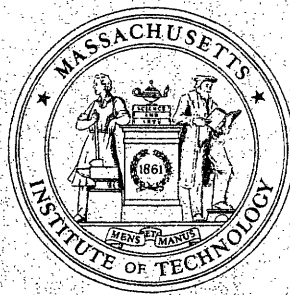


OPERATIONS RESEARCH CENTER

working paper



**MASSACHUSETTS INSTITUTE
OF TECHNOLOGY**

Methods for Allocating Urban
Emergency Units*

by

Jan M. Chaiken** and Richard C. Larson***

OR 003-71

June 1971

*This research was supported in part by the U. S. Department of Housing and Urban Development under Grant H-1056 and in part by the National Science Foundation under Grants GK-16471 and GI-5.

**New York City Rand Institute

***Electrical Engineering Department and Operations Research Center,
Massachusetts Institute of Technology. Consultant to New York City Rand
Institute

METHODS FOR ALLOCATING URBAN EMERGENCY UNITS

by

Jan M. Chaiken

and

Richard C. Larson

ABSTRACT

An urban emergency service system provides mobile units (vehicles) to respond to requests for service which can occur at any time and any place throughout a city. This paper surveys recent quantitative work aimed at improving the allocation policies of these systems, including determining the number of units on duty, designing response areas and patrol patterns, and locating service facilities. Recent models which provide insight into system operation are proposed to replace traditional rules-of-thumb as guides to allocation decision-making. The methods discussed are applicable to police and fire departments, emergency ambulance services, and certain other emergency services.

I. INTRODUCTION

Urban police and fire departments, emergency ambulance services, and similar urban emergency service systems comprise an important class of governmental service agencies that until recently has not benefited from systematic analyses of operational problems. These systems operate in a complicated environment that includes temporally and spatially varying demand patterns, both explicit and implicit administrative, legal, and political constraints, and often ill-defined mixtures of objectives.

Our purpose in this paper is to review those operational problems of these agencies which are related to the deployment of their vehicles and to report current progress on mathematical modeling approaches to these problems. The discussion focuses on the methods which are available, the extent of improvement that can be expected as a result of quantitative study, and the types of solutions that can be obtained. References are given so that the interested reader can pursue details which are not given here. Several of the discussed problems (e.g., determining how many units to have on duty) are common to many urban service systems we are considering, while others (e.g., allocation of preventive patrol effort) are experienced by only one agency.

Although we will not focus on the mathematical details of any of the methods, we hope to be sufficiently precise to bring out the following points:

- The performance of emergency service systems is often affected in counterintuitive ways by changes in procedures or deployment.
- Popular operational "rules of thumb" usually yield levels of performance which can be substantially improved by other methods.
- Simple changes in administrative procedure can often produce more significant improvements than expensive hardware systems or increases in manpower.

- The current state of research is sufficiently advanced in several areas so that many cities' emergency service systems could experience significant improvement simply by applying what is currently known.
- Additional research is needed in a number of areas.

II. DESCRIPTION OF AN URBAN EMERGENCY SERVICE SYSTEM

The class of urban service systems we are considering is characterized by the following properties:

- Incidents occur throughout the city which give rise to requests or calls for service (e.g., fire alarms, crime victim assists); the times and places at which these incidents occur cannot be specifically predicted advance.
- In response to each call, one or more emergency service units (vehicles) are dispatched to the scene of the incident.
- The rapidity with which the units arrive at the scene has some bearing on the actual or perceived quality of the service.

Examples of such emergency service units include fire engines and ladder trucks, police patrol cars and scooters, ambulances, emergency repair trucks for gas, electric and water services, and tow trucks.

Although all urban emergency service systems share the characteristics listed above, they may differ in certain significant details:

- Some emergency units are ordinarily found at fixed locations at time of dispatch. Others such as police patrol cars are mobile.

This distinction is important for both administrative and analytical purposes. For instance, in principle

it is possible to vary the location, size and shape of police patrol sectors at will, whereas the response areas of fire units must be designed in relation to the (fixed) locations of the fire stations. Also, the dispatch strategy for mobile units can often be improved by a variety of location-estimation techniques which are not needed if units are positioned at known locations. For instance, a police dispatcher could improve decisions by querying the cars as to their locations¹ or using information from an automatic car locator system.

The distinction between mobile and fixed-location units begins to break down during periods of high demand. At such times the units may be dispatched directly from one incident to the next, or they may be dispatched while enroute from a previous incident to their home location.² Under such conditions, system operation is not very sensitive to the distribution of initial locations, either fixed or mobile.

- Emergency services differ in the urgency of the calls they receive and in their ability to discriminate among types of calls in advance. For example, false alarms of fire are not at all urgent, but it is difficult to know which alarms are false.³ On the other hand, a telephone call to the police reporting a past burglary can be identified as not requiring the immediate response of a patrol car.⁴ In near saturation conditions the ability of an emergency service to distinguish the priorities of its calls critically affects operations. If a call can be identified as not urgent, the dispatcher may decide not to send any units, or he may hold the call in queue to await the availability of a unit near

the scene; he may even place a call in queue when some units are still available, thereby protecting his ability to dispatch units to future high-priority incidents. However, if each call must be assumed urgent, then none can be delayed in queue. Then, when many units are unavailable, the dispatcher must either send fewer units to each incident or dispatch units from greater distances.

- For some emergency services the time the units spend between servicing calls is used for another important activity. For example, it is widely believed that routine patrol by police cars acts as a deterrent to certain types of crime [33, 66]. If police cars spent nearly all of their time handling calls for service the preventive patrol function would suffer. Such an important secondary function is not present in all emergency service systems and should be distinguished from routine internal functions: rest, meals and training for the men, maintenance of equipment, and preparing written reports.

For units which do have an important secondary function, questions involving the dispatch of units cannot be answered exclusively in terms of the effectiveness of response to emergencies. For example, it may be desirable to place some calls to the police in queue simply to preserve the deterrent patrol. A fire dispatcher would rarely have occasion to make such a decision, since the available fire units are not engaged in any activity which could be judged more important than responding to an alarm.

III. FACTORS INFLUENCING THE SELECTION OF AN ALLOCATION POLICY

The allocation (or deployment) policy of an urban emergency service system determines the following properties of the system:

- (1) The total number of units of each type on duty at any one time. (This may differ by time of day, or day of the week, or season of the year.)
- (2) The number of men assigned to each unit.
- (3) The location or patrol area of each unit.
- (4) The priority attached to different types of calls, and the circumstances under which calls are queued.
- (5) How many units of each type are dispatched to each reported incident.
- (6) The particular units which are dispatched.
- (7) The circumstances under which the assigned locations of units are changed. (This operation is variously called relocation, move-up, redeployment, repositioning or reinforcement.)
- (8) When relocations are required:
 - the number of units which are relocated,
 - which particular units are relocated, and
 - what their new locations are.

We are concerned here with methods which can be used to select or improve allocation policies. But it should be noted that many operational aspects of an emergency service system which are not part of the allocation policy may have a major influence on the quality of service provided. Among the most important are:

- the assignments given to particular individuals,
- the procedures followed at the scene of an incident, and
- support functions: maintenance, supply, training, and administration.

Since cities differ widely in their properties and in the demands they place on their emergency services, it is clearly impractical to try to specify a single "optimal" allocation policy which can apply to all locales. Nor does "optimal" have much meaning for systems having a large number of objectives, many of them ill-defined. But the theory underlying the allocation of emergency units is now sufficiently well developed that we can identify the quantitative factors which should be taken into account:

- the nature of the geographical area to be used
- population density and land use patterns
- time and space distribution of calls for service
- number of units of each type required to service calls
- travel speeds
- service times at incidents

All of these factors will have some impact on the performance which can be expected from any particular allocation policy.

But the most fundamental difficulty arises in trying to measure performance. How do we know when an allocation policy is good? Which ones are better than others? These questions can only partially be answered in quantitative terms, and other factors specific to the type of service system being considered play an important role.

