Chapter 4
Probabilistic Coverage

4.1 Introduction

Much of the book is focused on facilities of various types, represented as points, nodes, lines, arcs, paths, tours, areas, etc., providing a wide range of services. The underlying assumption has generally been that facilities, or personnel at/from the facility, are available to serve when needed. This is not too surprising because a sizable number of location covering models are firmly rooted in the initial work Toregas et al. (1971) on the location set covering problem (LSCP), where they were specifically interested in public sector issues involving equity of access to service. In the LSCP facilities were viewed as available for service when needed. In particular, the application of the LSCP to site emergency services, like fire, ambulance and police response, helped to design and relocate such services so that they provide coverage to all. Little has changed, in fact, over the intervening years as emergency service contexts remain of great interest and coverage models have time and again been instrumental in helping to both understand existing service systems as well as develop management plans for emergency response while promoting fairness and equity in service access. Of course, there are many other areas of application for coverage models as well, but the emergency response context has continued to be both challenging and interesting as we better understand such systems and have better supporting data. The focus of this chapter involves the fact that facilities (or personnel) may not always be available when needed. That is, there is a non-zero probability that facility service coverage may not be provided even when every demand is within a desired maximal service standard of a facility. There are clearly many ways in which a facility would be unavailable for service. One situation is that personnel are already busy serving another demand. This is depicted in Fig. 4.1, where the fire engine has traveled from the fire station in response to a fire. However, while busy fighting this fire, another incident (vehicle crash and fire) has occurred across town. It is therefore not possible for a fire crew to respond immediately.

R. L. Church, A. Murray, *Location Covering Models*, Advances in Spatial Science,
https://doi.org/10.1007/978-3-319-99846-6_4

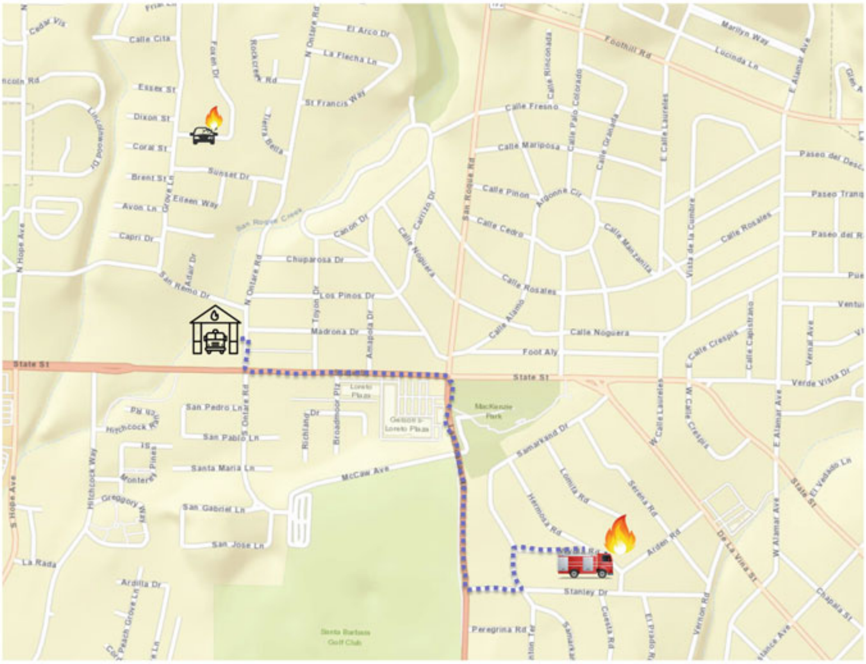

Fig. 4.1 Depiction of a facility busy serving demand

Another situation is that a facility may be unavailable due to a failure of some sort, such as equipment being broken, a power outage, a flood, or even an accident.

If one considers emergency services such as fire, ambulance/paramedic or police, as an example, then what is desired is speedy response to a call for service. As noted previously in this book, such response should be within a maximum service standard, S, in order to ensure safety, survivability, etc. As a result, the LSCP and MCLP detailed in Chap. 2 have been important modeling approaches in this area. However, there is an assumption that each facility is always available to provide service coverage. Such an assumption seems reasonable when considering fire stations and associated equipment, since they are usually idle and awaiting a call. Therefore, the probability of having two simultaneous nearby service calls is very small. Nevertheless, the time to serve calls can be substantial. Not taking this into account might result in the inability of a system of facilities to suitably respond to demand within the desired standard because nearby personnel are occupied with a previous service request. Accordingly, it might be critically important to take into account the probability or likelihood of facility unavailability, particularly given that the associated impacts on system service coverage may be significant.

Similar implications arise in the case of facility failure. Consider the case of cellular phone service. For a cell phone to work and be useful, it must be within the communication standard, S, of a service tower. However, proximity alone is not sufficient. The equipment at the service tower must be operable. While this equipment seems fairly reliable, it does fail, as power may be interrupted, infrastructure deteriorates, etc. As a result, system coverage is likely to be impacted at some point in time.

Several methods have been developed to take into account facility availability in deriving effective service coverage planning. These include the Probabilistic Location Set Covering Problem, the Maximal Expected Coverage Location Problem, the Maximum Availability Location Problem, and others. All of these models represent an attempt to address the availability of facilities in providing coverage. Simply put, response or service within the standard S may not always be possible, even though it is an underlying fundamental goal. Thus, from a probabilistic perspective one might expect that service of demand would most often be accomplished within the desired maximum service standard. Taking this a step further, we can establish a threshold under which "most often" can be stipulated. For example, we may want demand to be served within the S standard 85% of the time. Required standards along these lines are often codified in laws, regulations and guidelines. One example is that of the National Fire Protection Association in recommending that suppression resources be capable of arrival within 4 min travel time to 90% of the incidents (see Murray 2015). Many cities, counties and other local agencies have standards along these lines as well. Further, such standards are not limited to fire response. The California State Emergency Medical Service Authority, for example, promotes related standards for a range of EMS response categories, including Basic Life Support/CPR, defibrillation, Advanced Life Support, ambulance transport, etc. Irrespective of context, the stipulated threshold becomes another operational parameter. The literature often refers to this using the symbol α. In the above example, α would equal 85%.

