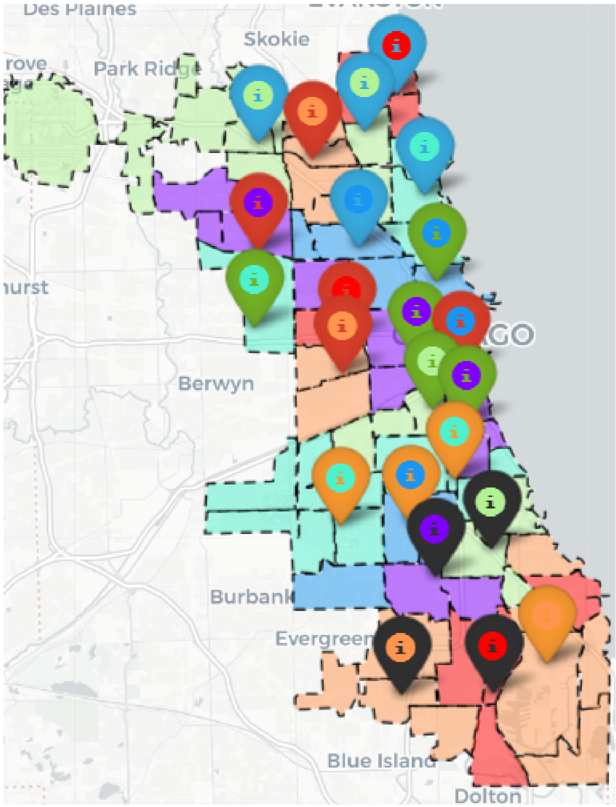

Fig. 2. New allocation of police districts in Chicago.

References