In the sections which follow, examples of various quantitative methods will be discussed and analyzed. Insights and interpretations derived from these methods will be discussed, and an attempt will be made to indicate how these insights can usefully guide an administrator's thinking about allocation policies.

There are basically two ways in which an agency administrator can use these methods. The first and most common application entails the

selection of improved methods for assigning a pre-determined total number of men. Even in cases where there is a possibility of hiring additional personnel, it is important to determine what level of performance can

be achieved with existing manpower. The cost of even a single additional emergency unit is usually large enough to justify whatever analysis may be required to bring about the same performance level without the added unit.⁵

The second application, which has been recommended in planning and administration tests [33, 67], is to use quantitative methods for determining the overall number of men required to meet some prespecified objectives. Since personnel costs often constitute as much as 90 to 98 per cent of the budget of an emergency service, this use of quantitative methods to derive an allocation policy is virtually equivalent to determining the budget level of the service. However, this happens only on rare occasions. More typically, the administrator of an emergency service in a major city is faced with a total budget (or authorized strength) which he cannot change very much in one year.

A challenge in future years will be to implement both uses of these methods, with current systems performing effectively under a given budget level and required future budget levels determined from analyses of predicted demands and models of operation.

IV. DETERMINING THE NUMBER OF UNITS TO HAVE ON DUTY IN EACH AREA

A. METHODS BASED ON GEOGRAPHY AND LAND USE

Decisions regarding the total number and locations of a city's fixed facilities (e. g., fire houses and police precinct stations) have usually been made solely on the basis of geography and land use patterns. This reliance on geographical factors has been reinforced by geographical standards and regulations which apply to many cities. For example, the Standard Grading Schedule of the American Insurance Association [2] is used in most U. S. cities (excluding New York) to establish fire insurance rates. As a rule, cities will attempt to meet as many standards in the schedule as possible, so as not to

have a lower rating than necessary. But for cities with population over 200,000, the only criteria provided by the Schedule for the number of fire engines and ladders to be located in each part of the city are based exclusively on geography and land use. For certain "high value districts" the Schedule requires every point to be no further than one mile from an engine company and no further than 1.25 miles from a ladder company. Moreover, within 1.5 miles of any point there must be at least 3 engine companies, and within 2 miles at least 2 ladders. These standards may vary slightly from area to area, but for each type of area, the same kind of geographical standard applies.

The main deficiency of geographical standards is that they are meant to be substitutes for standards involving the time between receipt of a call for service and the arrival of emergency units. But this response time depends on many factors aside from geographical ones: the delays incurred in dispatching the units, the speed at which the units can travel, and the probability that particular units will be available. (It is little comfort to know that a fire house is within a mile of your home if the units located at that house would very likely be busy at the time you had a fire.)

Thus, as a general rule, it is not possible to determine whether an adequate number of units are located in each geographical area solely by inspecting a map of the city which shows the home location of each unit.

B. ANOTHER TRADITIONAL APPROACH: WORKLOAD OR HAZARD FORMULAS

Instead of relying on a single factor such as geography, so-called workload or hazard formulas combine in a subjective manner virtually all factors which might be thought relevant for allocating units. They give an appearance of accuracy because of the large number of factors included.

Perhaps the most well-known such formula was developed for police use by O. W. Wilson in the late 1930's [65, 66, 67]. Wilson combined indicators of activity (such as number of arrests, number of calls for service of particular types, number of doors and windows to be checked) with other factors

(such as number of street miles, number of licensed premises, and number of crimes) to arrive at a "hazard score" for each area. An area's score is computed by taking a weighted sum of the fractions of each of the factors associated with the area. The weights are subjective indicators of relative importance. In applying the formula, the total number of men (or patrol cars) are to be distributed among the areas in direct proportion to their hazard scores.

This procedure often produces unsatisfactory allocations that may have to be "juggled" by hand computations to arrive at a "reasonable" allocation. For instance, the 5 or 10 per cent weighting often given to calls for service is not nearly adequate to avoid lengthy queue delays in certain areas during periods of high demand. The inherently linear form of a hazard formula precludes description of the highly nonlinear and complex interactions among system components which are often observed in practice. Such a formula also attempts a simple deterministic depiction of a system in which many of the variables are probabilistic. In addition, since factors such as arrests are more likely in sufficiently staffed areas, hazard formulas may have the perverse effect of indicating a need for additional personnel in areas which are already relatively overallocated. But the major difficulty arises in trying to determine how to improve the selection of the subjective weightings, a problem for which there seems to be no underlying principles or guidelines.

Wilson should be credited with introducing hand-calculable quantitative methods into a policy area which previously had depended on "command judgment" above. But increased implementation of digital computers and recent developments in the theory of allocating emergency units have made reliance on a hazard formula unnecessary.

C. MODELING METHODS

The approach emphasized in this paper relies on constructing and analyzing mathematical models of operation. From a modeling viewpoint

urban emergency service systems have two distinctive features: (1) probabilistic demands and service requirements over time and (2) probabilistic distribution of incidents and response units over the space of the city. The first gives rise to congestion when too much service is demanded in too short a time period. It is natural to examine such phenomena using queuing theory, a branch of operations research concerned with the performance of service systems where customers, ships, telephone calls, etc., may have to wait "in line" until they receive the desired service. The second feature, corresponding to the system's spatial characteristics, gives rise to travel time distributions, patrol patterns, etc. To the extent that space and time considerations can be separated, it is convenient to examine the spatially dependent properties of the system using essentially geometrical considerations. Then, in using these models in an allocation algorithm or other complicated decision application, we can incorporate the several criteria of each of the separate models, including both queuing and spatial properties of operation. These topics are discussed further in the following paragraphs.

Queuing Models

In applying queuing theory to urban emergency service systems, the "customers" are the calls requiring dispatch of a vehicle. The consequences of having to place such a call in queue may be quite serious. Indeed, it is characteristic of emergency systems that a person's life or well-being may well depend on the immediate dispatch of a unit.⁶

Thus, a primary objective of all urban emergency systems is to reduce to a low level the possibility that an urgent call will have to be placed in queue for more than a few seconds.

The probabilistic nature of the arrival times and service times of calls is such that one can never guarantee that every call will result in the immediate dispatch of a unit.⁶ Thus, the objective of a queuing analysis is to assure

that the probability of an important call encountering a queue is below some specified threshold (such as 1 in 50) or that the average time to wait in queue is below some specified limit (such as 1 minute).

To take a simple hypothetical example, we might imagine a city in which each police patrol car is assigned a geographical response area ("sector" or "beat") in such a way that no other car responds into his area. Then, whenever a given patrol car were busy, all calls from his sector would have to be placed in queue. Such an arrangement constitutes a "single-server queuing system," and, given reasonable assumptions, standard textbooks provide formulas which give good estimates of the probability that a call will experience a queue or the average time it will wait in queue [49, 51]. One could use these formulas to determine the required number of patrol units as follows: A threshold would be selected for the maximum value of the probability of a queue (or average waiting time) to be permitted in any sector; then the sectors would be selected small enough to assure that the threshold is not exceeded. The total number of sectors designed in this way would then determine how many units are needed.

Although this example is instructive, we are fortunate that no urban emergency service system actually operates in this fashion. At fixed manning levels, other arrangements lead to fewer delays; or, at fixed performance levels, other arrangements require fewer total response units.

The simplest generalization of this model which has been usefully applied to real emergency services is the following: There is a fixed number N of vehicles, either located at one place (e. g., a hospital) or distributed throughout a region, and each call requires the dispatch of one vehicle. A call is placed in queue only when all N vehicles are busy servicing prior calls. All calls are assumed to be identical in terms of their importance and service time distribution. With certain additional assumptions, the formulas describing this queuing system are easily calculated [51].

Stevenson [59] has applied this model to determining how large the number N of ambulances in a region must be (depending on the arrival rate of calls) to assure that only 1 in 100 (or 1 in 20, or 1 in 10) callers must wait for the dispatch of an ambulance. Given an estimate of the arrival rate during each time period, an administrator can select a desired threshold probability and determine how many ambulances to have on duty by time period.