4.2 Reliable Coverage

The first approach to take into account facility availability in coverage modeling was the Probabilistic Location Set Covering Problem (PLSCP) developed by Chapman and White (1974). The PLSCP essentially extends the LSCP by defining suitable coverage in terms of temporal availability as well as geographical proximity. That is, whether a demand location receives effective service is not solely dependent upon a site being within the service standard S of a facility, but also that this facility is available to serve. A critical question then is what is the likelihood of a facility being busy or otherwise unavailable? Suppose that we know this probability from historical records, quality guarantees associated with equipment, etc. If the likelihood of a facility being unavailable is q, then we can characterize the probability that at least one facility is within S and is not busy serving another demand. Further, ensuring that at least one facility is available for service to a given demand most of the time will likely require that that demand will have to be suitably covered by multiple facilities.

The notion that effective service should be at least some minimum threshold, α, in a probabilistic sense has been termed "α-reliable coverage". The reason for this is because the probability of effective coverage (facility availability within the service standard) must be at least α, where α is a fraction between zero and one. The PLSCP specifically seeks the minimum number and locations of facilities such that every demand receives α-reliable coverage.

Consider the following notation:

i = index representing demand for service
j = index representing potential facility sites
q = probability of a facility being unavailable or busy
d_{ij} = shortest distance or travel time from demand i to potential facility site j
S = the desired service standard (e.g. distance or travel time)
α = pre-specified level of reliability in service coverage

$$N_i = \{ j | d_{ij} \leq S \}$$

$$x_j = \begin{cases} 1, & \text{if a facility is located at site } j \\ 0, & \text{otherwise} \end{cases}$$

The PLSCP is based upon the assumption that all facilities are independent and that a facility being unavailable is independent of the availability of any other facility in the system. As before, suppose that we are able to derive or that we know from experience the probability that any given facility is unavailable, q. With this probability, we can estimate for any given demand location the probability that a suitable or timely response within the standard S is possible, based upon the number of facilities capable of serving this demand within the standard. For example, if there is only one facility that can cover a given demand i then that demand would be served with a probability of $1 - q$. Suppose that there are three sited facilities, each within the service standard S for a specific demand i. The probability that all three facilities are unavailable (or busy) is $q \times q \times q$, or q^3. This would imply that the probability of coverage would therefore be $1 - q^3$. If this probability is greater than the establish reliability threshold, α, then we could consider demand i sufficiently served, taking into account the possibility that one or more facilities may be unavailable for service. It is worth noting that q may be viewed as a system-wide average unavailability measure, applying equally to each facility or potential facility location.

For any configuration of facilities, $\sum_{j \in N_i} x_j$ represents the number of facilities that have been located within the coverage standard of demand i. One can then derive the probability that demand i will receive service coverage when needed as:

$$1 - q^{\sum_{j \in N_i} x_j} \tag{4.1}$$

Therefore, equation (4.1) would reflect the probability of suitable coverage within the service standard S. The stipulation then is that this should be greater than or equal to α:

$$1 - q^{\sum_{j \in N_i} x_j} \geq \alpha_i \tag{4.2}$$

Note that α is defined here for each demand i, using α_i. This was the original mathematical specification in Chapman and White (1974). In application, however, they simply utilized α. This may have been for convenience, or perhaps due to the fact that justification for varying levels of reliability among demand may be problematic in public sector siting contexts.

Given that the probability of each facility being unavailable is the same, q, equation (4.2) can be rewritten in a linear form. Through algebraic manipulation we have:

$$1 - \alpha_i \geq q^{\sum_{j \in N_i} x_j} \tag{4.3}$$

Taking the log of each side of the inequality in equation (4.3) yields:

$$\log(1 - \alpha_i) \geq \log q^{\sum_{j \in N_i} x_j} \tag{4.4}$$

This then simplifies to:

$$\log(1 - \alpha_i) \geq \sum_{j \in N_i} x_j \log q \tag{4.5}$$

Finally, recognizing that log q is less than zero because q is assumed to be less than one, then simplification reverses the inequality in equation (4.5) to give:

$$\frac{\log(1 - \alpha_i)}{\log q} \leq \sum_{j \in N_i} x_j \tag{4.6}$$

Therefore, equation (4.6) denotes that the number of facilities that serve a demand area i (right hand side of inequality) must be greater than or equal to the quantity on the left hand side of the inequality in order to satisfy the stipulated level of reliability, α_i, given the probability of each facility being unavailable, q. Empirical evaluation may be useful in better clarifying what is happening here. Consider a reliability standard of $\alpha_i = 0.85$ and a probability of unavailability of $q = 0.25$. The interpretation of this is that demand i is expected to see response for service within the standard S at least 85% of the time. However, facilities may be busy some 25% of the time, serving other calls for service. Equation (4.6) in this case gives a left hand side value of 1.368. Thus, two or more facilities would clearly be necessary to provide reliable coverage of 85% or better.

Chapman and White (1974) structured the following model, incorporating this condition:

$$\text{PLSCP}: \quad \textit{Minimize} \ \sum_j x_j \tag{4.7}$$

Subject to:

$$\sum_{j \in N_i} x_j \geq \left\lceil \frac{\log(1-\alpha_i)}{\log q} \right\rceil \quad \forall i \tag{4.8}$$

$$x_j = \{0, 1\} \quad \forall j \tag{4.9}$$

The objective, (4.7), seeks the smallest number of facilities. This is precisely the same objective as the LSCP. Constraints (4.8) specify that suitable coverage must be provided to each demand at the indicated level α_i, where the operator function $\lceil \ \rceil$ represents the smallest integer greater than or equal to the enclosed value. Constraints (4.9) impose binary integer requirements on the decision variables.

