Creating optimal patrol areas using the p-median model*

Andrew P. Wheeler[1,**],

*Data and code to replicate the findings can be obtained from https://www.dropbox.com/s/wdin3q22yfid0q9/Tutorial_PatrolAreas.zip?dl=0

**Corresponding author, email: apwheele@gmail.com

1. School of Economic, Political, and Policy Sciences, Program in Criminology and Criminal Justice,

The University of Texas at Dallas

800 West Campbell Road, Mail Station GR 31

Richardson, Texas 75080-3021

Andrew P. Wheeler is an Assistant Professor of criminology at the University of Texas at Dallas in the School of Economic, Political, and Policy Sciences. His research focuses on the spatial analysis of crime at micro places and practical problems faced by crime analysts. His recent work has been published in the *Journal of Research in Crime & Delinquency*, the *Journal of Quantitative Criminology*, *Crime & Delinquency*, *Homicide Studies,* the *Security Journal*, *Cartography and Geographic Information Sciences*, the *International Journal of Police Science and Management*, and the *Journal of Investigative Psychology and Offender Profiling*.

I would like to thank the Carrollton Police Department and in particular Cdr. Andrew Horn and Chief Derek Miller for the opportunity to collaborate on this project.

Title: Creating Optimal Patrol Areas using the P-median Model

Purpose: To illustrate the use of the p-median model to construct optimal patrol areas. This can improve both time spent traveling to calls, as well as equalize call load between patrol areas.

Design/Methodology: The paper provides an introduction to the use of integer linear programs to create optimal patrol areas, as many analysts and researchers in our field will not be familiar with such models. The analysis then introduces a set of linear constraints to the p-median problem that are applicable to police agencies, such as constraining call loads to be equal and making patrol areas geographically contiguous.

Findings: The analysis illustrates the technique on simplified simulated examples. The analysis then demonstrates the utility of the technique by showing how patrol areas in Carrollton, TX can be made both more efficient and equalize the call loads given the same number of patrol beats as currently in place.

Originality/Value: Unlike prior applications of creating patrol areas, this article introduces linear constraints into the p-median problem, making it much easier to solve than programs that have non-linear or multiple objective functions. Supplementary code using open source software is also provided, allowing other analysts or researchers to apply the model to their own data.

Keywords: Patrol Areas, Integer Linear Programming, p-median, Policing, Hot Spots

Supplementary materials illustrating how to conduct the analysis can be downloaded from https://www.dropbox.com/s/k3z1ww9ki5gyg6r/Tutorial_PatrolAreas_Anony.zip?dl=0

Introduction

Despite the fact that a quantitative model of selecting patrol areas was presented by Phillip S. Mitchell over 45 years ago (Mitchell, 1972), it is still the case that many police departments create patrol areas in an ad-hoc fashion (Larson, 1974b; Zhang & Brown, 2013). A likely explanation is that such models are not presented in a way that are accessible to police planners or crime analysts. The allocation model in Mitchell (1972), as well as more recent extensions (Bucarey, Ordóñez, & Bassaletti, 2015; Curtin, Qiu, Hayslett-McCall, & Bray, 2005; Camacho-Collados, Liberatore, & Angulo, 2015; Leigh, Dunnett, & Jackson, 2017; Zhang & Brown, 2012), are typically formulated using integer linear programming models. Linear programming models are within the domain of operations research (Maltz, 1996), but are not typically taught to students in the social sciences (Kringen, Sedelmaier, & Elink-Schuurman-Laura, 2017).

The motivation for creating optimal patrol areas is simple – an optimal spatial arrangement for patrol areas can improve efficiency in distance travelled to handle calls for service. This result has direct benefits in improved response time to calls, as well as savings in fuel and vehicle maintenance. There are several additional advantages for creating optimal patrol areas as well. Less time spent driving to calls means an officer has more discretionary time, which can be used for proactive policing (Wilson & Boland, 1978). Optimally sized patrol areas can naturally encourage hot spots policing (Braga, Papachristos, & Hureau, 2012). Areas with more calls for service will produce smaller patrol areas, so hot spots of crime have a larger dosage of officers compared to areas with fewer calls (Mitchell, 1972). Finally, an officer spending more time in their assigned area, as opposed to regularly travelling to different areas to answer calls, will become more familiar with the people and places in that area (Wilson, 1957). That greater familiarity can better facilitate community oriented policing or problem oriented policing approaches (Braga et al., 1999).

This essay provides a gentle introduction to spatial allocation models through simplified examples. First, it illustrates how to formulate the canonical linear programming problem by creating an N by N table (Hillier & Lieberman, 2005). Then it illustrates the model using simplified grids, as well as potential limitations and extensions of that model.

Unlike prior work that includes bi-objective functions or non-linear constraints to create patrol areas (Bucarey, Ordóñez, and Bassaletti , 2015; Camacho-Collados, Liberatore, & Angulo, 2015), the linear program demonstrated here only includes one objective function and linear constraints. This makes the problem much easier to solve in practice, and can be estimated in open source software using traditional integer linear programming. The supplementary materials provide a walk through and computer code to apply the model (Mitchell, O'Sullivan, and Dunning, 2011), so future analysts and researchers can easily apply the model to their own data.

The technique is illustrated with a set of real data using a case study in Carrollton, TX. The Carrollton Police Department was interested in redefining patrol areas, due to currently unequal call loads in their current set of patrol areas. The analysis demonstrates Carrollton can both make patrol areas much more efficient in terms of time travelled to calls, while still creating patrol areas that are much more equal in call load.

A Review of Prior Work on Police Allocation Models

Prior literature on allocating police resources has focused on either temporal dimensions or spatial dimensions. Temporal dimensions of allocating staffing resources frequently identify the fact that service needs vary throughout the day and the season of the year (Chaiken & Larson, 1972; Taylor & Huxley, 1989; Srinivasan et al., 2013). One often used rule is that an officer's workload should only be, on average, 60% (McCabe, 2013). That is, an officer should spend around 60% of their time on answering calls for service and initiating proactive incidents, such as conducting traffic stops (Sherman & Rogan, 1995) or pedestrian stops (Boydstun, 1975).

While McCabe (2013) suggests that workloads of under an average of 60% are suboptimal, this has important implications for the spatial allocation of resources. When officers are busier, it will result in more instances of officers from a different patrol area needing to take calls that are not within their assigned area (Larson, 1970). With very busy officers, it could be the case that optimal patrol areas are meaningless, as officers are simply going from call to call with no discretionary time spent within their assigned area. Many smaller police departments do not have patrol areas, and indeed it may not make sense for them to have them. Larger police departments do though regularly have assigned patrol beats and other hierarchically organized spatial areas (Curtin, Qiu, Hayslett-McCall, & Bray, 2005).

For those police departments that use a spatial patrol allocation strategy, they often have the following characteristics:

- Individual officer(s) are assigned to one particular area
- An officer is to conduct routine patrols in their assigned area, as well as answer calls for service that occur in that area (Kelling, Pate, Dieckman, & Brown, 1974)
- In the case a call is received for an area in which all assigned officers are busy, an officer from a nearby area is assigned to take the call (Larson, 1975). After they are finished with the call, they are to return to their initial assigned area.

There are variants to the above model – a common one is that each area is assigned only one officer, and each officer is expected to have near equal call load. There may be additional idiosyncrasies in any police department as well, such as officers with special assignments that are not assigned to one area, but may take calls in several different areas if needed.

There are then two different ways in which prior research has evaluated making the process of answering calls for service more efficient. One way is using a queueing model to assign calls (Larson, 1974a). In this model the patrol areas are fixed, but then one has a specific set of rules for *when* officers are assigned calls. While many police departments simply assign a call to the nearest available officer when it first arrives, this is not necessarily the most efficient way to process calls. Many police departments have different priority levels for different calls, for example a robbery in progress would likely have a higher priority than a theft from a motor vehicle. Some high priority calls demand immediate response from multiple officers, and it can be the case that even if officers are currently responding to one call they need to leave without resolving the incident (a.k.a. preempted) for the current higher priority incident.

Because of the different priority levels of calls, it often makes sense to place low priority calls in a queue, and only assign it when the officer assigned to that area becomes available. That is, instead of asking an officer to go outside of their assigned area to handle the call, the dispatcher will wait for the officer within the callers patrol area to become available. In such cases it is often the responsibility of the dispatcher to tell the caller an approximate time when an officer will be available to respond to the incident.