The results of Stevenson's calculations have a property which is common to nearly all realistic queuing models: the number of units needed increases with the call rate, but not in direct proportion to the call rate. Thus, a doubling of the call rate would produce a requirement for fewer than twice as many units. This observation constitutes an additional argument against using call-for-service rates in a linear fashion in workload formulas.

The same model has been used in St. Louis for the allocation of police patrol cars [48]. The city is divided into nine patrol districts, and a call is assumed to enter a queue whenever all the cars in its district are busy. For each four-hour time period, the Police Department estimates, using the N -server queuing model, how many cars will be needed so that at most 15 per cent of each district's calls will experience a queuing delay.⁷

The next step in complexity of queuing models occurs when various call types have different priorities. As an example, Cobham [20] developed a model which assumes that higher priority calls are served first, but retains the assumptions that one unit responds to each call and that all service times have the same exponential distribution. Although in most police departments calls are not explicitly assigned priorities according to specified rules, Larson has found this model useful as an approximation to current performance of police dispatchers and as a tool for analysis of the potential benefits of more precise priority schemes [43]. It has the advantage that it places emphasis on reducing the delays which are associated with important calls.

Greater realism could be introduced into this model by (1) permitting each priority level to have different service time distributions, or (2) allowing the

service time to vary with the number of units busy.⁸ But the effort required to design such models cannot be justified unless allocation decisions are found to be sensitive to the current model's assumptions and unless a comparable effort is devoted to collecting and analyzing service-time data.

One refinement of the multi-server queuing model has been found practical, and indeed necessary, for predicting the number of units busy at operations of a fire department. Fire dispatchers typically send several units to each alarm, while the previous models assume that one unit is sent to each incident. In addition, fire units do not all complete service at the same time since each may have distinct duties to perform.⁹

Chaiken [17] has developed a quite general queuing model which allows for these features of fire operations. In particular, in this model

- different types of alarms may require different numbers of units of various kinds;
- the units may arrive singly, or in groups, and they may depart in similar fashion; and
- the length of time the units are busy at the incident depends on the type of incident.

This model assumes that queues are never permitted to develop. Instead, whenever units are required in one region of the city, it is assumed they will be dispatched from there or from another region, if necessary. In applying this model, one does not try to assure that the probability of encountering a queue is small. Instead, one requires that the probability of needing to dispatch units from a distant region be small. The computations required for this model, as well as the other queuing models discussed, are readily carried out on a computer.

Applying the model in New York City [16], Chaiken found that at low alarm rates (such as occur in the early hours of the morning), the numbers of units needed to meet the requirements of the queuing model are well below the numbers needed to meet simple geographical requirements;¹⁰ therefore the geographical factors predominate. However, in some parts of the city at

times of high alarm rate the queuing model implies a need for more units than would be suggested by geography alone.

The same model could be utilized for analyzing operations in other emergency services which dispatch two or more units to certain types of incidents.

Travel Time Models

Although the typical travel times of four to ten minutes may be dominated by queuing delays during periods of saturation, travel time may comprise the greatest fraction of total response time during normal operating periods [43]. Thus, models are required which relate properties of travel time to the number of units on duty, geographical characteristics, arrival rates of calls, and service times at incidents. In periods of relatively light or moderate demand, the travel time models can replace traditional geographical factors¹¹ in determining the number of units to have on duty. In developing such models one has to take into account the probabilities of particular combinations of units being available, the dispatching rules, and other details of system operation.

Several geometrically oriented models show that average travel time is inversely proportional to the square root of the number of available units per square mile, with the proportionality constant dependent on the spatial distribution pattern of the units. If units are randomly located in a homogeneous region, with an average of r units per square mile, one model [43, p. 323] shows that the average right-angle travel time for the closest unit is approximately $0.63/(v\sqrt{r})$, where v is the response speed of the unit. If the units are not randomly located, but instead are located in such a way as to minimize average travel time, the constant 0.63 is replaced by 0.47 [46]. Kolesar [40] has found similar results for the average travel time of the second-arriving unit when two are dispatched. The models also provide estimates of the probability that travel time will exceed any specified threshold.

If these results are tied to the queuing methods described above, the resulting model can be applied to determining how many units to have on duty. The queuing model is used to determine the probability that n units will be busy, given N total units on duty and other characteristics of the system. Larson [43, p. 328] follows this approach to estimate average travel time, assuming each available unit is randomly positioned. It is similarly possible to calculate the expected travel time under the assumption that the available units are moved, if necessary, so as to occupy the locations which give minimum travel time.

Kolesar [40] has studied the numerical output from these models, as well as data derived from experiments or from more complex models, and has concluded that a reasonable approximation in cases where unavailabilities are not too severe is to assume that the average travel time is inversely proportional to the square root of the average number of units available. This result leads to the following method for determining how many units to have on duty in each region:

- Determine the constant of proportionality between average travel time and average number of units on duty, using data collected in the region.
 - Estimate the arrival rates for calls in the region.
 - Assuming N units are on duty, using queuing methods to calculate the average number of units available, and use this number to calculate the average travel time.
 - Select N so that the average queuing time (or the probability of incurring a queue) and the average travel time do not exceed specified thresholds.
- Such calculation provides sensible allocations for all regions, whether they experience high call rates or not.

Methods Using Several Criteria

The simple analytical models described above may not, by themselves, be sufficient to determine the number of units needed. First, for services which engage in important activities other than response to calls, criteria

relating to these activities have to be taken into account.¹² Second, and more fundamentally, easily quantifiable criteria (probability of encountering a queue, average travel time) do not often have a clear relationship to the true performance of the system. For instance, one would like to know the actual benefits which accrue by decreasing response time. Such benefits might be lives saved, stolen goods recovered, property damage averted, etc. Although preliminary research along these lines has been performed [34, 11, 27], the currently available empirical information is not an adequate foundation for an administrator's use in selecting allocation policies.

Thus, at present, one is forced to use available performance measures such as response times. A careful and realistic use of such measures can provide reliable proxies for more fundamental measures, as has been discussed by Carter and Ignall [12].

Given such a necessary reliance on performance measures, an administrator would usually want to employ several simultaneously to arrive at reasonable allocations. In addition to queuing and travel times, he could incorporate factors pertaining to other activities (e. g., preventive patrol) and to administrative matters (e. g., workloads).

Two criteria, travel time and response workload, are analyzed in Carter and Ignall's queuing model for determining the extent to which an added fire unit provides relief to overworked units in its area. It would be natural to assume that when units are added to a command, the number of responses made by each of its units would decrease, and this may be one of the secondary benefits of adding units which is particularly interesting to an agency administrator. However, since a full response to fires ordinarily requires several units, adding a new unit will increase the chances of a full rather than partial response, thereby increasing the total number of responses of units in the area. Thus their model shows that if it is desired to reduce the workloads of units in addition to improving the response time, a greater number of units may be needed than is suggested by simpler models.

Larson [43] has developed a dynamic programming model for allocating police patrol cars to commands (e. g., precincts) which incorporates a queuing model, a somewhat complex travel-time model, and a model of the frequency with which cars pass by an arbitrary point while on preventive patrol. In addition, it is possible to include restrictions on the smallest number of units which can be assigned to any one command and a variety of other criteria which may be supplied by police administrators. For each criterion a patrol administrator specifies a desired threshold, or "policy objective," which may vary with command. For instance, for a particular command it may be decided that the average travel time should not exceed four minutes. Then, the algorithm supplies the command with enough patrol units so that this objective and all other policy objectives (constraints) are satisfied. Once the objectives are met, the queuing delay is treated as a variable to be minimized using whatever additional patrol units are available.

With limited police resources, it is possible that a specific set of policy objectives is unobtainable. If so, the algorithm indicates the additional number of patrol units required to meet the stated objectives. To allocate the currently available number of units, the algorithm then requires a more modest set of objectives.