The main constraint of the PLSCP is to ensure that the probability of an available facility being within the S standard for each demand i is greater than or equal to α_i. The right hand side quantity is nothing other than a coefficient, derived as a function of α_i (or α) and q, inputs that are known in advance. Assuming that each facility has the same unavailability probability (or busyness level) yields a problem where each demand location needs to be covered at least some number of times. Specifically, one can view constraints (4.8) as follows:

$$\sum_{j \in N_i} x_j \geq \gamma_i \quad \forall i \tag{4.10}$$

where $\gamma_i = \left\lceil \frac{\log(1-\alpha_i)}{\log q} \right\rceil$. That is, γ_i is the minimum number of facilities necessary to ensure coverage at the reliability level of α_i. Constraints (4.10) represent processed right hand side values of constraints (4.8), something that would be necessary to compute prior to solving a problem instance. One should recognize that the value of γ_i, strictly speaking, does not need to be the smallest integer value that exceeds the right hand side of constraint (4.8). Since the left hand side of equations (4.8) and (4.10) are sums of decision variable values, all required to be integer, the values of the left hand sides of conditions (4.8) and (4.10) will be integer in value. Thus, there is no additional requirement being imposed when we define each γ_i as being an integer-valued parameter. Additionally, there may be some benefit to integer valued model coefficients when applying solution techniques like branch and bound.

Chapman and White (1974) also discussed a version of the PLSCP where the unavailability of each facility varies. They did not, however, specify any technique to be used for the derivation of q or for the unavailability of specific facilities. Formalizing this is possible, but requires some modified notation:

q_i = probability of a facility being unavailable or busy in demand area i
γ_i' = minimum number of facilities necessary to ensure coverage at α_i given local busyness q_i

Based on our previous derivation, the minimum number of necessary facilities for each demand area is then calculated as:

$$\gamma_i' = \left\lceil \frac{\log(1-\alpha_i)}{\log q_i} \right\rceil \tag{4.11}$$

Constraints (4.8) therefore become the following when a local busyness probability is utilized:

$$\sum_{j \in N_i} x_j \geq \gamma_i' \quad \forall i \tag{4.12}$$

Ball and Lin (1993) suggest this basic constraint form as well based on assumptions of a Poisson distribution for facility unavailability. At any rate, constraints (4.12) suggest two forms of the PLSCP. One is reflected in (4.7)–(4.9) where a system-wide facility unavailability probability is assumed. The other version we can refer to as PLSCP′ and consists of objective (4.7) and constraints (4.9) and (4.12). PLSCP′ incorporates a local probability of a facility being unavailable. Irrespective of the model formulation, important issues are input parameters, namely α, α_i, q and/or q_i.

Upon close examination, one should recognize the equivalence of the PLSCP to the multi-level covering problem discussed in Toregas (1970) and Church and Gerrard (2003), and detailed in Chap. 3. The multi-level covering problem can be structured for both the LSCP and the MCLP, where there is a desire to cover demands with more than one facility. For the LSCP, the multi-level covering problem entails minimizing the number of facilities needed to cover each demand a desired number of times. What is important here is that there is a probabilistic derivation for the values of γ_i' and γ_i. Further, the resulting model is an equivalent deterministic model. Although, this type of problem is often solved by the use of integer linear programming software, it can be solved partially or even completely by the method of reductions [see Toregas and ReVelle (1973) for single level and Church and Gerrard (2003) for multi-level reduction algorithms].

4.3 Expected Coverage

Based on the discussion and models detailed in previous chapters, a natural extension of the PLSCP is to address the basic premise of the MCLP, where there are limited resources in siting facilities thereby preventing coverage of all demand within the desired service standard. Daskin (1982, 1983) incorporated the probability of a facility being unavailable, q, in creating a stochastic form of the MCLP. This model has been referred to as the Maximal Expected Covering Problem (MEXCLP), taking into account the probabilistic nature of facility availability.

Differing from the PLCSP, however, Daskin (1982, 1983) counted any level of coverage as providing some benefit, rather than requiring a minimum coverage reliability threshold. This was done by viewing availability as a binomial probability function. Consider the following additional notation:

k = index corresponding to the number of facilities sites
a_i = demand for service at location i
p = number of facilities to be sited

$$y_{ik} = \begin{cases} 1, \text{if demand } i \text{ is covered by at least } k \text{ facilities} \\ 0, \text{otherwise} \end{cases}$$

What is unique is that the y_{ik} variables are introduced to track the number of facilities capable of covering an individual demand i. Assume as before that each facility is unavailable some portion of the time, q. Thus, a demand site that is covered exactly once will be served $1 - q$ fraction of the time. We can represent this level of coverage by multiplying this fraction of coverage by the population at that site as the term $a_i(1 - q)y_{i1}$, where y_{i1} equals one in value if that demand is covered once. Technically, this is also equivalent to $a_i(1 - q)q^{k-1}y_{i1}$ as the term q^{k-1} equals 1 when $k = 1$. When two facilities cover demand i, then the probability of coverage is $1 - q^2$. The added fraction of coverage in increasing coverage from one facility to two facilities is the difference between $1 - q^2$ and $1 - q$, which is $(1 - q^2) - (1 - q) = q - q^2$. Thus, the added coverage is equivalent to $(1 - q)q^1$ or $(1 - q)q^{k-1}$ when $k = 2$. In general, the added probability of coverage for k facilities is $(1 - q)q^{k-1}$. Thus, the total increase in coverage for demand i when adding a kth facility that covers demand i is $a_i(1 - q)q^{k-1}$. Therefore, we can account for added levels of expected coverage. Based upon this, we can formulate the Maximal Expected Coverage Location Problem (MEXCLP) as follows:

$$\text{MEXCLP}: \quad \textit{Maximize} \sum_i \sum_{k=1}^{p} a_i(1-q)q^{k-1}y_{ik} \tag{4.13}$$