Jan Chaiken and Richard Larson worked extensively on different computer systems used to better allocate patrol resources and simulate different scenarios (Chaiken & Dormont, 1978; Chaiken & Larson, 1972; Larson, 1973). These solutions, however, assumed that particular patrol areas were fixed, and that one would simply shift around resources to better accommodate calls occurring in those pre-set areas (D'Amico et al., 2002). Thus they do not help with the task of creating patrol areas to begin with. Hence a second set of research has focused on different ways to optimally create patrol areas given the spatial distribution of calls for service.

Patrol areas are often created based on constructing natural areas in a city, frequently with a goal of making call loads for officers as equal as possible (Wilson, 1957; Larson, 1974). It is possible that such areas created in an ad-hoc manner tend to be near optimal (Curtin, Qiu, Hayslett-McCall, & Bray, 2005; D'Amico et al., 2002), and there would be no (or only slight) additional benefit to redesigning new areas for a police department. Because crime within hot spots and neighborhoods tends to be quite temporally stable (Sampson, 2012; Weisburd et al., 2004), even if such areas were made many years ago it is possible they still currently provide near optimal spatial arrangement of patrol areas. Simply having naturally defined areas could be important for having officers effectively stay within their patrol area (Larson, 1974; Sorg, Wood, Groff, & Ratcliffe, 2014), or even be important for implementing crime prevention initiatives (Caulkins, Larson, & Rich, 1993).

But, some cities are faced with changing dynamics that require them to reevaluate their current patrol areas. Annexation of neighboring areas to a city, extensive commercial and residential development (or decline), and decreasing (or increasing) crime at particular places may force a police department to reconsider patrol areas. Additional development of major thoroughfares may divide particular neighborhoods (Jacobs, 1993), but may make travel in a particular predominate direction faster (Larson, 1970). In any of these cases models of how to spatially distribute patrol areas are particularly relevant.

As previously mentioned, Mitchell (1972) presented a model for creating patrol areas. All one needs to implement this model is the number of calls one expects in particular subareas that are smaller than the expected patrol areas, as well as a measure of the distance between those subareas. Although Mitchell (1972) suggested to use a particular algorithm to solve the problem, it is an example of a more general integer linear program, with which there are currently a wide array of tools one can use to solve the problem (Hillier & Lieberman, 2005). In particular, this problem is often referred to as the *p-median* problem (Mladenović et al., 2007).

There have of course been additional refinements of the patrol allocation models suggested since Mitchell (1972). Curtin et al. (2010) illustrates a maximal coverage model that includes the specific location of particular incidents, incorporates minimal travel distances in the coverage, and includes the potential to have incidents that need backup, or multiple responses. Zhang & Brown (2013; 2014) discuss different simulation approaches to find optimal areas, which can incorporate more realistic behavior of agents in the system and can incorporate a wider set of characteristics to optimize than Mitchells (1972) model. Bucarey, Ordóñez, and Bassaletti (2015) include constraints to make the resulting regions compact and balance the workload.

Several different extensions incorporate the idea of hot spots into either creating optimal patrol areas (Camacho-Collados, Liberatore, & Angulo, 2015), or in dispatching units to recent hot spot crime areas (Leigh, Dunnett, & Jackson; 2017). Both of these models subsequently produce dynamic solutions – ones that can change given different seasons or projections for future crimes.

Despite these advances though, the original model presented by Mitchell (1972) is still applicable and provides a simple introduction to creating optimal patrol areas. It is also readily implementable in current open source software, so it can be practically applied by any interested researcher or crime analyst.

Illustrations of Choosing Patrol Areas

As opposed to starting with a mathematical formulation, the next section will walk the reader through an example of applying the model to a very simple set of data. Figure 1 displays an example of a set of micro *atoms* that we wish to partition into patrol areas. In practice these atoms can represent any micro level of geography that the researcher is interested in. They can be a set of small grid cells (Mitchell, 1972), or they could represent different geographies, such as street segments or census blocks. The atoms are labelled individually using the letters A, B, C and D.

[INSERT FIGURE 1 HERE]

Above each atom is a number. In this example these numbers will represent the total number of calls that each atom is expected to receive. Although here these are integer numbers, in practice one would likely want to weight calls according to different factors. For example, one might want to give higher priority calls and/or calls that take longer to service higher weight (e.g. a call that takes on average 30 minutes to service gets twice the weight compared to a call that takes 15 minutes to service). Fractional values for these call weights can be used for this model.

Finally, each atom has a measured distance to other atoms. In this example, each adjacent atom is one distance unit apart. So the distance between A and B is one, the distance between A and C is two, etc. For actual applications it is preferable to use either the road distance between the two atoms or the travel time distance. Note that the travel time for actual roads may be different depending on the source and destination, such as travelling from A to B may take less time than travelling from B to A (travelling opposite a one way street would be an example). The way the model is formulated such asymmetry is fine. Additionally the distance within one atom does not need to be set equal to zero. One could use the distance or travel time it takes to cross from one end to the other of the atom (Bucarey, Ordóñez, and Bassaletti , 2015).

This example will partition the units into two patrol areas. With such a simple example, an obvious partition would be one patrol area will encompass the A and B atoms, and the other patrol area would encompass the C and D atoms. The following section will quantify that intuitive partition into a mathematical model that can be extended to more complicated situations.

A first step in many linear programs is to formulate the problem into a contingency table, where the columns are the end choices that need to be decided amongst, and the rows are different options (Hillier & Lieberman, 2005). In this particular problem, the contingency table is created by turning the original four locations into a distance matrix. So in this four by four distance matrix, the distance between source row A and destination column A is zero, the distance between source row B and destination A is one, etc. Then one multiplies that distance by the total number of calls in the *destination* atom. This final value is then the *cost* if a patrol car was located in source atom *s*, but had to travel to destination atom *d* to answer the calls. So for example, the cost between source C and destination A is 2 (the distance between the two atoms) multiplied by 4 (the number of calls in atom A), for a total cost of 8. This then forms the cost matrix for the costs between all pairs of atoms, which is displayed in Table 1.

[INSERT TABLE 1 HERE]

The final part of the process is to describe a set of *decision variables,* labelled X_{sd}, that identify whether source atom *s* will be assigned destination area d. As a note, most regression models are written where the *X* or *Y* variables are known measures, and one wishes to obtain estimates for different parameters, often denoted using Greek letters (e.g. for the linear regression equation $y = \alpha + \beta x + \epsilon$, one knows the values of *x* and *y*, and one wants to estimate alpha, beta and epsilon). The notation is the opposite though for linear programs, *X* and *Y* are often the values that need to be estimated via a model. This problem is a binary integer linear program, so the X_{sd} decision variables can only take the values of either 0 or 1. The linear model in this case can then be written as:

$$
\begin{aligned}
\text{Total Weighted Travel} = \quad & 0X_{AA} + 1X_{AB} + 4X_{AC} + 9X_{AD} \\
& + 4X_{BA} + 0X_{BB} + 2X_{BC} + 6X_{BD} \\
& + 8X_{CA} + 1X_{CB} + 0X_{CC} + 3X_{CD} \\
& + 12X_{DA} + 2X_{DB} + 2X_{DC} + 0X_{DD}
\end{aligned}
$$

The X_{sd} are decision variables that can either be zero or one in the final estimate, and one wants to *minimize* the total weighted travel from source *s* to destination *d*. To turn this equation into a set of mutually exclusive patrol areas though, the solution needs to adhere to several constraints. First, one needs to set an X in each column (as written) to a value of 1 one time and only one time. This means that each destination is *covered* by only one source location. So for example, among the variables $X_{AA}, X_{BA}, X_{CA}, X_{DA}$, one of those has to be set to be equal to 1, and the others are set to be equal to 0.

Second, to make a set of at most only two patrol areas, one can only select X variables that are within two rows (since we are constructing only two areas). So if one selects X_{AA} and X_{DD}, then none of the X's within source row B or source row C can be equal to 1.

Given these two constraints, the set of variables that minimizes the total weighted travel is to set $X_{AA}, X_{AB}, X_{DC}, X_{DD}$ to 1, and all other X's to 0. This means that source atom A covers atoms A and B. Subsequently those two atoms make up one patrol area. The other patrol area includes atoms C and D, which are covered by source atom D. This is mostly due to the fact that atoms A and D have the highest number of calls in this example, so assigning them as the source minimizes their weight.

This simplified example shows how the model is limited in application to assigning patrol areas. First, the total weighted travel time calculation presumes that a patrol car will be dispatched from a centralized location. This makes sense for some emergency services, like fire or ambulances, but does not for police cars, which have a random patrol function. Also, such a model does not consider the fact that cars will have to cross patrol areas to answer calls on occasion, as well as the fact that some calls may need multiple units to respond. These however are limitations shared with many of the more recent model extensions as well (Curtin et al., 2010). Typically, one will want to conduct simulations after the patrol areas are created to determine if this simplification will cause unexpected issues (Larson, 1973; Zhang & Brown, 2013; 2014).