Compared to the allocations derived from a hazard formula, the algorithm-derived allocations appear to reflect more fully the operating characteristics of the system. For instance, in one large city the results suggest that during periods of relative congestion (e. g., Friday and Saturday evenings), average queuing delay can be decreased significantly by diminishing resources in residential commands with relatively light demands and increasing resources in "core area" commands which are heavily loaded. Such a redeployment of resources does not noticeably degrade performance in residential commands since sufficiently many patrol units are retained to satisfy all policy constraints. Yet, average queuing delays in core areas can be reduced often from thirty minutes to less than two or three minutes.

Although such a finding may not be surprising, the calculation of this reallocation would be extremely tedious without the assistance of a computer

algorithm which can compute the effects of each alternative and quickly discard "bad" ones. And, without models of patrol activity, it would not be possible to predict whether each alternative allocation satisfies the policy constraints and reduces the delay at the dispatcher's position in the best possible way.

In general, the quality of the allocations which an administrator can expect from any of the models described above depends on how much effort he is willing to have his staff devote to the application. An analyst who is not a member of the concerned agency cannot make an appropriate determination of what constitutes an "excessively long" delay before the arrival of a unit, or how much preventive patrol will be considered adequate, or what level of workload is "too great."

In the case of fire departments, where the various units dispatched to a single incident may arrive at different times, the analyst is not even in a position to know which arrival patterns are "better" than others. However, the field chiefs, who are completely familiar with their department's operating procedures at a fire, can provide valuable information. Through asking a series of questions such as "Would you prefer two fire engines arriving 1.5 minutes after an alarm, or one arriving at the 1-minute mark and one at the 2-minute mark?", it is possible to derive a chief's utility function for arrival times. Work in progress by Keeney [37] to develop such utility functions should make it possible to select the allocation of units so as to maximize a chief's utility of the resulting patterns of arrival times.

One final comment: in regard to any of these methods for determining how many units to have on duty, it should be noted that there may be some difficulty in assigning individuals to shifts or tours of duty which best "fit" the desired assignments. Legal and administrative constraints can make this problem quite difficult. A heuristic approach is discussed by Edie [24]. A more general approach using a computer algorithm has recently been reported by Heller [29].

V. DESIGN OF RESPONSE AREAS

A problem commonly shared by all spatially distributed urban systems is the design of response areas (districts or sectors or beats) that indicate where a particular patrol unit, fire engine, or ambulance is to have primary responsibility. In designing these administrative areas, agency administrators have stated several diverse (often conflicting) objectives:

- minimization of response time;
- equalization of workload;
- demographic homogeneity of each area;
- administrative convenience.

No single mathematical technique for design of districts is likely to take into account all the relevant factors. Yet, even some of the more simple, recently developed models have provided more insight into the problem than was previously available with ad hoc "rules of thumb" and, in fact, have shown several such rules to be invalid in most cases.

SINGLE SECTOR MODELS

Traditionally, police planners have been instructed to design patrol sectors as squares, circles, or as straight lines along particular streets. The idea behind square or circular sectors is to keep at a minimum the time required for the patrol unit to travel to the scene of a reported incident in its sector. For instance, O. W. Wilson states that "... a square beat (sector) permits a maximum quadrilateral area with a minimum distance between any two possible points within it."¹³ [67, p. 228]

One factor not considered in this statement is that, with mobile patrol units, travel speeds may depend on direction of travel; in such a situation, it will be desirable to design the sector so that the longer sector dimension corresponds to the direction with higher travel speeds. Using quantitative techniques, it is possible to predict the travel time characteristics of any

proposed sector design and thereby determine which designs actually do minimize some indicator of travel time.

As an example, consider an urban region in which the streets form a mutually perpendicular grid (e.g., as occurs in central Manhattan and certain other cities), running, say, east-west and north-south. Then, a shortest route of travel for the assigned patrol unit requires the unit to traverse the total east-west distance, plus the total north-south distance, between the unit's initial position and the position of the incident.¹⁴ Given some simplifying assumptions Larson has shown that average intrasector travel time is minimized by designing the sector so that the average time required to travel east-west equals the average north-south travel time. Since it is not unusual to find regions (such as in Manhattan) where the north-south speed is about 4 times as great as the east-west speed, this implies that the sectors should also be 4 times as long in their north-south direction. In this case, such a sector design can be expected to reduce average intrasector travel time by approximately 20% over that obtained by the rule-of-thumb design - square or circular sectors.

Some of these ideas have already been applied by Bottoms, Nilsson, and Olson in the city of Chicago [7]. They have constructed a new sector plan of the city using rectangular sectors designed so that the average intrasector travel time never exceeds approximately three minutes.

Certain complications to travel involving one-way streets or obstacles such as railroad tracks would increase average travel time. Larson has computed the mean extra distance traveled due to these complications. Although the results indicate a general insensitivity to most complications, certain responses involving one-way streets may require three or more additional minutes for the unit to reach the scene.

Similar results are found when drawing boundaries between the response areas of units which have fixed locations. If all the calls in a district are to be served by units from the fixed facility in that district, then the dividing lines

must consist of points which are an equal travel time from two facilities, rather than an equal travel distance [46].

INFLUENCE OF INTERSECTOR COOPERATION

When an incident is reported from a response area whose units are busy, most emergency service systems will dispatch an available unit nearby in another response area. Such an arrangement is nearly mandated by queuing considerations, but it introduces subtle complications into the design of response districts. In the case of mobile units, it even raises questions about the need for restricting response areas of the units to be nonoverlapping. These consequences of intersector cooperation will be discussed below.

Police Patrol: "flying"

Police administrators are often heard to argue in favor of assigning patrol units to nonoverlapping sectors in order to establish a "sector identity" on the part of the patrol officer. This identity, which derives from patrolling and from citizen contacts made while responding to calls for service, is supposed to cause the officer to feel responsible for public order in his sector. However, given nonoverlapping sectors, one can show by a simple probabilistic argument [44] that the fraction of dispatches which cause a unit to travel outside its own sector is usually equal to or greater than the fraction of time that units in that area are unavailable for dispatch. This result does not appear obvious at first glance, and it has been quite difficult to convince police administrators of the following type of statement: "If your patrol units are 'busy' 40 per cent of the time (a typical value), then at least 40 per cent of all dispatch assignments cause the assigned patrol unit to leave its 'own' patrol sector. In turn, at least, 40 per cent of all citizen contact occurring while responding to calls-for-service takes place in sectors other than the patrol unit's 'own' sector."

The predicted amount of intersector dispatches (called "flying" in some police departments) has been substantially verified both by our own work [44] and by the reports of others [47]. These data showed that the amount of intersector dispatching is never significantly less than the percentage of time unavailable, and it may be significantly more. Intersector dispatches ranged from 37 to 57 per cent of the total.

The extent of flying brings into question not only the philosophy behind nonoverlapping sectors but also a widely popular statistical procedure for computing workloads of police patrol cars. Usually a sector is associated with a "workload" which is proportional to the number of calls for service generated from within the sector. Thus, for instance, sector A would have a "workload" three times as great as sector B if sector A generated three times as many calls for service as sector B. And, it would usually be assumed that a patrol unit assigned to sector A would work three times as "hard" as a unit assigned to sector B. Using our knowledge of the flying phenomenon, we know that the latter assumption is false and, in fact, the car assigned to sector B may be dispatched to calls in sector A sufficiently often so that both may work about equally "hard." Thus, to keep track of the workload of a patrol unit one must record the dispatch assignments of each unit and not the rate of calls for service from individual sectors.

There is one additional property of nonoverlapping sector systems that we should mention. It involves the "randomness" of preventive patrol. With nonoverlapping sectors, preventive patrol coverage in a sector is reduced to zero whenever the corresponding patrol unit is busy. Anyone, including potential criminals, can monitor a patrol unit's activity in some manner (e. g., visual observation, listening to the police radio) and determine when a particular car is not patrolling. Then, since units are assigned to nonoverlapping sectors, a crime can be committed with near zero probability that the patrol unit will pass during the commission of the crime.