Subject to:

$$\sum_{j \in N_i} x_j - \sum_{k=1}^{p} y_{ik} \geq 0 \quad \forall i \tag{4.14}$$

$$\sum_{j} x_j = p \tag{4.15}$$

$$x_j = \{0, 1\} \quad \forall j \tag{4.16}$$

$$y_{ik} = \{0, 1\} \quad \forall i, k \tag{4.17}$$

The objective (4.13) seeks to maximize expected coverage by taking into account the probability that each facility may be unavailable to serve a certain portion of the time. Constraints (4.14) account for the number of facilities that are capable of providing coverage to a particular demand. If demand i is covered only once then constraint (4.14) allows the sum of $\sum_{k=1}^{p} y_{ik}$ to be one in value. But since the additions to expected coverage in the objective function decrease as the number of units covering a demand increase, y_{i1} will then equal 1 in value. For this same reason, if $\sum_{k=1}^{p} y_{ik}$ equals two in value, both y_{i1} and y_{i2} will be one in value. In general, $y_{ik-1} \geq y_{ik}$. Consequently, the objective function combined with constraints (4.14) will accurately account for the expected coverage that is provided by the facilities that have been located. Constraint (4.15) limits the number of facilities to be located to exactly p facilities, which is a specific form of a budget constraint. Constraints (4.16) and (4.17) impose binary integer restrictions on the decision variables. Technically speaking, the integer restrictions on the y_{ik} variables are not necessary as long as these variables are constrained to be no larger than 1 in value. When all of the x_j are integer in value, so too will the y_{ik} variables be binary at optimality.

This formulation differs from the original version presented in Daskin (1982, 1983). It is quite possible in some circumstances that multiple facilities could be located/allocated to a site (called co-location). This may be necessary to ensure satisfactory response requirements. For example, one could house multiple ambulances at an EMS facility, or multiple response teams at a fire station. Accordingly, it may be necessary to allow the facility siting variable x_j to take on general, positive integer values in some situations. As the discussion in this chapter has focused on simply a facility, a decision was made to limit variables to binary. Extension along the lines originally presented in Daskin (1982, 1983) is a simple modification of the MEXCLP formulation presented here.

Suggested in the previous section is that the facility unavailability measure q may be known or unknown. If it is unknown, which would generally be expected for most service systems, then some sort of estimation is necessary. Daskin (1982) and

ReVelle and Hogan (1988) suggested the following basic system-wide busyness fraction estimate:

$$q = \bar{t} \sum_i a_i \Big/ T\,p \tag{4.18}$$

where $\bar{t}$ is the average time to complete service for a demand; and, T is the total availability possible (e.g., 24 h in a day) for a facility. With this, or any potential variation, it is possible to calculate the added amount of availability for the k^{th} added facility using the derivation above as:

$$\beta_k = (1 - q)q^{k-1} \tag{4.19}$$

With this term, we can re-write objective (4.13) as follows:

$$Maximize \sum_i \sum_{k=1}^{p} a_i \beta_k y_{ik} \tag{4.20}$$

The added feature is that β_k may be defined in different ways, depending on the probabilistic function deemed most suitable for a particular application context. In this particular case, Daskin (1982, 1983) assumed a binomial function. Of course, any changes in attribute coefficients in this model (or others) may impact results, changing the configuration of facility sites and associated coverage in an optimal solution.

While objective (4.20), and objective (4.13), specifies one form of criteria to be optimized, other objectives may be possible. Perhaps there is a preference to move away from a system wide measure of unavailability. If so, then one might view that there could be a local area specific measure of busyness. ReVelle and Hogan (1988) suggested a local busyness fraction, q_{ik}, defined as a function of the number of facilities that could serve a particular demand. This local busyness measure is:

$$q_{ik} = \bar{t} \sum_{i' \in N_i} a_{i'} \Big/ T\,k \tag{4.21}$$

With this local measure, facility availability becomes more location specific, and can be summarized as follows:

$$\beta_{ik} = (1 - q_{ik})q_{ik}^{k-1} \tag{4.22}$$

This gives rise to an alternative objective, replacing objective (4.20) with:

$$\text{MEXCLP}' : \quad \textit{Maximize} \ \sum_i \sum_{k=1}^{p} a_i \beta_{ik} y_{ik} \tag{4.23}$$

This objective combined with constraints (4.14)–(4.17) defines MEXCLP′, specifically accounting for and incorporating a local busyness measure. Such a model variant was, in fact, detailed in Sorensen and Church (2010).

While it is possible to present a formulation of the MEXCLP in a number of ways, there are two important distinctions to keep in mind. First, the unavailability or busyness of a facility may be viewed (and modeled) from a system-wide perspective, assuming that the probability of a facility being busy and unavailable for service is the same across a region, or it may be conceived that facility service loads are heterogeneous due to variations in demand across a region. Accordingly, there are two different objective functions of the MEXCLP that have been adopted, one generalized as (4.20) representing system-wide busyness and the other as (4.23) that takes into account a local busyness estimate, MEXCLP′. Second, facility siting may permit only one service entity at a location, or it may allow the co-location of facilities. Again, the formulation of the MEXCLP presented here assumes that co-location is not permitted. However, modification of this formulation is possible to allow for such a situation.

4.4 Maximal Reliable Coverage

The PLSCP and MEXCLP have been important formulations in the progression of developing capabilities that address issues of facilities being busy and unavailable when attempting to provide coverage. The PLSCP represents an attempt to ensure that a sufficient number of facilities provide coverage to each demand, doing so such that the probability of having a facility available when needed is at least α (or α_i if more specific detail is needed). The MEXCLP attempts to maximize the total amount of demand that is likely to receive service coverage without any delay. While a distinction can be made that the PLSCP extends the LSCP and the MEXCLP extends the MCLP, there are fundamental differences in the basic approaches taken. Accordingly, there is potential for a more direct extension of the PLSCP that is perhaps more consistent with the MCLP, but differs from the MEXCLP. Specifically, the MEXCLP does not require coverage of each demand be at least α (or α_i) as is the case with the PLSCP. Rather, the MEXCLP seeks to maximize the expected coverage provided.