An additional limitation is that such a solution is only cross-sectional. It could be the case that certain areas have more calls during different shifts or days of the week, making the average number of calls not representative of the actual call load in any particular shift. Again, one will want to conduct additional analysis as to whether the patrol areas result in large differences in call load across different shifts. While some of the models discussed are intended to be more dynamic (Camacho-Collados, Liberatore, & Angulo, 2015; Leigh, Dunnett, & Jackson; 2017), in many circumstances patrol areas are going to need to be fixed. It is often the case that one takes the solution provided by the p-median algorithm and then adjusts the areas to accommodate several of these factors not included in the model.

For those interested in how to formally write this model, the appendix contains the mathematical notation to describe and implement the p-median model with a set of relevant constraints, in particular spatial contiguity as well as workload minimum and maximum constraints.

The next section will present some additional examples applying this patrol area algorithm to simplified examples. These will be on a slightly larger set, given a set of small atoms on a 6 by 6 grid, but will still be assigned to only two patrol areas. They are presented to show how the model behaves given different potential inputs. In each grid, the number of calls is displayed as a circle with varying sizes – a larger circle indicates more calls in that particular atom. The distance between each atom will be measured by Manhattan distances (not straight-line Euclidean distances), and each grid cell will be 1 unit apart in each X and Y direction. So this means that grid cells that are touching corners are 2 units apart. Contiguity between the units is defined as sharing an edge (e.g. Rook contiguity). In all cases the resulting areas were constrained to have near equal call loads (within 5% of the average call load).

Figure 2 displays four example grids with the solutions for two patrol areas. To read the graph, the variable sized points display the total number of calls within an atom grid cell. The colors represent what patrol areas were assigned to the same source, and the source atom is symbolized with a black outline.

[INSERT FIGURE 2 HERE]

In the case for even calls (top left), it split the patrol areas evenly. While in this example the algorithm split the patrol areas north-south, splitting them east-west would be an equally valid solution. In the top right, it partitions the one hot spot into a slightly smaller area, whereas it makes the grey partition cover a larger swath. This is an example of how creating optimal patrol areas can help facilitate hot spots policing. The smaller patrol area with more calls will have a higher police presence per area, compared to the larger patrol area that is more spread out, but does not have any high concentration of crime. Additionally the source area is more likely to be located at or nearby the hot spot. Thus the algorithm provides a suggested area that an officer should spend more of their discretionary time that simultaneously reduces travel times to calls and encourages hot spots policing.

Another aspect of the p-median solution is that the assigned beat in the upper right has an elongated pattern. One might assume that one patrol area would encompass a more square like pattern, starting from the hot spot and emanate outwards, and the other patrol area would encompass the remainder. The p-median solution does not consider the shape of the resulting patrol area though, only the distances between the source atom and the destination atom. Many definitions of how patrol areas should look typically want them to be a polygon that is approximately square (Larson, 1974b). The patrol area being a nicely shaped polygon (i.e. convex) is not guaranteed though by this algorithm. In some instances this could be reasonable, such as a long and skinny patrol area running along one major thoroughfare. In other cases it may produce irregular borders though that would divide up more natural areas.

In some cases there may be natural borders the police department does not want patrol areas to cross, such as a river or a highway dividing a city. In these cases there are two simple ways to amend the algorithm to obey those boundaries. One is if the boundaries entirely split the jurisdiction into subsections, you can simply run the algorithm on each subsection. Another that allows the model to run on the whole sample at once is that if a particular path between a source and destination atom crosses the barrier, you can set the distance between the source and destination in that circumstance to some arbitrary but very large value. In that case the cost to assign those areas to the same source will be too large, and thus they will always be divided in the solution.

In Figure 2 in the lower left, the hot spot in the middle ends up being split between the two patrol areas. This is because if the hot spot were entirely within one area it would cause disparity in the call loads between the two patrol areas. Thus the hot spot ends up being split up among the two patrol areas, but each source atom is located closer to the highest crime hot spot. Again this facilitates hot spot policing, as the patrol area that contains the largest center of the hot spot has a smaller amount of total area to cover.

The final example in the lower right shows how with two hot spots the two patrol areas are again split evenly between the two areas. In this case though the areas are not perfect rectangles, but are weighted closer to the one hot spot area in the corner. Again there are additional ways that such a configuration could be split to return slightly different patrol areas with the same minimum objective function, but the results are intuitive given the location of the hot spots and which source atom is chosen for each patrol area.

Case Study of Creating Optimal Patrol Areas in Carrollton, Texas

Carrollton is in northeast Texas and is a suburb of Dallas. The city's population is currently over 120,000 in an area over 37 square miles, but like many cities in the southern United States is experiencing population growth and regularly annexes neighboring areas. Figure 3 displays the current spatial configuration of the Carrollton Police Departments (CPD) geographic areas, which encompasses twelve beats, nested within three divisions (North, Central and South).

[INSERT FIGURE 3 AND FIGURE 4 HERE]

CPD was interested in reorganizing their beats, as recent changes within the city had produced unequal call loads among the beats. CPD also wished to incorporate future planned development of single family and multi-family housing, which they expected would also potentially change call loads in specific areas of the city. The main constraint of interest was to balance the resulting patrol areas, as each patrol area is intended to be assigned one officer. There are a total of 325 reporting areas in Carrollton (with a mean area of 0.1 miles^2), and these will be the micro atoms with which the resulting optimized patrol areas will be clustered into. Figure 4 displays those reporting areas, as well as the spatial variation of calls for service from

3/12/16 through 8/29/17 – a total of over 130,000 calls for service.[1] Calls were only obtained back until March 2016 as Carrollton PD had recently changed their RMS system and call classification scheme. Weighting calls according to the time they took to service (from the time the first officer was on-scene to when the incident was cleared) produced very similar estimates compared to just using the counts of calls aggregated to the reporting area level – a correlation of over 0.95. So for simplicity the counts of calls for service in a reporting area are used as the call weight in creating optimal patrol areas.

Network distances in terms of *minutes* of travel between each reporting area centroid were calculated using a street centerline file with speed limits (provided by the City of Carrollton). Rooks contiguity was subsequently used to identify whether two reporting areas were adjacent to one another.

The last arbitrary parameter to determine would be the amount of inequality that is currently acceptable in workloads. Given the historical data, the minimum proportion of calls within a beat was 6%, and the maximum was 10%. Given twelve beats, a perfectly equal call load would be just over 8% of all calls. Subsequently the current amount of disparity in call loads across beats is over 20%. Any solution should at a minimum not produce worse disparity, and should attempt to improve upon this disparity. To accommodate this, several solutions reducing the inequality were run, from 20%, 10%, and 5%. The 5% solution returned a set of reasonable areas, and still dramatically improved the total weighted travel compared to the current set of patrol areas, so this solution was used for illustration.[2]

---

[1] Calls with missing date-times for when the officer was on-scene or cleared the incident, as well as calls assigned to the police headquarters were eliminated from this analysis.

[2] Interactive maps of the solutions with different constraints on the workload equality, as well as the original patrol areas, is provided at https://dl.dropbox.com/s/pasiulzd86nn814/Carrollton_DiffSolutions.html. They tended to produce very similar patrol areas, so the analysis only focuses on the one solution with the strictest 5% workload inequality constraint.

[INSERT FIGURES 5 AND 6 HERE]

Figure 5 illustrates the workload equity constraint. The dots in the figure each represent a patrol area; on the left it shows the current workload as an average number of calls per day (Y-axis) for each of the twelve patrol areas. The dots on the right illustrate how the current proposed solution constrains the new areas to have very near equal workload, with a range from 19.8 to 21.8 calls per day. This is compared to the current workload inequality across patrol areas, with a range of 15.1 to 24.3.

Figure 6 maps the results of the optimal patrol areas. Under the current set of beats, the minimum estimated current weighted travel is 248,680, whereas the solution setting the amount of disparity to 5% is 198,758. This is a reduction in the total weighted travel of over 20%. This potential travel time reduction is not guaranteed, as the model does not take into account random patrol, cross-area dispatches, or the fact that an officer may not be dispatched from the source area. But making the areas much more equal in workload is the simplest way to prevent more cross-area dispatches, and thus it is possible the new suggested patrol areas result in *even larger* travel time reductions than 20%.

