**Carnegie Mellon University Police Reallocation**

Andy Chung

Jee Yeon Shin

Jieun Lee

Jun Eung Kak

**Table of Contents**

**i. Abstract** 1

**I. Problem Statement** 1

Introduction to the Problem 1

Current System 2

**II. Data Analysis** 2

**III. Solution** 3

Assumption 3

Formulation 4

Implementation 4

Interpretation 5

**IV. Conclusion** 6

Comparison 6

Limitation and Further Studies 6

Conclusion 7

**V. Appendix**

 A ………………………………………….. Crime Log

 B ………………………………………….. Histograms

 C …………………………………………... Schedule

 D …………………………………………… Number of Crimes

 E …………………………………………… Solution

 F …………………………………………… Solution with the Map

**Abstract**

 The goal of this research was to find the optimal placement of police officers such that they cover the maximum crime rates within the areas where Carnegie Mellon University Police Department has jurisdiction. In order to approach the problem, crime rates at each approximate location was found. To ensure that each officer was patrolling an area that they could respond within a predetermined time constraint, each location was given a set of neighboring locations within a circular area. Using these sets of neighboring locations, the maximal coverage algorithm was used to determine the optimal placement of police officers such that they cover the areas with high crime rates, given crime rates in each location.

**I. Problem Statement**

Introduction to the Problem

The idea for this project was introduced through a click of an email with the simple subject: “Crime Alert”. The email described the scene as an armed assault on North Craig Street which is only a few blocks off of Carnegie Mellon’s campus. Although the victim was uninjured, the email alone inevitably planted a drop of fear into every student.

Throughout each semester, it is likely that each student receives at least a few emails about dangerous crimes that occur near campus. According to the annual reports on Carnegie Mellon University Police Department’s (CMUPD) website, the crime rate has generally been increasing near Carnegie Mellon each year. Aside from the rare violent and potentially lethal crimes, CMUPD also deals with other accidents and lesser crimes such as fire, theft, and drug abuse. Ideally, the police would want to forestall as much of these accidents and crimes as possible, but realistically, preventing crimes as a whole is a very difficult goal to accomplish. Is there a strategy that the police can devise to reduce the number of crimes without heavy budgeting? Our initial problem was to find a strategy such that we minimize the cost and increase the safety in the areas of jurisdiction.

There were many different ways to approach the problem. Among these algorithms, the simplest, most intuitive method was to use set covering problem such that the optimal solution covered the most area on the map. However, while the set covering problem made the problem very simple, it brought up new problems of its own: can the problem really be solved just by a need covering the most area possible, how can we determine the costs, and finally, how will we consider other factors such as population density and crime density? Being overly simple, the set covering problem ignored too many variables and we had yet to find a lead on this problem. In hopes to find some answers, we contacted CMUPD’s Lieutenant Scheimer.

Current System

 Current strategy that CMUPD currently uses, according to Lt. Scheimer, is as follows.

* Three shifts (2300-700, 700-1500, 1500-2300)
* 11 Guards either stationary or patrolling in designated areas
* Campus is divided to North and South
* No other strategy or algorithm is used

 In addition, we were given the crime and accident logs from January 2009 to early December 2011. Given the current allocation strategy and the data, it was clear that our initial problem had to change. Our new goal was to find the optimal placement of the guards such that they cover the areas with high crime rates, given crime rates in each location.

**II. Data Analysis**

 When we received the data from Lt. Scheimer, the file took up over 90000 lines on excel, filled with information that we did not need such as technicalities. Once we deleted all the unnecessary data, we categorized each data by following:

* Location
* Time, shifts
* Date
* Crime types

 Once we sorted the data, we found the crime density, and thus crime rate, for each of the categories. For our project, we specifically focused on the crime rates based on locations.