To incorporate such a modeling feature ReVelle and Hogan (1989) introduced the Maximal Availability Location Problem (MALP). This model is characterized by an intent to maximize demand that is provided α-reliable coverage. Accomplishing this requires the use of γ_i, introduced in the discussion of the PLSCP. Recall that γ_i was defined as the minimum number of facilities that is necessary for demand i to have coverage reliability of at least α. Rather than require that all demand must be covered with this level of reliability, we can seek to maximize the demand that is provided α reliable coverage while locating p facilities. This of course could be considered to be an extension to the original maximal covering location problem. We can formulate the model of ReVelle and Hogan (1989) as follows:

$$\text{MALP}: \quad \textit{Maximize} \ \sum_i a_i y_{i\gamma_i} \tag{4.24}$$

Subject to:

$$\sum_{j \in N_i} x_j - \sum_{k=1}^{\gamma_i} y_{ik} \geq 0 \quad \forall i \tag{4.25}$$

$$\sum_j x_j = p \tag{4.26}$$

$$y_{ik-1} - y_{ik} \geq 0 \quad \forall i, k \tag{4.27}$$

$$x_j = \{0, 1\} \quad \forall j \tag{4.28}$$

$$y_{ik} = \{0, 1\} \quad \forall i, k \tag{4.29}$$

The objective, (4.24), maximizes the amount of demand that is reliably covered. Constraints (4.25) operate in a similar fashion as constraints (4.14) and help to define the level of coverage that is provided to a given demand i. What is different here is that constraints (4.14) in conjunction with the objective function (4.13) ensure that in the MEXCLP the variable y_{ik} was always less than or equal to the value of $y_{ik\,-\,1}$, without expressly forcing such a condition. Since the objective (4.24) counts only those demands that have received α-reliable coverage (i.e., when $y_{i\gamma_i} = 1$), without expressly forcing $y_{ik\,-\,1} \geq y_{ik}$, constraints (4.25) would allow $y_{i\gamma_i} = 1$ when there is only one facility that has been located that covers demand i. Therefore, constraints (4.27) are included to enforce this condition. That is, coverage at a higher level k cannot be counted unless it is provided at the preceding lower level, $k - 1$. Also note that coverage is counted up to and including the level γ_i with the use of y_{ik} variables and where the value of k is limited to be less than or equal to γ_i for each i. Constraint (4.26) limits the number of facilities sited and constraints (4.28) and (4.29) impose binary integer restrictions on the decision variables. Worth noting is the fact that constraints (4.27) can be added to the PLSCP model as well without any loss of generality. It is possible that these constraints may help to tighten the formulation and reduce solution times.

An important distinction with MALP, similar to MEXCLP, is that it can accommodate either system-wide or local busyness measures. Thus, there is a variant of

MALP using a system-wide busyness but also a version with a local busyness measure, which we can refer to as MALP′.[1]

One thing we have attempted to do in this book is highlight coverage model formulation nuances. In doing this, there have been many instances were alternative mathematical specifications are possible. Associated with MALP, the use of system-wide or local busyness is but one aspect of specification differences. It turns out that an alternative formulation is also possible, and there may be benefits to rely on such an alternative.

Consider the following additional notation:

$$y_i = \begin{cases} 1, \text{if demand } i \text{ covered at or above the reliability standard} \\ 0, \text{otherwise} \end{cases}$$

Given this variable, an alternative MALP formulation is the following:

$$\mathit{Maximize} \sum_i a_i y_i \tag{4.30}$$

Subject to:

$$\sum_{j \in N_i} x_j - \gamma_i y_i \geq 0 \quad \forall i \tag{4.31}$$

$$\sum_j x_j = p \tag{4.32}$$

$$x_j = \{0, 1\} \quad \forall j \tag{4.33}$$

$$y_i = \{0, 1\} \quad \forall i \tag{4.34}$$

The objective, (4.30), maximizes the total amount of demand that is covered with α reliable coverage. Constraints (4.31) specify that the number of facilities that cover demand i must equal or exceed the value of γ_i in order for α reliable coverage to be counted for demand i. Constraint (4.32) limits the number of facilities to equal p. Constraints (4.33) and (4.34) impose binary integer restrictions on the decision variables.

Marianov and ReVelle (1996) as well as Sorensen and Church (2010) relied on this alternative formulation of MALP. There are a number of differences from (4.24) to (4.29). One of the benefits of this alternative is that it requires fewer variables and fewer constraints. This may be appealing for a number of reasons. It is more concise, and depending on the solver may actually lead to faster solution times. However, doing so likely reduces the "integer friendly" properties characteristic of the original

[1] The literature has often called these variants MALP I and MALP II, as introduced in ReVelle and Hogan (1989). Such a distinction is avoided in this chapter because the only difference is a change in model coefficients, with most of the variables and constraint structure remaining the same.

MALP formulation. More traditional approaches for solving integer programs, such as branch and bound without cuts as well as other modern advances for improving performance, may still encounter substantially more computational difficulty.

4.5 Queuing

An obvious concern in addressing a facility being unavailable or busy is a meaningful and accurate characterization of the associated probability of busyness. The models presented in this chapter up to this point are based on simple definitions of q, q_i and q_{ik}. These probabilities are either assumed to be known or are estimated in some fashion. Of course, many have questioned whether assumed or estimated probabilities are accurate, and the very existence of alternative measures and models suggests that enhancements have and continue to be necessary. Murray and Church (1992) examined system-wide and local busyness estimates, comparing measures relied upon in MALP to simulated busyness estimates associated with identified service system configurations. Their findings suggested that reliability estimates are not particularly accurate, and can negatively impact optimal facility configuration identification. Ultimately, the implications are that system performance may not conform to steady state expectations, which means that the reliability conditions within the model may not be achieved in actual application. As a result, these models may inadvertently identify a suboptimal solution, thinking it performs better.