While these maps are illustrative of the p-median method, they are often just a useful starting point, not the end of the analysis itself. It is likely the case that CPD will adjust the boundaries based on various aspects not considered by the model. In particular the resulting models tend to produce small peninsulas that the department may wish to smooth out (Bucarey, Ordóñez, and Bassaletti, 2015). In addition to this, the CPD may wish to make the boundaries adhere to particular major roadways. For example, the long dark blue patrol area in the southwestern part of the city may appear awkward, but follows a major highway running north-south in that area of the city. Also the reporting area at the southernmost tip of the city that is not included in the dark blue patrol area appears strange at first glance as well. This reporting area happens to be a residential area that *does not* have direct access to that same highway. So although it creates an awkward looking boundary, given the local street network those two delineations appear to be quite reasonable.

The author is continuing to work with Carrollton PD to incorporate forecasts of future call load into creating such areas, as well as analyzing whether the proposed changes may have unintended consequences for allocating officer resources that cannot be incorporated into the model. This will likely involve ad-hoc changes to the optimal solution. While this would seem to be disparaging to using automated methods at all, the optimal solutions provided by the p-median model are likely a much better starting point than constructing the patrol areas entirely from scratch in an ad-hoc fashion.

Conclusion

Optimally creating patrol areas is a task that is quite regular for police departments, especially in growing cities. The creation of such optimal patrol areas not only can reduce response times and equalize call load between areas, but can also naturally encourage hot spots policing. Any arrangement that reduces an officers time spent travelling to calls will also allow more time for an officer to become more familiar with their assigned area, which can also help encourage problem and community oriented police tactics.

The main aim of this article was to introduce a model of how to create patrol areas in a relatively simple way. A particular hurdle, even for more advanced quantitative researchers, is understanding the typical formulation of linear programming models, which have several distinctions from regression models that are more standard fare for social scientists. The supplementary materials include computer code implementing such models using the open source python programming language.

Using such a model, the analysis demonstrates that the Carrollton Police Department can create much more efficient patrol areas in terms of the expected time travelling to calls. It simultaneously created patrol areas much more equal in call load. Police departments in growing or shrinking cities will often need to redraw such boundaries, so having programs to help ease that task is important. Even if a department wants to manipulate the resulting areas to account for factors not included in the model, the results from the p-median model will provide a much better starting point that manually constructing areas in an ad-hoc fashion.

One should be aware of the fundamental limitations of the p-median model though. Its main limitations are that is assumes travel will be coming from one particular source atom, which is unrealistic and does not take into account how patrol officers will randomly move about (and beyond) their assigned patrol area (Sorg, Wood, Groff, & Ratcliffe, 2017). Another limitation is that the patrol areas created can have irregular borders. While other models take into account the convexity of the resulting patrol areas (Camacho-Collados, Liberatore, & Angulo, 2015), it will often be the case that one needs to conduct manual adjustments to the final p-median solution to account for factors that cannot be encoded in the model. A final limitation is that this model does not consider aspects such as providing back up to officers (Curtin et al., 2010).

Given these limitations though, it does quite well in equalizing call loads, an important consideration for most police departments when creating patrol areas. The novel set of constraints introduced in this article to set the minimum and maximum call load to specific values make the p-median model much more applicable to the task of creating patrol areas.

When call loads are unequal, officers will more often be dispatched to calls within other patrol areas, which can cascade into more inefficient call times throughout the system. To create more efficient patrol areas equalizing call loads should be given high weight in any particular model.

While this article focuses on creating the optimal patrol areas and does not currently discuss queueing models in any detail, they often go hand in hand. That is, one formulates areas and then evaluates different potential queueing models for allocating resources, as the areas are then fixed. Often such different solutions are evaluated using simulation of individual agents travelling to calls (Larson, 1973; Zhang & Brown, 2012; 2013). An important future research endeavor is simply re-creating a system similar to that described in Larson (1973) that can simulate different outcomes for police departments, but take advantages of different modern software languages.[3] Such open source software could be developed with grant resources from the National Institute of Justice, similar to how the CrimeStat (Levine, 2006) and GeoDa (Anselin, Syabri, & Kho, 2006) programs are currently supported. Having easy to use software applications would make applying such models much easier for both practitioners and researchers.

While the creation of optimal patrol areas can be motivated simply be equalizing workloads and reducing travel time to calls, they can also facilitate a hot spots policing strategy. One way optimal patrol areas facilitate hot spots policing is that the created patrol areas should be more compact, and in particular patrol areas with a high density of crime should be smaller. It is possible for the algorithm to assign one small hot spot its own patrol area if the call load in that hot spot is sufficiently high to warrant it.

---

[3] Current versions of the PAM software algorithms developed by Jan Chaiken are available from the National Highway Traffic Administration at https://one.nhtsa.gov/Driving-Safety/Enforcement-&-Justice-Services/Personnel-Allocation-Model-(PAM)-for-Law-Enforcement-Agencies and are programmed using an Access database.

A second way creating optimal patrol areas facilitate hot spots policing is through the creation of equal call-loads; officers assigned to patrol areas that contain hot spots of crime do not have an additional burden of extra calls compared to officers covering areas that crime is less prevalent. Thus they should have an equal amount of time to pursue proactive activities as do officers assigned to not high crime locations. Additionally this should prevent being dispatched to other areas more frequently, allowing officers to spend more time and thus become more familiar with their assigned patrol area. Such familiarity is a necessary component of either community or problem oriented policing.

An additional way that creating optimal patrol areas can facilitate hot spots policing is that when officers are constrained to small hot spot areas they tend to roam from their assigned locations (Sorg, Wood, Groff, & Ratcliffe, 2014; Sorg, Wood, Groff, & Ratcliffe, 2017). The creation of optimal patrol areas provides a larger box that does not necessarily constrain officers to a very small area, but provides a suggested source area that officers should return to frequently to minimize total travel. That source area will often be in or nearby a hot spot of crime (see the black boxes in Figure 2). Thus they are a potential way of incorporating hot spots policing into the current way patrol resources are distributed that do not require creating additional resources, such as a specialized unit assigned to predicted areas of high crime. Officers can simply be nudged into more frequently visiting those long term high crime areas within their assigned patrol areas, without totally restricting their discretion to patrol throughout their assigned area.

Appendix: The Formal Model

This technical appendix is provided to detail the formal integer linear program used in this research. Table A.1 introduces the variables that will be used to write the linear programming model more formally. The more general way our integer linear model can be formulated is:

$$\text{Total Weighted Travel } = \text{minimize}\left[\sum_s \sum_d (C_d \cdot D_{sd}) \cdot X_{sd}\right]$$

Subject to the constraints:

1. $X_{sd} \in (0,1)$
2. $\sum_s X_{sd} = 1,\ \forall\, d$
3. $\sum_s X_{ss} = P$
4. $X_{sd} \leq X_{ss}, \forall\, s, d$
5. $\sum_a W_{ad} \cdot X_{sa} \geq X_{sd},\ \forall\, s, d$ subject to $s \neq d$
6. $\sum_d C_d \cdot X_{sd} \leq I_{\max},\ \forall$ s
7. $\sum_d C_d \cdot X_{sd} \geq I_{\min} \cdot X_{ss},\ \forall$ s

The next section will walk through each step in turn and explain the mathematical formulation.

[INSERT Table A.1 HERE]

First, one wants to minimize the statement, $\sum_s \sum_d (C_d \cdot d_{sd}) \cdot X_{sd}$. The s and the d subscripts again refer to either a *source* atom or a *destination* atom in our application. This means that C_d equals the number of calls in a source destination, and this is multiplied by D_{sd}, the distance from the source atom. These are then the costs, which were illustrated in Table 1. The number of calls and the distance between a source and a destination are known quantities, but we need to estimate what the values of the X_{sd} variables will be to minimize the equation. This is a more concise way to write the model, instead of writing out the entire distance matrix.[4]

[4] This will subsequently result in n^2 decision variables if one has an original n atom locations. A simple way to reduce the size of the problem is to set a threshold on the distance, t, for whether two atoms could be feasibly linked. This creates a moving neighborhood around every source location (D'Amico et al., 2002), and can dramatically reduce the size of the integer linear programming problem. With this restriction, the total number of decision

Which X_{sd} we choose though needs to have some constraints relevant to the problem. With no constraints, the trivial solution is to simply set all X_{sd} variables to zero. This however will not help us identify optimal patrol areas. Setting a certain set of constraints on the system will produce a set of feasible solutions (ones that satisfy the constraints), from among which we can assess which solutions minimize the weighted distance travelled.