**III. Solution**

Assumption

Before we derived our model, it was necessary to simplify some of the conditions in the university policy system. We started with simplification that truncated trivial part of our data. As we sorted the data, we observed that some crimes have occurred less frequently at specific locations that were far off the main campus. Such locations include Alumni House, Business Service, electric garage, Frank Curto, Gesling Stadium, ROTC, Tartan House, Residence at 5th and Robotics Engineering Consortium. Although the university police had jurisdiction on these locations, they were too scattered away and the number of crime occurrences was trivial to be considered. So we decided to disregard the small number of crimes, and exclude them when we consider the areas needed for the placement of a police officer. This assumption was reasonable to make, since the total number of crimes at those areas summed to 170, which takes up only 2% of our total crime.

The next consideration was regarding 486 crime occurrences that were taken place off campus, where our university police has no jurisdiction. Although our university policy does respond to crime occurrence off campus areas, we assumed that the university police system is not obliged to cover such areas and removed those areas when we analyzed the crime rate of areas.

Another condition was necessary when we considered the streets and avenues that were reported as crime areas. Due to the nature of such areas, they could not be made as a sector and be covered by a police officer. Therefore, we decided the streets and avenues to be covered by an extra police officer who patrols in a car who is not assigned an area. Our assumption was realistic, since our optimal location should naturally cover part of the streets and avenues. As noted in our collected data, the type of crimes that is the most frequent on the streets and avenues deal with cars, which must be inspected by a police officer in a car.

We made our final simplification considering the police officers. We assumed that police officers share the same authority, and once they were assigned a region to cover, they patrol on foot in a random manner. This simplification was made for two reasons. First, all the police officers are given the same authority regardless of their titles, because our objective was to place a police officer to areas where crimes are the most frequent and to guarantee their response time. Therefore, police officers should not be distinguished from each other or by the type of crime they can handle. Second, police officers patrol on foot in random, because assigning them a path within the area could incur a strategic crimes, and criminals who may study the pattern of the officer’s patrol.

Formulation

Since our goal was to find an optimal placement of police officers such that it covers as much crime area as possible, we found that this problem could be expressed as a maximum coverage. First, we observed all the sites where crimes have occurred from 2009 to 2011, and numbered those sites from 1 to 92. Each site referred to a specific location. Since we wanted to obtain areas in our final solution, we grouped them by their proximity. When we grouped those sites, we also considered a response time of a police officer, and created what was called a zone. A zone was decided so that within the zone, a police officer could respond to a crime within 2 minutes. After we grouped them, we were able to reduce our final zones down to 47.

Our decision to be made was in which zones to place police officers. This was represented by a decision variable $x\_{j}$, for each zone j, which is set to 1 if the zone j is covered by a police officer, and 0 otherwise. A site that was chosen to be a candidate for placing a police officer was weighted by the total number of crime occurrences $c\_{i}$, for each candidate zone i. Another decision variable, $y\_{i}$ was used to indicate whether the candidate zone covers its crime rate.

As we stated above, we wish to maximize the crime occurrences that are covered. By using our decision variable and crime rate, this can be expressed as the following maximal coverage:

max $\sum\_{}^{}c\_{i}y\_{i } ∀ i$

subject to $ \sum\_{}^{}x\_{j}-y\_{i}\geq 0 ∀ i$

$ \sum\_{}^{}x\_{j}= n ∀j $

Our first constraint guarantees that a zone is chosen as a candidate and therefore, its crime rate is covered, only if there a police officer is placed at the zone. The second constraint limits the total number of police officers to n.

Implementation

 We used the Excel Solver to solve our problem, because the problem involved a simple objective function and a set of constraints. The Excel Solver (see Appendix E) is a direct reflection of the problem formulation. The numbers 1 to 47 denote our zones that represent area within 2 minutes response time from its center. They are circular areas of a radius 0.1 mile, and this will guarantee a police officer at 3 mph walking speed to respond to a crime within 2 minutes.

 All the zones are weighted by their demand, or the total number of crimes occurred from 2009 to 2011. In each numbered row, since the zone that corresponds to the number contains itself and its neighboring zones, their intersecting cells are marked 1’s, and the remaining cells are marked 0‘s.