One way that probability estimates have been improved is through the use of queuing models. For example, Batta et al. (1989) examined the MEXCLP and proposed an extension for adjusting expected coverage based on viewing service as a queuing system. Their intent was to address facility independence, varying busyness probabilities, and location influence on busyness probabilities. Goldberg et al. (1990), too, were interested in expected coverage, and relied on queuing and simulation to derive and update probabilities of facility response and service. Improved estimates of q_i values based on a queuing model were also used to enhance the derivation of the γ_i coefficients, the minimum number of needed facilities capable of covering demand i at or above the level α_i, in Marianov and ReVelle (1994) for improving the PLSCP and in Marianov and ReVelle (1996) for improving MALP.

The basic idea is that viewing the facility response process as a queuing system enables performance characteristics in steady state to be derived. In particular, assume that a system with k facilities observes a service arrival that is Poisson distributed at rate λ. Further, assume that service by a facility is exponentially distributed at rate μ. The rate diagram for this associated system is shown in Fig. 4.2 for this associated *M/M/k* queuing system. Such a system can be viewed as an *M/M/k* queuing system (k servers serving a system where arrival rates follow a Poisson process and where service times follow an exponential distribution). It is also often assumed that service calls are lost (no queue) when no facilities are available. For example, when a call for an EMS occurs and all ambulances are busy, then a rescue vehicle or some non-emergency patient transport company is

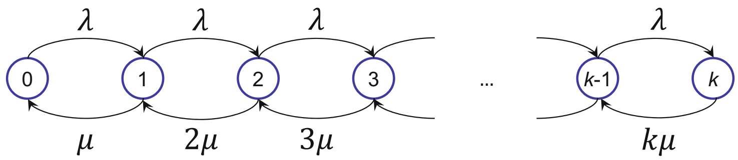

Fig. 4.2 Rate diagram for an assumed *M*/*M*/*k*/*k* queuing system

called to assist. Marianov and ReVelle (1996) used this queuing model to estimate within each local neighborhood of a demand how many servers would be necessary to satisfy the reliability constraint for that neighborhood. To do this, one needs to assume that the queue of one neighborhood is independent of the servers of other neighborhoods or vice versa. Unfortunately, this assumption has not been fully tested although questions have been raised as to the validity of this assumption (see for example Baron et al. 2009). This is essentially the same assumption made by ReVelle and Hogan (1989).

One of the underlying issues in locating facilities that can be congested is to ensure that each facility is not over committed in providing service. Marianov and Serra (1998) suggested a queuing based-form of the maximal covering problem, which explicitly addressed the capacity of individual servers. To formulate their approach we need to define additional notation:

$\lambda_i =$ service (arrival) rate for demand area i
$\mu_j =$ service completion rate for facility j
$c_j =$ coefficient adjusting for service capacity of facility j given reliability level α

$$z_{ij} = \begin{cases} 1, \text{if demand } i \text{ is covered and served by facility } j \\ 0, \text{otherwise} \end{cases}$$

With this notation, Marianov and Serra (1998) formulated a model that accounts for associated facility capacities. Doing this requires the use of allocation variables, z_{ij}, that tracks which demand is served by which facilities in order to ensure that, under steady state conditions, service to local areas of demand can be accomplished within desired α standards. This discrete model (maximal covering location problem with queuing, MCLP-Q) is as follows:

$$\text{MCLP-Q}: \quad \textit{Maximize} \ \sum_i \sum_{j \in N_i} a_i z_{ij} \tag{4.35}$$

Subject to:

$$\sum_{j \in N_i} z_{ij} \leq 1 \quad \forall i \tag{4.36}$$

$$\sum_{j} x_j = p \tag{4.37}$$

$$z_{ij} \leq x_j \quad \forall i, j \in N_i \tag{4.38}$$

$$\sum_{i \in N'_j} \lambda_i z_{ij} \leq c_j \mu_j x_j \quad \forall j \tag{4.39}$$

$$x_j = \{0, 1\} \quad \forall j \tag{4.40}$$

$$z_{ij} = \{0, 1\} \quad \forall i, j \in N_i \tag{4.41}$$

The objective, (4.35), seeks to maximize the demand suitably covered. Constraints (4.36) indicate that each demand is allocated to at most one facility. Constraint (4.37) specifies the number of facilities to be sited. Constraints (4.38) restrict allocations to only those sites that are selected for a facility. Constraints (4.39) limit the arrival rates that are allocated from demands to a given site to be less than the capacity of the server at that site, based upon the service completion rate. The individual facility service capacities are based on a given reliability standard assuming queuing based steady state probabilities. Constraints (4.40) and (4.41) impose binary integer restrictions on decision variables.

While the use of queuing is an improvement in some ways, especially when considering the capacity of the servers as in MCLP-Q, many of the developed models rely on system-wide or local busyness measures. Murray and Church (1992) highlight the ways that this is problematic. Beyond this, there are a number of limitations associated with queuing approaches for supporting coverage modeling.

Often the demand/arrivals for an area i are based upon the assumption that service is provided by the set of facilities covering a demand, $\sum_{j \in N_i} x_j$. Unfortunately this set of facilities may also serve other demands outside the neighborhood N_i. Thus, an $M/M/\sum_{j \in N_i} x_i$ queuing system can overestimate service potential in a given demand area i because on average not all of the sited facilities will be available for potential service to demand i.

Further, questions have been raised about the degree to which a local estimate of busyness meets the assumptions under which the actual busyness estimate was made. For example, if you have a system that can cover all demand with one location and multiple units can be located at that site, then estimating the probability of busyness can be accurately accomplished either probabilistically, through the use of simulation, or by the use of a queuing model that meets the conditions of the system. But, when a system cannot be covered by one central location, then facilities will need to be scattered in order to cover all demands, which is more problematic. However, local busyness estimates can be somewhat accurate, if a local area does not share its servers with other demands outside the local coverage area or when units outside a local coverage neighborhood help to serve demands with that local neighborhood. But, the assumption that calls handled from outside servers to a local neighborhood N_i is balanced by servers within the neighborhood N_i handling calls that occur outside the neighborhood is likely to be violated. In fact, it is easy to overcommit the capacity of servers when service across the boundary of a neighborhood N_i occurs. Consequently, one of the foundational assumptions of the simple local busyness estimates is likely to be violated when call volumes are high and vary over a region.