Given the undesirable features of a nonoverlapping sector system, how can an administrator revise and improve operations? First, if the sector concept is to be retained, the large fraction of calls which are low-priority¹⁵ (i. e., they do not require rapid response) can be "stacked" and handled by the car assigned to the sector of the call when that car becomes available. This procedure reduces the amount of flying and enhances "sector identity."

Second, the sector concept can be modified so that patrol units are assigned to overlapping areas (or sectors). This procedure enlarges the area with which each patrol officer should develop an "identity." In addition, it increases the "randomness" of patrol, a desirable outcome which is not achieved simply by stacking on nonoverlapping sectors.

Clearly, the number of possible combinations of alternatives is very large. Fortunately, quantitative methods using mathematical models of operations can structure one's thinking about these alternatives and, in fact, can predict the extent of improvement to be gained by an particular combination [43].

Response Districts for Fire Units

A fire unit's primary response district consists of all points to which it would be dispatched if an alarm were generated there, even if all units were available. In the event of unavailabilities, the unit may also respond to alarms elsewhere. Fire departments have traditionally designed districts so that the dispatched units are the ones closest to the fire. This means that all points on the dividing line between two districts are equally close to some pair of companies.

With the modification of interpreting "closest" in the sense of "shortest travel time," one might expect this procedure to minimize overall response time. However intuitively reasonable this rule-of-thumb may appear, a recent analytical study by Carter, Chaiken, and Ignall [14] has shown that

"equal travel time" dividing lines are usually not optimal and that overall average travel time is minimized by following a policy that often requires a unit other than the closest unit to be assigned to a particular fire.

The philosophy underlying this result is one that often appears in systems with unpredictable demands in the near future - it may be preferable to incur an immediate cost (e. g. , travel time) that is slightly greater than the minimum possible immediate cost so that the system (e. g. , the collection of all fire apparatus) is left in a state which best anticipates future demands. That is, assigning, say, the second closest unit to the most recent fire may result in favorable positioning of units for the next reported fire, thus minimizing overall response time. Assignment of the closest unit to the first fire might have required an unusually large amount of time to respond to the next reported fire.

Carter, Chaiken and Ignall have also shown that the boundaries which minimize overall average response times will, in many cases, also reduce workload imbalance (where workload imbalance is defined to be the difference in the fraction of time worked by the busiest unit and by the least busy unit). Thus, implementation of their derived procedures results in two types of gains - response time reduction and workload imbalance reduction.

Their boundary structuring procedures have been worked out in detail for the case of two cooperating units; current research is being directed at extending the results to systems with many cooperating units. The qualitative features of the results have already been used to arrive at preferable dividing lines in New York City Fire Department operations - and these results are currently being implemented.

VI. LOCATING UNITS AND FACILITIES

Closely related to problems of response area design are problems of location, including

- which site to select for an additional facility;
- when consolidating two or more existing facilities into one new facility, where to place the new facility;
- pre-positioning, or where to locate units at the start of a tour;¹⁶
- repositioning, or how, and under what circumstances, to change the locations of units during a tour to correct for unavailabilities as they develop.

Although there is an extensive literature on the subject of "facility location," most of it is presented in economic terms and ignores probabilistic aspects of operations. ReVelle, Marks and Liebman [53] have recently reviewed a variety of applications of location theory to public sector problems, but none has yet been applied to the allocation of urban emergency units.

The work of Larson and Stevenson [46] is the beginning of a theory of facility location specifically designed for emergency services. Although further research is needed to eliminate some of their simplifying assumptions, this work tends to show that the optimal location of a new facility is rather insensitive to the precise location of existing facilities.

A considerably large body of analytical work has been completed, or is underway, concerning the repositioning of units during the course of a tour. Two examples are discussed below.

Local Repositioning: Police Patrol Cars

Consider the case of two square patrol sectors which have a north-south boundary in common. We will assume that each unit patrols its sector randomly, and the demands are uniformly distributed over the entire two-sector region. Each unit responds in its own sector, unless it is unavailable, in which case the other one responds.¹⁷ The question of interest is, "At the moment when one of the units become busy, is there any advantage to repositioning the remaining available unit? If so, how should this be accomplished?"

Whenever one unit becomes unavailable, consider the following three alternatives for the free unit:

Alternative 1: The free unit does nothing (i. e., no repositioning)

Alternative 2: Assign the free unit to patrol both sectors uniformly (i. e., "uniform repositioning")

Alternative 3: Station the free unit on the boundary line between the two sectors at the north-south halfway point (i. e., "fixed point repositioning").

It is straightforward to show [43, p. 351] that Alternatives 1 and 2 have the same average travel distance¹⁸ and the distance for Alternative 3 is 75 per cent as large. Thus, in an average travel distance sense, uniform patrol repositioning (Alternative 2) offers no advantage over no repositioning (Alternative 1). On the other hand, fixed point repositioning offers a 25 per cent reduction in average travel distance when compared to Alternatives 1 and 2. Similar results hold [43, p. 353] for regions of four cooperating sectors and for more complicated examples.

The results suggest that any local repositioning (among nearby sectors) is advantageous in a travel distance sense only if patrol is concentrated near the boundaries of the appropriate sectors. In practice, strict fixed point repositioning may not be advisable because of lost preventive patrol coverage; still, if the free unit must remain patrolling, a large part of the travel distance reduction can be retained provided the patrol occurs near the appropriate sector boundaries. In fact, we have heard patrolmen remark that on an informal basis two units will occasionally agree to "cover" both sectors when the other unit is unavailable; this "covering" usually takes the form of concentrated patrol near the common sector boundary. To gain travel distance reductions when such covering occurs, it is necessary that the dispatcher be aware of the identity of the cooperating units so that he can assign the covering unit to a call in the busy unit's sector.

Global Repositioning: Relocation of Fire Units

By "global repositioning," we mean the reassignment of one or more available units to areas which may be at some distance from the areas to which they are currently assigned. For many years urban fire departments have relocated available units when a number of units in one area become busy fighting a large fire. Indeed, these relocations are pre-planned, so that when a second alarm (or higher) is sounded, specified units respond to the fire while other units simultaneously move into certain fire houses which have been vacated by units at the fire. Such large-scale repositioning is not as widely used in other urban service systems, although situations continually arise (e.g., police precinct-level congestion) in which repositioning of forces would reduce travel times and dispatch delays or provide some preventive patrol.¹⁹

The following factors are important in determining whether to make a relocation:

- How long is the expected duration of the existing unavailability?
- How many units will have to relocate to accomplish the desired final locations?
- How long will it take for the units to travel to their new locations?
- Is the magnitude of the expected improvement in performance large enough to warrant the effort required to move units?
- Is there a good reason to believe that the units to be moved will be needed at their present locations in the near future?

Work still in progress at the New York City-Rand Institute is designed to produce an algorithm which will operate in a computer assisted dispatch system and will recommend relocations both for large fires and for unavailabilities which occur through an accumulation of smaller fires.

Several approaches have been tried. Swersey [61] developed an integer programming model to determine which fire houses should be empty and which full. His objective was basically to minimize the average travel time to incidents, taking into account the average time that busy units would remain busy. In addition his procedure provided a penalty for each unit relocated. Once a solution to this model has been found, a standard assignment problem can be solved to recommend which units should move to which empty houses. Unfortunately, it was not possible to solve Swersey's model rapidly enough, using either branch-and-bound or approximate heuristic techniques [63], to make it an appropriate tool for real-time applications.²⁰

The relocation method which is now planned for implementation has been developed by Kolesar and Walker [39] based on a suggestion of Chaiken. Instead of focusing on average travel time, which is not sensitive to many states of the system, this method utilizes ideas of "coverage." A point is said to be "covered" if at least one available engine company (or ladder company) is within T minutes of the location. If one or more neighborhoods are expected to be uncovered for an undesirable amount of time, a heuristic algorithm first determines which vacant houses to fill, then which available units to relocate to the vacant houses. While this algorithm is not "optimal" in any sense, it appears to compute very reasonable relocations using a comparatively small amount of computer time.