The first constraint listed is that either a source-destination pair is chosen or it is not chosen, we do not want our algorithm to return fractional values. This is stipulated in the first constraint, that $X_{sd} \in (0,1)$. This mathematical statement means that our decision variables are an *element* within the set of 0 and 1. Or otherwise it needs to be a binary yes-no decision whether that source variable covers the destination variable. The fact that the decision variables are limited to integers is what makes this an integer programming problem. In the end, destination atoms covered by the same source atom will represent one patrol area. For example, if the solution found the variables X_{AA}, X_{AB}, X_{AC} were each chosen as equal to 1, this would mean atoms A, B, and C are all contained in one patrol area, with the atom A as the source for each.

The second constraint, $\sum_s X_{sd} = 1, \ \forall\, d$, symbolizes that each destination needs to be assigned to only one source. This can be thought of as summing each column in the original table formulation, so for example it is saying that $X_{AA} + X_{BA} + X_{CA} + X_{DA}$ must equal 1. You then have this constraint *for each* (the ∀ symbol) destination *d*, so this symbolizes four separate constraints on our original simplified example.

The third and fourth constraints are each necessary to limit the number of source locations to our pre-determined number of patrol areas, *P*. The third constraint, $\sum_s X_{ss} = P$,

---

variables to estimate will be less than $n \cdot a$, where *a* is the maximum number of atom areas within a circle with a radius of *t*. The size of *t* can be set by noting the approximate maximum distance that a reasonable patrol area may encompass (e.g. a patrol area will never be longer than 5 miles), or limit the distance or time travelled to pre-defined thresholds (Curtin et al., 2010).

means that the total number of source areas assigned to themselves equals the number of patrol areas. Without this constraint, the system would simply assign each destination to its source area, so you would have only the variables $X_{AA}, X_{BB}, X_{CC}, X_{DD}$ as equal to 1 and the rest 0. The third constraint forces the system to choose P locations to assign the destination to itself.

The third constraint by itself is not sufficient to force the solution to only choose P locations. For our simplified example, one could choose $X_{AA}, X_{AB}, X_{BC}, X_{DD}$, and this is a valid solution when only using constraints 1 to 3. But this solution ends up returning source atom B to cover destination atom C as a feasible solution, which does not make sense for our application. The fourth constraint on the system, $X_{sd} \leq X_{ss}, \forall\ s, d$, prevents this from occurring. Intuitively what this constraint does is that in each row, if the variable X_{ss} equals 0, then the other variables in that row cannot equal 1. So with this constraint, for example since X_{BB} was not selected, it means that X_{BC} is not within the feasible solution. You then make a set of constraints like this for every pairwise combination *s* and *d*. So with our original example, this results in an additional 16 constraints.[5]

The fifth constraint is not typically included for the p-median problem, but enforces that the solution return only contiguous areas. This introduces a new term, W_{ad}, which equals one if atom *a* and atom *d* are contiguous, and zero otherwise. It also equals zero for W_{ii}, that is an atom is not contiguous with itself. The sum then iterates over all atoms in the sample, *a*, for each source atom *s* and destination atom *d*, except when the source and the destination are the same location. What this term accomplishes is that if the decision variable assigns X_{AC} a value of one, that means that its adjacent areas, $X_{AB} + X_{AD}$, has to be greater than or equal to one. You cannot

[5] There is an additional way that previous literature has formulated this constraint, but this involves introducing a set of additional decision variables. So if one has additional binary variables Y_s, then the third constraint is replaced with $\sum_s Y_s = P$, and the fourth constraint is replaced with $X_{sd} - Y_s \leq 0, \forall\ s, d$. Here $X_{ss} = Y_s$, so these formulations are equivalent.

have in the end an assigned area that is not contiguous, such as X_{AA} and X_{AC} both equal to 1 if X_{AB} is equal to 0.[6]

Finally, the 6th and 7th constraints are also not needed for the typical p-median problem, but it is a common one when police departments wish to make patrol areas. This constraint says that for each source area, it cannot be assigned any more calls than I_{max}, or any fewer calls than I_{min}. What this accomplishes is that each patrol area has to have near equal call load, with the minimum and maximum set by the analyst. To describe how this works, think back to the original simplified four areas we categorized. The one assigned patrol area including atoms A and B had a total of 5 calls, and the other assigned patrol area including atoms C and D also had 5 calls. In that case, the call load was equal. Now imagine changing the number of calls in atom A to 6. With our original four constraints, it will still return the same set of patrol areas, but in that case the call load for the set of areas A and B would be 8, and the set C and D are still 5. It may be stipulated at the onset that call loads need to be approximately equal, especially in the cases that one officer is assigned to each area. Given a stipulation that call loads need to be under seven and above five calls, that would produce a set of patrol areas where atom A was one area, and atoms B, C and D were another. This would result in an equal call load for each patrol area.

References

Anselin, L., Syabri, I., & Kho, Y. (2006). GeoDa: An introduction to spatial data analysis. *Geographical Analysis*, 38, 5-22.

Boydstun, J.E. (1975). *San Diego Field Interrogation – Final Report*. Washington, D.C.: Police Foundation.

Braga, A.A., Papachristos, A.V., & Hureau, D.M. (2012). The effects of hot spots policing on crime: An updated systematic review and meta-analysis. *Justice Quarterly*, 31, 633-663.

[6] To further reduce the chance of creating islands, one can only count particular contiguous atoms closer to the source atom (Caro et al., 2004). Closer can be in terms of geographical distance, or in terms of a minimum path between the source and destination. The supplementary materials show how to construct this constraint.

Braga, A.A., Weisburd, D.L., Waring, E.J., Mazerolle, L.G., Spelman, W., & Gajewski, F. (1999). Problem-oriented policing in violent crime places: A randomized controlled experiment. *Criminology*, 37, 541-580.

Bucarey, V., Ordóñez, F., & Bassaletti, E. (2015). Shape and balance in police districting. In. H.A. Eiselt & V. Marianov (Ed.), *Applications of Location Analysis* (pp. 329-347). Cham, Switzerland: Springer.

Camacho-Collados, M., Liberatore, F., & Angulo, J.M. (2015). A multi-criteria police districting problem for the efficient and effective design of patrol sector. *European Journal of Operational Research*, 246, 674-684.

Caro, F., Shirabe, T., Guignard, M., and Weintraub, A. (2004). School redistricting: Embedding GIS tools with integer programming. *Journal of the Operational Research Society*, 55, 836-849.

Caulkins, J.P., Larson, R.C., & Rich, T.F. (1993). Geography's impact on the success of focused local drug enforcement operations. *Socio-Economic Planning Sciences*, 27, 119-130.

Chaiken, J.M., & Dormont, P. (1978). A patrol car allocation model: Background. *Management Science*, 24, 1280-1290.

Chaiken, J.M. & Larson, R.C. (1972). Methods for allocating urban emergency units: A survey. *Management Science*, 19, 110-130.

Curtin, K.M., Qui, F., Hayslett-McCall, K., & Bray, T.M. (2005). Integrating GIS and maximal covering models to determine optimal police patrol areas. In F. Wang (Ed.), *Geographic Information Systems and Crime Analysis* (pp. 214-235). Hershey, PA: Idea Group Publishing.

Curtin, K.M., Hayslett-McCall, K, & Qiu, F. (2010). Determining optimal police patrol areas with maximal covering and backup covering location models. *Networks and Spatial Economics*, 10, 125-145.

D'Amico, S.J., Wang, S.J., Batta, R., & Rump, C.M. (2002). A simulated annealing approach to police district design. *Computers & Operations Research*, 29, 667-684.

Hillier, F.S. & Liberman, G.J. (2005). *Introduction to operations research*. Boston, M.A.: McGraw Hill Higher Education.

Jacobs, J. (1993). *The death and life of great American cities.* New York, NY: Modern Library.

Kelling, G.L., Pate, T., Dieckman, D. & Brown, C.E. (1974) *The Kansas City Preventative Patrol Experiment: A Summary Report*. Washington, D.C.: Police Foundation.

Kringen, J.A., Sedelmaier, C.M, & Elink-Schuurman-Laura, K.D. (2017). Assessing the relevance of statistics and crime analysis courses for working crime analysts. *Journal of Criminal Justice Education*, 28, 155-173.

Larson, R.C. (1970). On quantitative approaches to urban police patrol problems. *Journal of Research in Crime and Delinquency*, 7, 157-166.

Larson, R.C. (1973). On-line simulation of urban police patrol dispatching. *Proceedings of the 6th Conference on Winter Simulation*, 371-385.

Larson, R.C. (1974a). A hypercube queuing model for facility location and redistricting in urban emergency services. *Computers & Operations Research*, 1, 67-95.