4.6 Facility Availability

Daskin (1983) suggested that facilities could be unavailable for reasons other than service loads. For example, a facility may have reliability issues, needs regular maintenance (which requires downtime), or be subject to disruption due to infrastructure failure or flooding. Each of these examples involve aspects of a system that are unrelated to the demands that are being served, but where there may be an easily calculated probability of being unavailable for service. If all of the facilities have an equal probability of being unavailable (e.g., q) and each event of a facility losing service capabilities is independent of other service disruption events, then the models like MEXCLP and PLSCP can be used. The key element in the formulation of the MALP relies on the assumption that all facilities in the neighborhood of a demand have an average busyness given this number (i.e., q_{ik}), rather than specific facility availability estimates.

In general, busyness or unavailability of a facility to provide service may vary and not be the same (Lei et al. 2014), or that a facility's capability to provide service degrades as one moves further from that facility (Altınel et al. 2008). Both factors are issues in designing a wireless sensor network. Let's suppose that we are designing a wireless sensor network. There are two main choices in this design: (1) where to place each sensor, and (2) which type of sensor do we use at each chosen location. Let's say that this network is designed to detect an intruder or threat. Each sensor needs to be within communication of at least another sensor and altogether, the network of sensors must be able to communicate with each other (thus, the design of communication between sensors must represent a connected graph). For the purpose here, we will assume that the area in question does not hinder communication between neighboring sensors, given that enough sensors will be needed and are close enough together in order to adequately surveil an area. Furthermore, sensors are not perfect and there is a probability that they cannot detect an intruder. Also, the capability of a given sensor detecting an intruder declines with distance. Consider the following notation:

t = an index used to refer to sensor type, where $t = 1, 2, \ldots, T$
p_{ijt} = the probability of sensing point i with sensor type t located at point j
M_i = the maximum acceptable probability of missing an intruder at point i
c_{jt} = the cost of placing sensor type t at location j

$$x_{jt} = \begin{cases} 1, & \text{if sensor type } t \text{ is placed at location } j \\ 0, & \text{otherwise} \end{cases}$$

The probability of detection is commonly calculated as an exponential decay function, $e^{-\beta_t d}$, where the decay rate, β_t, is associated with the sensor type and d is the distance between the target point and the sensor. Using this function, we can calculate:

$$p_{ijt} = e^{-\beta_t d_{ij}} \tag{4.42}$$

This means that the probability of missing an intruder at point i can be calculated as:

$$\prod_{j \in J} \prod_{t=1}^{T} \left(1 - p_{ijt} x_{jt}\right) \tag{4.43}$$

With this function, we can now formulate the Effective Covering Problem (ECP) of Altınel et al. (2008) for multi-type sensor placement:

$$\text{ECP}: \quad \text{Minimize} \sum_{j \in J} \sum_{t=1}^{T} c_{jt} x_{ijt} \tag{4.44}$$

Subject to:

$$\sum_{t=1}^{T} x_{jt} \leq 1 \quad \forall j \tag{4.45}$$

$$\prod_{j \in J} \prod_{t=1}^{T} \left(1 - p_{ijt} x_{jt}\right) \leq M_i \quad \forall i \tag{4.46}$$

$$x_{jt} = \{0, 1\} \quad \forall j, \forall t \tag{4.47}$$

The objective, (4.44), involves minimizing the cost of the sensors that are used in the network. Note that the costs of sensors varies by type. Constraint (4.45) limits the location of only one sensor type at a given site j. Constraints (4.46) ensure that probability of missing an intruder at a given point i is less than the maximum acceptable probability of missing an intruder at point i. Constraints (4.47) restrict the decision variables to be binary in value. Unfortunately, this formulation is highly non-linear due to constraints (4.46), and not easily solvable by existing software. To address this formulation directly, one would need to design a heuristic.

There is, however, another approach to solving this problem and it involves a transformation. If we take the natural log of expression (4.46) we obtain:

$$\sum_{j} \sum_{t=1}^{T} \ln\left(1 - p_{ijt} x_{jt}\right) \leq \ln\left(M_i\right) \tag{4.48}$$

which is mathematically equivalent to constraint (4.46). Note that when $x_{jt} = 1$, the term associated with this variable in the summation on the left hand side of the inequality equals $\ln(1 - p_{ijt})$. Thus, when $x_{jt} = 1$ the term is equivalent to $\ln(1 - p_{ijt})$ x_{jt}. Also note that when $x_{jt} = 0$, the associated term in the left hand side of the inequality equals $\ln(1 - p_{ijt} \times 0) = \ln(1)$. Since the $\ln(1)$ equals zero, this term is equivalent to $\ln(1 - p_{ijt})x_{ijt}$ when $x_{jt} = 0$. Thus, we can reformulate constraint (4.46) by taking the natural log of each side and using the properties when x_{jt} is binary in value to form the following equivalent constraint:

$$\sum_{j} \sum_{t=1}^{T} \ln\left(1 - p_{ijt}\right) x_{jt} \leq \ln\left(M_i\right) \tag{4.49}$$

The beauty of this constraint is that it is linear. Note that the probabilities p_{ijt} and M_i are less than one in value. The natural logs of these values will be negative. If we define $a_{ijt} = -\ln(1 - p_{ijt})$ and $b_i = -\ln(M_i)$, we can write constraints (4.49) in the following form:

$$\sum_{j} \sum_{t=1}^{T} a_{ijt} x_{jt} \geq b_i \tag{4.50}$$

Altogether, we can now formulate an equivalent form of the ECP as follows:

$$\text{Minimize} \sum_{j \in J} \sum_{t=1}^{T} c_{jt} x_{jt} \tag{4.51}$$

Subject to:

$$\sum_{t=1}^{T} x_{jt} \leq 1 \quad \forall j \tag{4.52}$$

$$\sum_{j} \sum_{t=1}^{T} a_{ijt} x_{jt} \geq b_i \forall i \tag{4.53}$$

$$x_{jt} = \{0, 1\} \quad \forall j, \forall t \tag{4.54}$$

This model is now in the form of a classical multi-level covering problem, except that the contributions to covering a given area i are real values rather than integer values and the demand for coverage, b_i, is a real value rather than integer. Altogether, this is a general form of multi-level location set covering problem presented in Chap. 2. It can also be viewed as a generalization of PLSCP presented in Sect. 4.2.