VII. CRIME PREVENTIVE PATROL

Although other urban service agencies have certain patrolling activities (e. g., fire departments "patrol" areas looking for fire hazards), the patrolling function is most important in urban police operations. A patrol unit is said to be performing "crime preventive patrol" when passing through an assigned area, with the officers checking for crime hazards (e. g., open doors and windows) and attempting to intercept any crimes in progress. By removing

opportunities for crime, preventive patrol activity is supposed to prevent crime. By posing the threat of apprehension, preventive patrol is supposed to deter individuals from committing crimes.

Most mathematical studies of police preventive patrol have occurred in the past several years, although some earlier work in "search theory" is also relevant to the problem. The term "random patrol" was introduced into police literature in 1960 by Smith [58] who stressed the need for unpredictable patrol patterns. Blumstein and Larson [5] developed a simple analytical model for spatially homogeneous random patrol in order to estimate the probability that police would pass a crime in progress. Elliott [25] quoted one of Koopman's [41] search theory results and attempted also to compute probabilities of space-time coincidence of crime and patrol. Recently, Rosenshine [24] analyzed certain problems which arise from the fact that the topology of streets may prohibit certain desired patrol patterns. Bottoms, Nilsson and Olson [7] have applied some of these ideas to operational problems in the Chicago Police Department.

To illustrate one simple model, consider the problem of estimating the probability that a patrolling unit will intercept a crime while in progress. For a crime of short duration T_c which occurs on street segment i one can argue that a reasonable upper bound estimate of the probability of space-time coincidence of crime and patrol is

$$P_c = s e_i P_p T_c / L,$$

where

s = speed of patrolling vehicle

e_i = a number between 0 and 1 reflecting the relative patrol coverage of segment i

P_p = fraction of total time spent patrolling²¹

L = a weighted sum of the segment lengths in the patrol sector, the weights being the e_i 's.

Even this simple formula provides certain insights. It illustrates that crime-intercept probability is directly proportional to coverage (e_i), fraction of time spent patrolling (P_p), duration of the crime (T_c), patrol speed (s),²² and inversely proportional to a weighted sum of segment lengths (L). The interaction of the response and patrol activities is also apparent: during periods of heavy call-for-service demand (i. e., when P_p is small), crime intercept probability is small. A potential trade-off exists between the amount of screening and/or delaying of calls for service, reflected in the value of P_p , and the likelihood of intercepting a crime in progress.

In applications of this formula one typically finds intercept probabilities below 1 in 100. Such low detection probabilities bring to question whether the threat of apprehension provided by preventive patrol is actually great enough to deter crime.

Given such scarcity of preventive patrol effort, any effective allocation of effort must reflect the relative likelihoods of crimes occurring at various times and places. Even raising intercept probabilities from 0.01 to 0.02, say, could result in a doubling of on-scene apprehensions. By structuring a model of preventive patrol operation one finds that the allocation of preventive patrol effort is mathematically similar to an allocation of search effort problem studied by Koopman [41]. Application of Koopman's search theory ideas to allocating relative patrol effort (e_i 's) to maximize detection probabilities yields the following properties:

- (1) On street segments with sufficiently low crime rates, no preventive patrol effort should be allocated.
- (2) On segments which should receive preventive patrol effort, the effort should grow slower than linearly with crime rate.

This behavior again illustrates the inadequacy of linear hazard formulas which imply that preventive patrol coverage should be directly proportional to frequency of crime occurrence. Although much more refinement of

Koopman's techniques is required before they can be implemented by police, we would expect the qualitative features of his solution to hold.

VIII. SIMULATION MODELS FOR EVALUATING PROPOSED CHANGES IN ALLOCATION PROCEDURES

An agency administrator is typically faced with a number of proposed changes in his allocation policy at one time. For example, the results of the models described in previous sections may suggest that he should add units at certain times of day, select new locations for some units, change response areas or patrol patterns, and modify the procedures for relocating units. In addition, certain technological innovations such as automatic car locator systems may have been proposed to accomplish some of the same objectives. Before making a choice among the alternatives, the administrator will want to have a realistic comparison of the benefits which can be expected from each approach.

For a thorough evaluation of such a comprehensive change in allocation procedures, one generally has to turn to much more complex and detailed models than the ones already discussed. Large-scale simulation models are typically used for this purpose. They can provide information about the effect of a proposed policy change on a wide range of variables: response times to particular types of calls, workload of units, queuing delays, availability of units, etc. Such simulation studies have been undertaken by Savas [56] to investigate the reduction of travel times which could be achieved by spatially repositioning ambulances, by Swersey [60] to analyze the operations of the dispatch centers of the New York City Fire Department, by Carter and Ignall [12] to compare a wide range of combinations of fire department allocation policies, by Larson [43] in a study of the allocation of police patrol and the potential benefits of utilizing a car locator system, and by Adams and Barnard [1] to study the value of an automated dispatch system for the San Jose Police Department. Recent work on efficient computer coding of

geographical data [35, 8] has been of some assistance in designing such simulation models of urban emergency service systems.

A common feature of these studies has been the finding that rather simple and inexpensive administrative innovations can often make a contribution to system performance which is equivalent to that of much more expensive hardware or increases in manpower. Swersey's study [60] provides such an example. In this case, the fire dispatching office in Brooklyn was experiencing an increase in alarm rates and consequent delays prior to dispatch of units which were beginning to be of some concern to the Department. The simulation showed that computerized methods for recording, storing, or retrieving location information about alarms would not solve the essential difficulty, which had to do with the fact that a single man had final responsibility for every dispatch decision. Swersey's suggestions for dividing this responsibility, a basically administrative change which has been implemented, provided substantially decreased delays during peak-alarm hours.

Similarly, Larson's simulation [43] has demonstrated that the absence of an explicit priority structure for calls to police departments produces unnecessary delays for urgent and moderately important calls. Most departments have been reluctant to implement such a structure, stating that their policy is to provide rapid service to all citizens. However, some departments [21, 22] have begun to implement priority structures based on quantitative information derived from such models.

In addition, the Larson simulation was used to study the best use of automatic vehicle locator systems in police departments. The technology of such systems is well developed [38, 55], and recently field tests and operational installations have been reported [10, 19, 23, 30, 32, 68]. Each system provides a central dispatcher with estimates of the positions of all service units (e. g. , buses, patrol cars, taxicabs) and with other status information (e. g. , current speed and direction, current type of activity).

In the Larson study [43, p. 289] analysis showed that superimposing an automatic vehicle locator system on a patrol force assigned to nonoverlapping sectors causes an average travel time reduction in the order of 10-20 per cent, the exact value depending on the fraction of time each car is busy, number of sectors in a command, spatial distribution of calls, etc. Analysis also showed that a system with fully overlapping sectors utilizing car position information has approximately the same travel time characteristics as current nonoverlapping sector systems without car position information. Thus, if there are reasons to want overlapping sectors, even to the extent of not assigning sectors to cars, there would be little or no degradation in travel time characteristics of the overlapping sector system, compared to current systems, provided high resolution car position information is available. Apparently, the pre-positioning advantages gained by assigning cars to mutually exclusive sectors are recovered by knowing exact car positions in a system with no deliberate spatial prepositioning.²³ (As mentioned in Section V, arguments based on "regional identity" and "randomness of patrol" seem to favor some type of overlapping sector plan.)

This analysis is an example of an instance in which applying technology to a system "operating as usual" may not fully utilize the new technological capabilities.

SUMMARY AND CONCLUSION

Many urban emergency service systems, including police departments, fire departments, and ambulance services, share common operational problems related to the allocation and dispatch of vehicles which respond to calls for service from the public. Recent studies have been directed at the use of quantitative methods for solving such allocation problems. This report has reviewed the major research topics and described those results which have produced, or are likely to produce, substantial improvements in system performance when implemented.