Solutions

After we solved our problem with the Excel Solver, we obtained 11 optimal locations to be covered by the 11 guards (see Appendix E). These locations included: 300 Craig, Mellon Institute, Resnik, Morewood Gardens, Fraternity Quadrangle, Gate Center, University Center, East Campus Garage, The Hill, College of Fine Arts, and Wean Hall. Recall that each location represents a circular area of 0.1 mile with its center at the selected location.

Interpretation

 Our optimal locations for the placement of police officer should not be surprising. We can see it clearly that most of the areas surrounding campus is covered by the union of our sets (See appendix F). Our optimal location must cover the place where it has a high crime rate. Out of 2810 number of crimes in total, below is the number of crimes took place at each location and their rates:

|  |  |  |
| --- | --- | --- |
| **Area** | **Number of Crimes** | **Rate** |
| 300 Craig | 116 | 4.13% |
| Mellon Institute | 129 | 4.59% |
| Resnik House | 255 | 9.07% |
| Morewood Gardens | 321 | 11.42% |
| Fraternity Quadrangle | 321 | 11.42% |
| Gate Center | 230 | 8.19% |
| University Center | 380 | 13.52% |
| East Campus Garage | 255 | 9.07% |
| The Hill | 343 | 12.20% |
| College of Fine Arts | 210 | 7.47% |
| Wean Hall | 243 | 8.65% |

 Although the percentages may look small, recall that the total number of locations was 47. Therefore, our solution accurately predicts the locations where the area covered by set is high in its crime rate.

**IV. Conclusion**

Comparison

 The reason why our allocation is superior to the current allocation is that our algorithm allocates guards based on crime rates. Carnegie Mellon Police Department assumed that the places such as the University Center had high crime rates. But instead, our data showed a relatively low crime rate in the University Center itself. By reducing the number of police officers in the areas with low crime rates, our solution spreads them into areas with higher crime rates.

Limitations and Further Studies

 The biggest limitation for this research was that the center of every zone lies on top of a location due to the fact that we use binary xi and yi to denote whether the zone is used or not. If we wanted to include locations that were not one of our initial set of locations, many new problems would be introduced such as the minimum distance between the centers of zones and how we would go about solving it. For integer programming, with the application of a proper algorithm, the Excel helps solve for the final answer. However, if we changed each zone to have an indefinite center, the problem would require algorithms and codes that are beyond the scopes of this class.

 The next feasible limitation is further scrutinizing the data. If we wanted to have a more specific solution, one of many options would be to analyze the data by crime rate per hourly basis and allocate accordingly 24 times. Analyzing the data by adding on specific details will take an enormous amount of time and it may not change the results greatly. However, the biggest concern regarding this limitation is not the amount of time it will consume. By introducing hourly reallocations, we disregard the randomness patrol method by making predictable movements. In such cases, we may be required to introduce larger assumptions that may completely invalidate our results.

Lastly, applying our algorithm to larger areas such as Pittsburgh is a limitation due to the amount of time it takes, but it is certainly a feasible extension. The area of jurisdiction becomes exponentially larger, but regarding the final solution, due to our algorithm’s flexibility, it can help find the optimal placement when given supporting data.

Conclusion

 Our goal was to find the optimal placement of police officers such that they cover the areas with high crime rates, given crime rates in each location by using maximal coverage algorithm. Other than optimal placement of officers, our algorithm can be used for any other application that requires maximal covering such as the optimal placement of facilities or services given demand in regions. Although our problem took us a long time to compute, it was mainly due to the fact that we had to sort the data by hand. When given sorted data, our algorithm will be able to find correct optimal placement for any areas with demands or density. Hence, our flexible algorithm can provide solutions to a variety of fields.

Acknowledgement

 We would like to thank Lt. Scheimer for providing us data and kindly answering any questions we had regarding this project.