Larson, R.C. (1974b). Illustrative police sector redesign in district 4 in Boston. *Urban Analysis*, 2, 51-91.

Larson, R.C. (1975). What happened to patrol operations in Kansas City? A review of the Kansas City preventative patrol experiment. *Journal of Criminal Justice*, 3, 267-297.

Leigh, J., Dunnett, S., & Jackson, L. (2017). Predictive police patrolling to target hotspots and cover response demand. *Annals of Operations Research*, Online First, doi:10.1007/s10479-017-2528-x.

Levine, N. (2006). Crime mapping and the CrimeStat program. *Geographical Analysis*, 38, 41-56.

Maltz. M. (1996). From Poisson to present: Applying operations research to problems of crime and justice. *Journal of Quantitative Criminology*, 12, 3-61.

McCabe, J. (2013). An analysis of police department staffing: How many officers do you really need? *ICMA Center for Public Safety Management White Paper*, available at https://icma.org/documents/how-many-officers-do-you-really-need.

Mitchell, P.S. (1972). Optimal selection of police patrol beats. *Journal of Criminal Law, Criminology, & Police Science*, 63, 577-584.

Mitchell, S., O'Sullivan, M., & Dunning, I. (2011). *PuLP: A linear programming toolkit for python*, available at http://www.optimization-online.org/DB_FILE/2011/09/3178.pdf.

Mladenović, N., Brimberg, J., Hansen, P., & Moreno-Pérez, J.A. (2007). The *p*-median problem: A survey of metaheuristic approaches. *European Journal of Operational Research*, 179, 927-939.

Sampson, R.J. (2012). *Great American City: Chicago and the enduring neighborhood effect.* Chicago, IL: University of Chicago Press.

Sherman, L.W., & Rogan, D.P. (1995). Effects of gun seizures on gun violence: "Hot spots" patrol in Kansas City. *Justice Quarterly*, 12, 673-693.

Sorg, E.T., Wood, J.D., Groff, E.R., & Ratcliffe, J.H. (2014). Boundary adherence during place-based policing evaluations: A research note. *Journal of Research in Crime and Delinquency*, 51, 377-393.

Sorg, E.T., Wood, J.D., Groff, E.R., & Ratcliffe, J.H. (2017). Explaining dosage diffusion during hot spot patrols: An application of optimal forager theory to police officer behavior. *Justice Quarterly*, 34, 1044-1068.

Srinivasan, S., Sorrell, T.P., Brooks, J.P., Edwards, D.J., & McDougle, R.D. (2013). Workforce assessment method for an urban police department. *Policing: An International Journal of Police Strategies & Management*, 36, 702-718.

Taylor, P.E., & Huxley, S.J. (1989). A break from tradition for the San Francisco Police: Patrol officer scheduling using an optimization-based decision support system. *Interfaces*, 19, 4-24.

Weisburd, D., Bushway, S.D., Lum, C., & Yang S.M. (2004). Trajectories of crime at places: A longitudinal study of street segments in the city of Seattle. *Criminology*, 42, 283-322.

Wilson, J.Q. & Boland, B. (1978). The effect of the police on crime. *Law and Society Review*, 12, 367-390.

Wilson, O.W. (1957). *Police Planning*. Thomas Publisher.

Zhang, Y. & Brown, D.E. (2013). Police patrol districting method and simulation evaluation using agent-based model & GIS. *Security Informatics*, 2, 1-13.

Zhang, Y. & Brown, D.E. (2014). Simulation optimization of police patrol districting plans using response surfaces. *Simulation*, 90, 687-705.

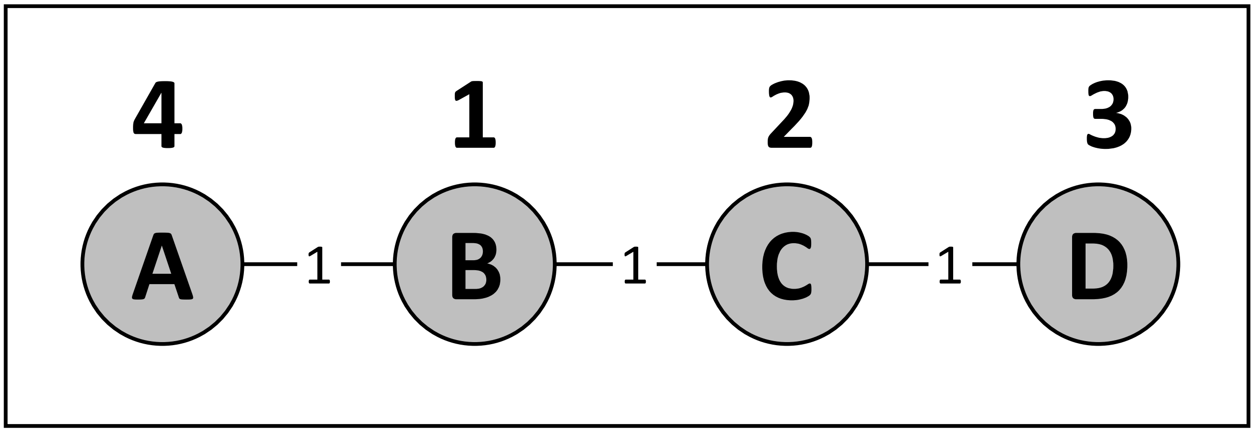

Figure 1: Example set of atoms intended to partition into patrol areas. Above each atom are the total number of calls within that location. So atom D has a total of 3 calls. The distances between each adjacent atom are labelled as 1, so the distance between A and B is 1, the distance between A and C is 2, and the distance between A and D is 3.

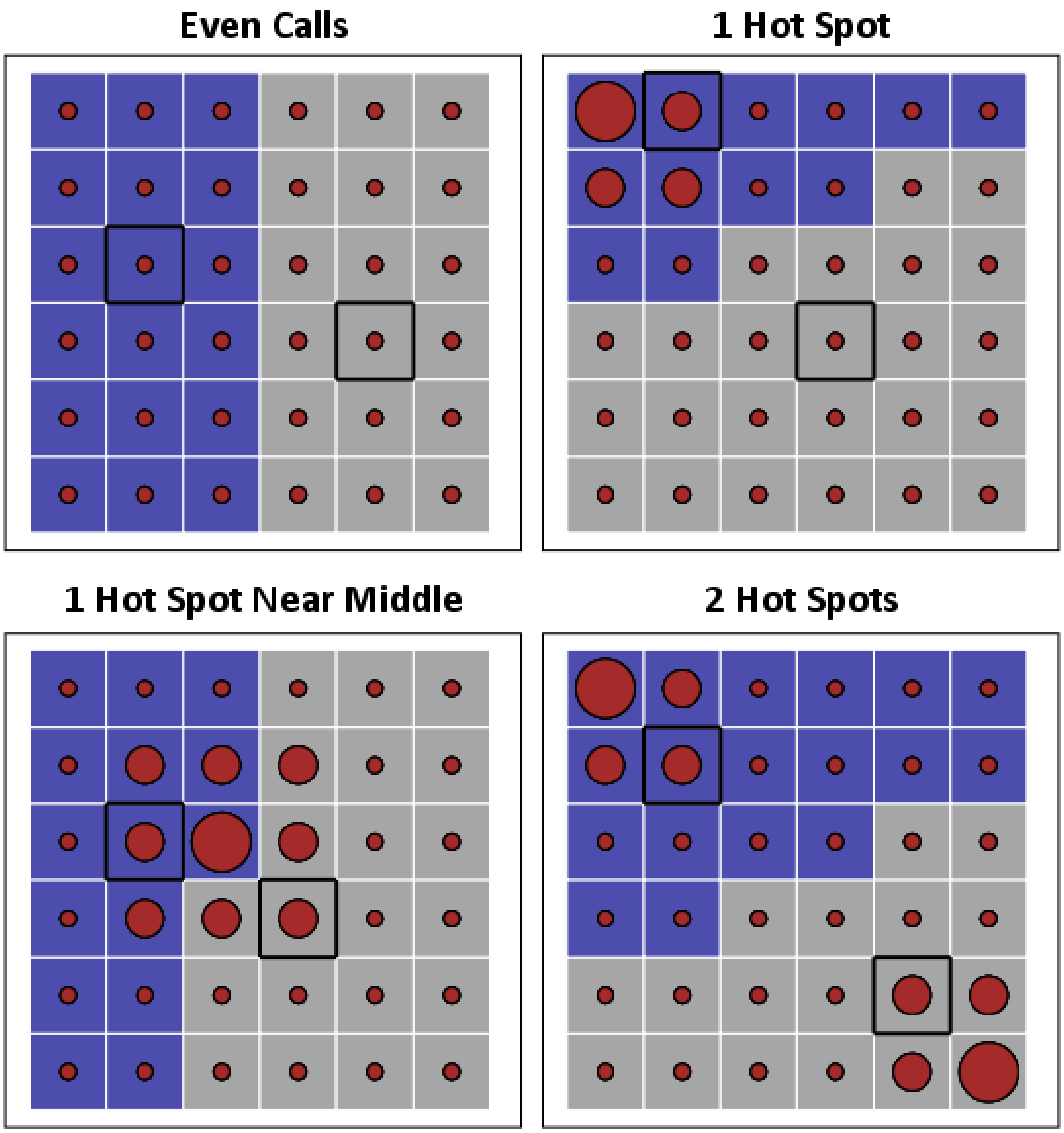

Figure 2: Example outputs for generating patrol areas on a regular grid. Circles represent number of crimes, with larger circles representing more crimes (proportional to the circles area), and colored squares represent the patrol areas. The square with the black outline is the source area for each patrol zone.