The following aspects of allocation policy were discussed:

- determining the number of units (vehicles) to have on duty in each geographical area at different times of the day or week.
- selecting the unit(s) to respond to a particular incident.
- determining the locations or patrol areas for the units on duty, and designing patrol coverage patterns.
- deciding when units should be re-deployed to improve service in areas where a large number of units are temporarily busy.

Although in many instances we do not yet know how to make the link between true measures of performance of emergency systems and the quantities which can be studied using analytical models, it is now apparent that many models are sufficiently developed to be of great assistance to agency administrators when carefully used. Many of the research goals for allocation of police patrol forces proposed in 1968 in a study for the Department of Justice [6, p. 168] are now being approached, if not achieved. Wide interest is now apparent, as illustrated by reported applications of quantitative techniques in police departments in Boston [42], New York [64], St. Louis [48], Chicago [18], Cleveland [28], Tucson, Arizona [3], Phoenix, Arizona [15], and Great Britain [9]. The whole subject of the allocation of fire units has been developed in the past two or three years and has given an entirely new complexion to fire research. In the next few years we expect that the models will improve in their sophistication and utility and agency administrators will make increasing use of quantitative models as their virtues become apparent.

ACKNOWLEDGMENTS

We wish to thank our colleagues at the New York City-Rand Institute and members of the Boston Police Department and the New York City Police and Fire Departments for their assistance to us during our discussions of the topics in this report. This work was supported in part by the U. S. Department of Housing and Urban Development under Grant H-1056 and in part by the National Science Foundation under Grants GK-16471 and GI-5.

1. Congested radio frequencies often prohibit this type of location procedure.
2. This pattern is also common for emergency repair services, in which the driver may contact the dispatcher at the end of one service to determine where he should go next.
3. Some cities maintain records of false alarm rates for each fire alarm box and adjust their initial response. These typically vary from 20 to 90 per cent false.
4. Some requests for service are generated by field personnel, as when a fire chief signals a second alarm at a fire, or a patrolman calls for assistance over his radio. Such calls can be immediately identified as reliable and of high priority.
5. Typically, a little more than 5 man-years are required annually to fill a single post around the clock. Thus, the direct manpower costs of operating a two-man patrol car in New York are approximately \$120,000 per year, and a single fire engine may cost over \$500,000 per year.
6. Except if the number of emergency vehicles is infinite - or, as a practical approximation, much too large for the budgets of most cities.
7. The results of this queuing calculation are not the sole basis on which cars are assigned to districts, since St. Louis also has a "preventive patrol" force which does not respond to calls, unless they have a very high priority. However, the use of queuing theory is an essential component of the allocation policy of the St. Louis Police Department.
8. The dependence of service times on the number of busy units is characteristic of most urban emergency systems, but it is difficult to measure quantitatively. One cause of the variation, which can be estimated at least roughly [44], is the increase in average travel time which occurs when distant units must be dispatched to calls. More important, however, is the fact that an incident may escalate when units do not arrive promptly; a small fire may become much larger and require a longer time to extinguish, or a reported marital dispute may result in serious assault before a patrol car arrives. Available data are rarely adequate to model these phenomena [34].
9. Some units may leave the fire scene when the fire is under control, while others will remain until extinguishment, and still others will continue to work after extinguishment on some duties known as "overhaul."

10. See Section IV. A.
11. See Section IV. A.
12. For a general discussion of criteria which are appropriate for police patrol, see the report of Kakalik and Wildhorn [36].
13. More precisely: among all quadrilaterals of a given area, the one with the smallest maximum distance between two points is a square.
14. Several interesting applications of this "right-angle distance metric" and other metrics have been discussed in References 26 and 57.
15. For instance, even for those calls which are related to crimes, typically 75 per cent are "nonemergency" and thus do not require rapid response. [62, p. 91]
16. "Tour" refers to the period of time during which a specific group of men will be working.
17. In this simple example we assume that calls arriving when both units are busy are handled by units outside the two-sector region.
18. Similar results hold if response time rather than response distance is used.
19. The absence of global repositioning as a standard technique in police patrol operation may be explained by the fact that an accumulation of small incidents, rather than a single large incident, is most often the cause of whatever unavailabilities exist. Even fire departments are not likely to provide relocation guidelines for dispatchers to use in cases when several small fires produce as many vacant fire houses as a large fire might.
20. The model can, however, be used to solve the simpler problem of determining where to reposition n units (fewer than the number of houses) in order to minimize expected response time when all n units are available.
21. The remainder of the time is spent answering calls and performing other duties.
22. This might be bothersome if one considers s to be a controllable parameter. But usually s is in the range of 5-15 mph and cannot be readjusted at will.

23. These statements are subject to the assumptions of the models used, the most critical of which is the assumption that the cars patrol independently of each other.

REFERENCES

1. Adams, R., and S. Barnard, "Simulation of Police Field Forces for Decision-Making in Resource Allocation," Law Enforcement Science and Technology III, IIT Research Institute, Port City Press, 1970.
2. American Insurance Association, Standard Schedule for Grading Cities and Towns of the United States with Reference to their Fire Defenses and Physical Conditions, New York, 1956.
3. Bergstrom, K., "The Fluid Patrol System," Traffic Police Administration Training Program, Traffic Institute, Northwestern University, 1966.
4. Blum, E., The Deployment of Municipal Emergency Services: Overview and Summary of Results, New York City-Rand Institute, R-685-HUD, 1971.
5. Blumstein, A., and R. Larson, "A Systems Approach to the Study of Crime and Criminal Justice," Operations Research for Public Systems, P. M. Morse and L. W. Bacon, eds., M.I.T. Press, Cambridge, Massachusetts, 1967.
6. Blumstein, A., et al., "A National Program of Research, Development, Test, and Evaluation on Law Enforcement and Criminal Justice," Institute for Defense Analysis, Arlington, Virginia, 1968.
7. Bottoms, A., E. Nilsson, and D. Olson, "Third Quarterly Progress Report, Operations Research Task Force," Chicago Police Department, 1968.
8. Brownstein, S. H., "Some Concepts and Techniques for Constructing and Using a Geographically-Oriented Urban Data Base," Socioeconomic Planning Sciences, Vol. 1, (1968), p. 309.
9. Bryant, J. W., M. L. Chambers, and D. Falcon, "Patrol Effectiveness and Patrol Deployment," Report on Home Office Project in Lancaster Division of Lancashire Constabulary, University of Lancaster, Department of Operational Research, 1968.
10. Business Week, "Electronic Eye on Auto Fleets," No. 2107, (January 17, 1970), p. 88.
11. Capaul, J., N. Heller, and E. Meisenheimer, "A Stochastic Model for Allocating Police Patrol Units to Districts Which Reflects Each District's Rates of Injury, Property Loss, and Fear," Paper presented at the 38th National Meeting of Operations Research Society of America, 1970.
12. Carter, G., and E. Ignall, "A Simulation Model of Fire Department Operations: Design and Preliminary Results," IEEE Transactions on Systems Science and Cybernetics, SSC-6, (1970), p. 282.
13. Carter, G., and E. Ignall, "Predicting the Actual Number of Fire-Fighting Units Dispatched," New York City-Rand Institute (Unpublished Mimeograph).

4. Carter, G., J. Chaiken, and E. Ignall, Response Areas for Two Emergency Units, New York City-Rand Institute, R-532-NYC/HUD, March 1971.
5. Casey, P., "Determining Police Patrol Car Requirements by Computer Simulation," M.S. Thesis, Arizona State University, 1968.
6. Chaiken, J., Estimating Numbers of Fire Engines Needed in New York City Fire Divisions, New York City-Rand Institute, R-508-NYC, 1971.
7. Chaiken, J., Number of Emergency Units Busy at Alarms Which Require Multiple Servers, New York City-Rand Institute, R-531-NYC/HUD, March 1971.
8. Chicago Police Department, Operations Research Task Force, Quarterly Progress Reports, (A. Bottoms, Project Director), 1968, 1969.
9. Chicago Transit Authority, Progress Reports, Chicago, Illinois, 1969, 1970.
10. Cobham, A., "Priority Assignment in Waiting Line Problems," Operations Research, Vol. 2, (1954), p. 70.
11. Crime Control Digest, Detroit Police Department, September 11, 1968.
12. Crime Control Digest, St. Louis Police Department, January 14, 1969.
13. Distler, R., "An Active Infrared System for Vehicle Tracking and Ranging," Law Enforcement Science and Technology I, Proceedings of the First National Symposium on Law Enforcement Science and Technology, Thompson Book, Publisher, 1967, pp. 913-918.
24. Edie, L. C., "Traffic Delays at Toll Booths," Operations Research, Vol. 2, (1954), p. 107.
25. Elliott, J., "Random Patrol," Law Enforcement Science and Technology II, Proceedings of the Second National Symposium on Law Enforcement Science and Technology, IIT Research Institute, Port City Press, 1969, pp. 557-560.
26. Fairthorne, D., "The Distances Between Pairs of Points in Towns of Simple Geometrical Shapes," Proceedings of the Second International Symposium on the Theory of Traffic Flow, London, 1963, published by Organization for a European Cooperation and Development, 1965, pp. 391-406.
27. Fisher, I., and J. Midler, Analyzing Municipal Fire Protection: A Theoretical Framework, New York City-Rand Institute, R-599-NYC, 1971.
28. Gass, S., "On the Division of Police Districts into Patrol Beats," Paper presented at National Conference of the Association for Computing Machinery, New York, 1968.

29. Heller, N., "Proportional Rotating Schedules," Ph.D. Thesis, University of Pennsylvania, 1969.
30. Heubrocker, H., "Feasibility of Radar Beacon Techniques for Patrol Vehicle Tracking and Position Display," Law Enforcement Science and Technology I, Proceedings of the First National Symposium on Law Enforcement Science and Technology, Thompson Book, Publisher, 1967, pp. 903-912.
31. Hogg, J. M., "Planning for Fire Stations in Glasgow in 1980," Home Office (U.K.), Scientific Advisory Branch, Report No. 1/68, October 1968.
32. Institute for Public Administration and Teknekron, "An Analytic and Experimental Evaluation of Alternative Methods for Automatic Vehicle Monitoring," 1968.
33. International City Managers' Association, Municipal Police Administration, Fifth ed., Chicago, 1961.
34. Isaacs, H., "A Study of Communications, Crimes, and Arrests in a Metropolitan Police Department," Task Force Report: Science and Technology, A Report to the President's Commission on Law Enforcement and Administration of Justice, U.S. Government Printing Office, Washington, D.C., 1967.
35. Jacobsen, J., "Geographic Retrieval Techniques," Proceedings of the IBM Information Systems Symposium, Washington, D.C., 1968, pp. 121-124.
36. Kakalik, J., and S. Wildhorn, Aids to Decisionmaking in Police Patrol, The Rand Corporation, R-593/4-HUD/RC, February 1971.
37. Keeney, R., "Preferences for the Response Times of Engines and Ladders to Fires," New York City-Rand Institute, to appear.
38. Knickel, E., "Electronics Equipment Associated with the Police Car," Task Force Report: Science and Technology, A Report to the President's Commission on Law Enforcement and Administration of Justice, Appendix E, U.S. Government Printing Office, Washington, D.C., 1967.
39. Kolesar, P., and W. Walker, "A Relocation Algorithm for the MICS," New York City-Rand Institute (Unpublished Mimeograph).
40. Kolesar, P., "Inverse Square Root Laws for Fire Engine Effectiveness," New York City-Rand Institute (Unpublished Mimeograph).
41. Koopman, B., "The Theory of Search III, The Optimum Distribution of Searching Effort," Operations Research, Vol. 5, (1957), p. 613.
42. Larson, R., "Hourly Allocation of Complaint Clerks, Dispatchers, and Radio-Dispatchable Patrol Personnel," Proceedings of Second National Symposium on Law Enforcement Science and Technology, IIT Research Institute, Port City Press, 1968.

43. Larson, R., "Models for the Allocation of Urban Police Patrol Forces," M.I.T. Operations Research Center Technical Report No. 44, 1969.
44. Larson, R., Measuring the Response Patterns of New York City Police Patrol Cars, New York City-Rand Institute, R-673-NYC/HUD, April 1971.
45. Larson R., Response of Emergency Units: The Effects of Barriers, Discrete Streets, and One-Way Streets, New York City-Rand Institute, R-675-HUD, April 1971.
46. Larson, R., and K. Stevenson, On Insensitivities in Urban Redistricting and Facility Location, New York City-Rand Institute, R-533-NYC/HUD, March 1971.
47. McCormack, R. J., Jr., and J. L. Moen, "San Francisco's Mission Police District: A Study of Resource Allocation," Center for Planning and Development Research, Institute of Urban and Regional Development, University of California, Berkeley, 1968.
48. McEwen, T., Project Director, "Allocation of Patrol Manpower Resources in the Saint Louis Police Department," Vols. I, II, 1966.
49. Morse, P. M., Queues, Inventories, and Maintenance, John Wiley, New York, 1958.
50. Parzen, E., Stochastic Processes, Holden-Day, San Francisco, 1962.
51. Prabhu, N., Queues and Inventories, A Study of Their Basic Stochastic Processes, John Wiley, New York, 1965.
52. The President's Commission on Law Enforcement and Administration of Justice, The Challenge of Crime in a Free Society, U.S. Government Printing Office, Washington, D.C., 1967, pp. 20-22.
53. ReVelle, C., D. Marks, and J. Liebman, "An Analysis of Private and Public Sector Location Models," Management Science: Theory, Vol. 16, (1970), p. 692.
54. Rosenshine, M., "Contributions to a Theory of Patrol Scheduling," Operational Research Quarterly, Vol. 21, (1970), p. 99.
55. Rypinski, C., "Police Message, Status and Location Reporting System," Law Enforcement Science and Technology III, IIT Research Institute, Port City Press, 1970.
56. Savas, E. S., "Simulation and Cost-Effectiveness Analysis of New York's Emergency Ambulance Service," Management Science, Vol. 15, (1969), p. B-608.
57. Smeed, R. J., and G. O. Jeffcoate, "Traffic Flow During the Journey to Work in the Central Area of a Town Which has a Rectangular Grid for its Road System," Proceedings of the Second International Symposium on the Theory of Traffic Flow, London, 1963, Published by Organization for a European Cooperation and Development, 1965, pp. 369-390.

58. Smith, R. D., "Random Patrol," Field Service Division of the International Association of Chiefs of Police, Washington, D.C., 1970.
59. Stevenson, K., "Operational Aspects of Emergency Ambulance Services," M.S. Thesis, M.I.T., 1970.
60. Swersey, A., "Dispatching, Deployment and Relocation of Fire Engines," Paper presented at the 37th National Meeting of Operations Research Society of America, Washington, D.C., 1970.
61. Swersey, A., "A Mathematical Formulation of the Fire Engine Relocation Problem," New York City-Rand Institute (Unpublished Mimeograph).
62. Task Force Report: Science and Technology, A Report to the President's Commission on Law Enforcement and Administration of Justice, U.S. Government Printing Office, Washington, D.C., 1967.
63. Walker, W., and S. Shinmar, "Approaches to the Solution of the Fire Engine Relocation Problem," New York City-Rand Institute (Unpublished Mimeograph).
64. Wildhorn, S., Public Order Studies in New York City, The Rand Corporation, P-4250, November 1969.
65. Wilson, O. W., "Distribution of the Police Patrol Force," Publication 74, Public Administration Service, Chicago, Illinois, 1941.
66. Wilson, O. W., Police Planning, Charles C. Thomas, Publisher, Springfield, Illinois, 2nd ed., 1958.
67. Wilson, O. W., Police Administration, 2nd ed., New York, McGraw-Hill, 1963.
68. Zauderer, J., "Field Testing of an Automatic Vehicle Monitoring System," Law Enforcement Science and Technology III, IIT Research Institute, Port City Press, 1970.