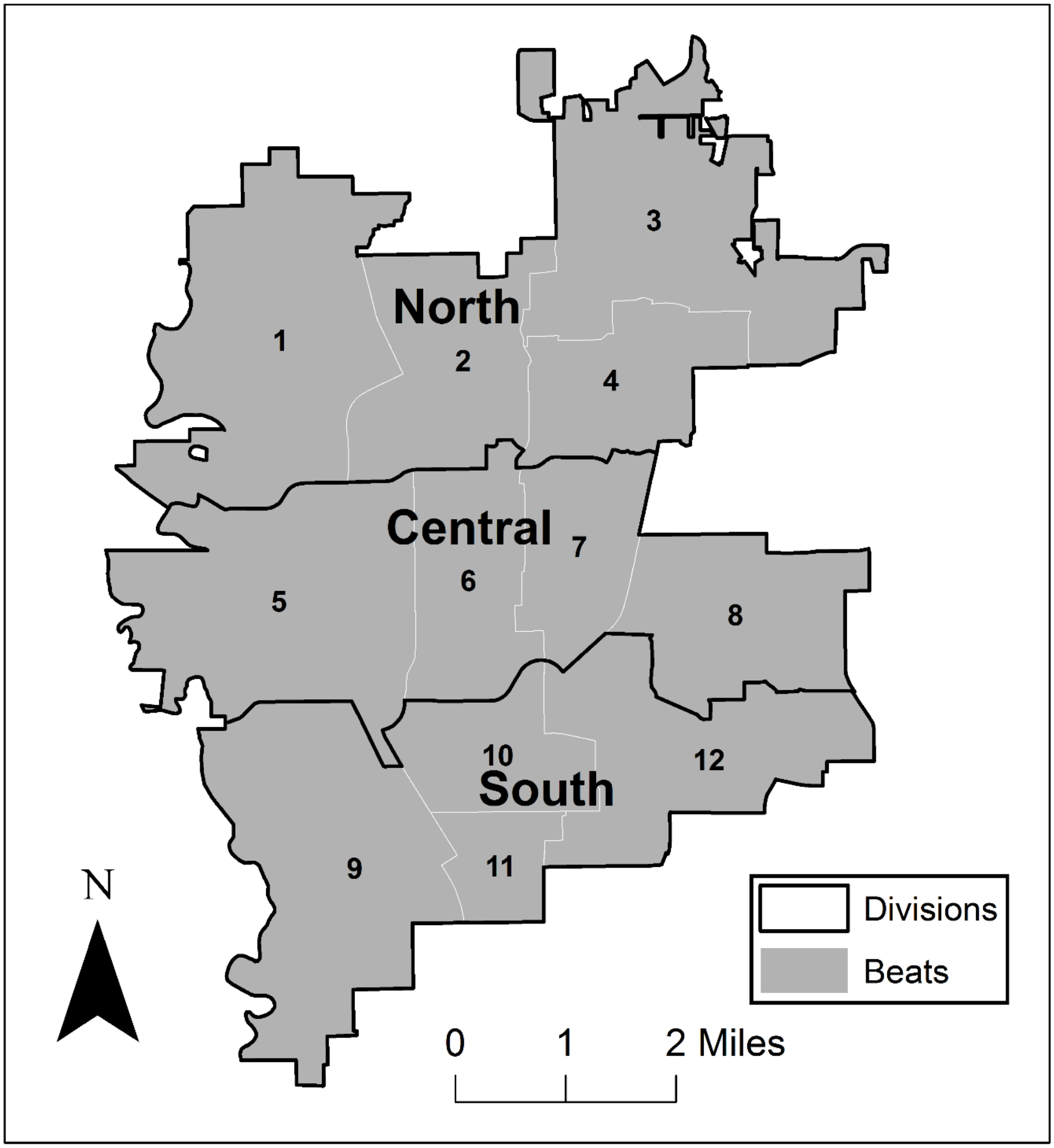

Figure 3: Current organization of police beats (numbered 1 through 12) nested within three divisions (North, Central, and South) in Carrollton, Texas.

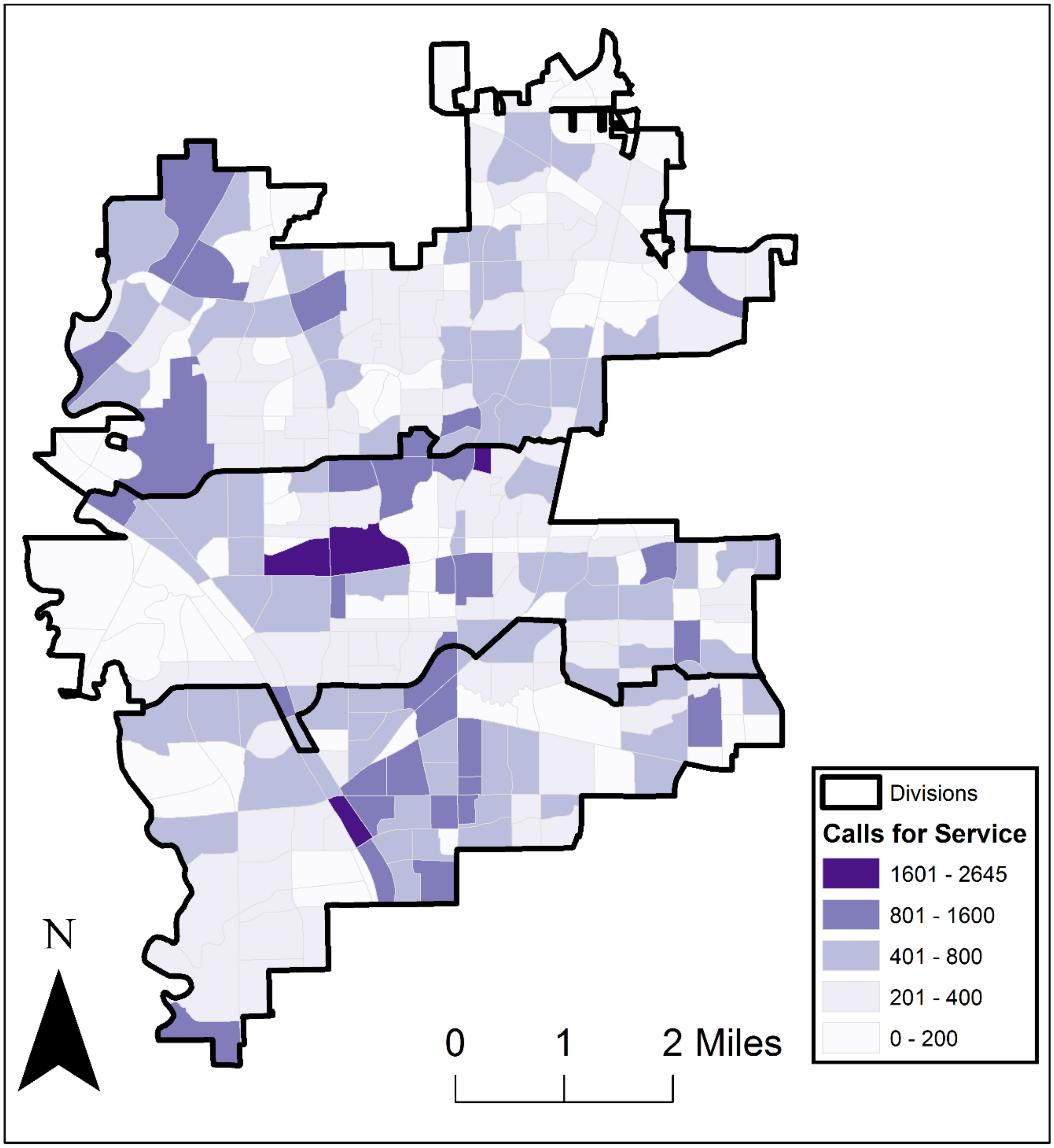

Figure 4: Spatial distribution of calls for service, from 3/12/16 to 8/29/17, aggregated to the Reporting Area level.

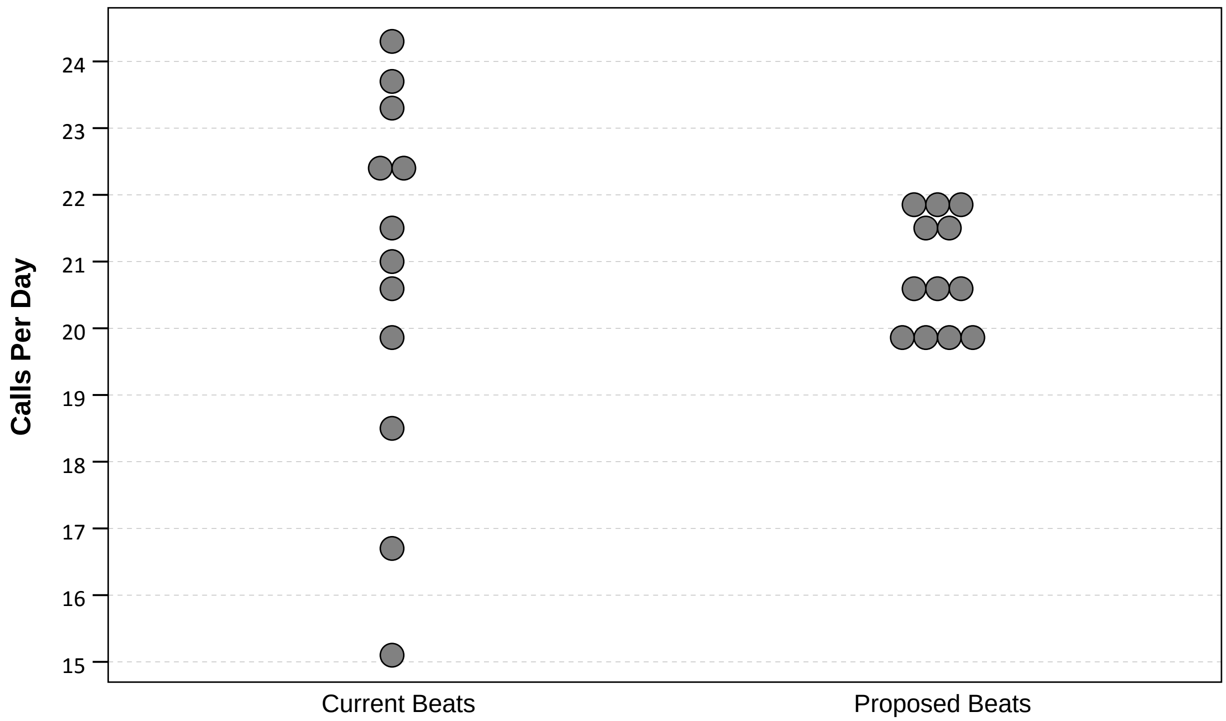

Figure 5: Workload distribution under current patrol areas versus workload distribution for the suggested patrol areas. Each dot is a patrol area, and the Y axis shows the average number of calls per day. This illustrates how the model constrains workloads to be very equal. The current patrol areas range from averaging 15.1 to over 24.3 calls per day, whereas the suggested solution constrains the constructed areas to have between 19.8 and 21.8 calls per day.

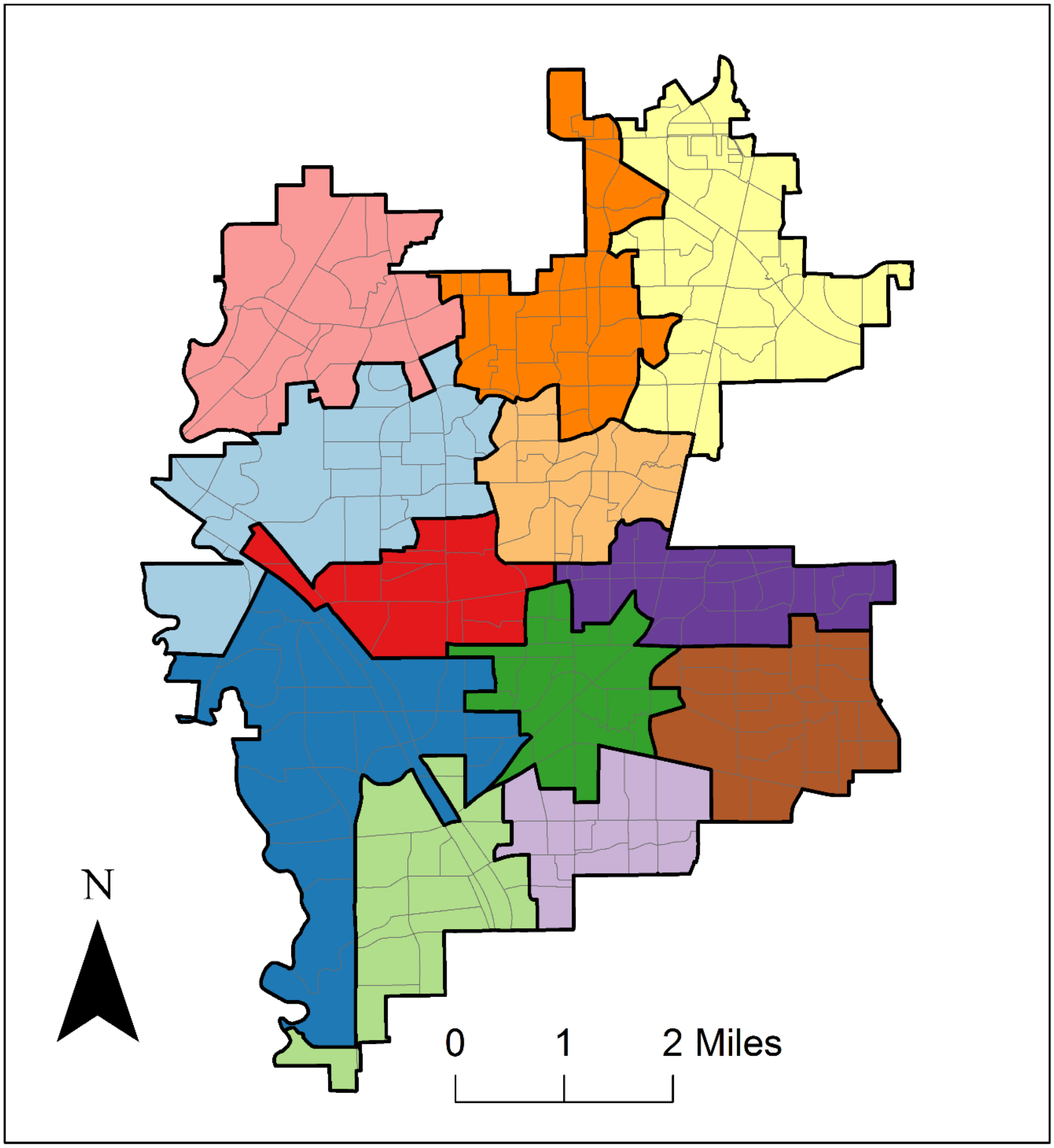

Figure 6: Optimal spatial arrangement of clustering the reporting areas into twelve new beats based on the p-median model. The maximum inequality in the beats is 5% or less.

Table 1: Cost Table for Assigning Source Atoms to Destination Atoms. Numbers are the distance between the Source row and the Destination column multiplied by the number of calls in the destination.

	Destination			
Source	*A*	*B*	*C*	*D*
A	0	1	4	9
B	4	0	2	6
C	8	1	0	3
D	12	2	2	0

Table A.1: Variable Definitions

Notation	Description
s	Subscript indicating source atom
d	Subscript indicating destination atom
C_d	number of calls in destination *d* (can be weighted)
D_{sd}	distance between source *s* and destination *d* (can be travel time or road distance)
W_{sd}	When atoms *s* and *d* are contiguous, equals 1, otherwise equals 0
X_{sd}	decision variable whether source *s* is assigned to destination *d*, can equal 0 or 1
P	The total number of patrol areas to create
I_{max}	The maximum number of calls any patrol area can be allocated
I_{min}	The minimum number of calls any patrol area can be allocated