We can also formulate a generalized form of MALP′, where facility availability varies geographically rather than being the same across a region. To do this, first recognize that we will consider only one type of facility being located, but where the probability of being available is site specific, p_j, or site and distance sensitive, p_{ij} (like in the sensor case). The site and distance sensitive form is appealing because not only can we handle a site specific availability of a facility, but we can also include the probability of reaching a given demand within a standard amount of time. Thus, we can define $a_{ij} = -\ln(1 - p_{ij})$ or $a_{ij} = -\ln(1 - p_j)$. With this, we can formulate a general form MALP′ (MALP′-G) as:

$$\text{MALP}'\text{-G}: \quad \text{Maximize} \sum_i a_i y_i \tag{4.55}$$

Subject to:

$$\sum_{j \in J} x_j = p \quad \forall j \tag{4.56}$$

$$\sum_{j \in N_i} \sum_{t=1}^{T} a_{ij} x_j \geq b_i y_i \forall i \tag{4.57}$$

$$x_j = \{0, 1\} \quad \forall j \tag{4.58}$$

$$y_i = \{0, 1\} \quad \forall i \tag{4.59}$$

where vehicle or facility availability varies geographically and where reaching a given demand from a specific site may also be probabilistic. The objective, (4.55), involves maximizing the demand that is provided α-reliable coverage. Constraint (4.56) represents the constraint on the number of facilities that are being located. Constraints (4.57) define whether α-reliable coverage has been provided to a specific demand i or not. Note that the value of $b_i = -\ln(M_i)$ where $M_i = 1 - \alpha_i$.

The problem becomes a bit more complicated if we want to develop a model to maximize expected coverage of demand, like that of Daskin (1983), when probabilities of facility unavailability vary geographically. MALP′-G is straight forward when we have a standard for whether a demand is covered or not, like an α-reliable covering standard, but if we wish to maximize expected coverage then determining and accounting for the amount of expected coverage to a given demand is considerably more difficult. The reason for this is that with the Daskin (1983) construct we could easily calculate the level of expected coverage for a demand based upon the number of facilities that covered that demand. But, if the probabilities of coverage vary in possibly covering a given demand, then the actual value of expected coverage varies on which sites are selected as well and not just on the number of selected site. This was the subject of recent work by Lei et al. (2014). To calculate expected coverage of a given demand, they first identified all of the different covering configurations of sites within the coverage area of that demand. For each covering configuration for that demand, they then determined the expected coverage that would be provided by that configuration given site independent availabilities. Their overall model then was designed to maximize expected coverage, where each demand is represented by the coverage provided by the best coverage configuration of that demand that exists among the selected facility locations. It is a relatively ingenious approach, but fraught with a significant degree of complexity, especially when there may exist many alternate covering configurations among those sites that can cover a given demand.

Rather than present the details of the maximal expected coverage model of Lei et al. (2014), we will present a far simpler, but approximate form of MEXCLP given varying probabilities of facility availability (site specific, p_j, or site and distance sensitive, p_{ij}). To do this, consider the following additional/modified notation:

α_k = expected coverage level k where $k = 1, 2, \ldots, K$, for example $\alpha_1 = 0.50$ or 50%.

$$y_{ik} = \begin{cases} 1, \text{if demand } i \text{ is covered at an expected coverage level } k \\ 0, \text{if otherwise} \end{cases}$$

The basic idea is that we discretize expected coverage levels into a small number, say 0.50, 0.60, 0.70, 0.80 and 0.90, or in this case $K = 5$. We can use the basic structure of constraint (4.53) to determine if a given level of coverage, say 50%, has been provided to a given demand. Using this discretized definition of expected coverage, we can define an approximate or discretized-level form of MEXCLP using site specific probabilities of availability:

$$\text{MEXCLP-GD}: \quad \text{Maximize} \sum_{j \in J} \sum_{k=1}^{K} \alpha_k a_i y_{ik} \tag{4.60}$$

Subject to:

$$\sum_{j \in J} x_j = p \quad \forall j \tag{4.61}$$

$$\sum_{k=1}^{K} y_{ik} \leq 1 \quad \forall i \tag{4.62}$$

$$\sum_{j \in N_i} \sum_{t=1}^{T} a_{ij} x_j \geq b_k y_{ik} \forall i \forall k \tag{4.63}$$

$$x_j = \{0, 1\} \quad \forall j \tag{4.64}$$

$$y_i = \{0, 1\} \quad \forall i \tag{4.65}$$

where $b_k = -\ln(1 - \alpha_k)$. The objective (4.60) maximizes the sum of expected coverages based upon what probabilistic level of coverage is attained (e.g., 60%) for each demand. Constraint (4.61) limits the number of facilities being located to equal p. Constraints (4.62) limit the levels of accounted coverage for each demand to be at most one. Constraints (4.63) are used to define whether a specific level k of expected coverage has been met for demand i. If so, then the constraint allows y_{ik} to equal one, denoting that expected coverage level k has been provided to demand i. Because only one level of coverage can be counted for a given demand, only the highest achieved level of expected coverage will be accounted for that demand in the objective function. The remainder of the constraints list the binary restrictions on the decision variables.

4.7 Extensions

There are a number of pathways in which the models presented in this chapter have been extended, often with the purpose of addressing issues experienced in applications. From the very beginnings of Location Science and the development of the classic LSCP and MCLP approaches, refinements have been proposed to address specific issues that arise in application. For example, Plane and Hendrick (1977) relocated fire stations, where the objective was to improve coverage as much as possible while keeping as many of the existing stations as possible. Often, there are many issues when taking a system and improving it by making selected changes as compared to designing something from scratch. That is, in application most problems are not "green field" in nature, but start with an existing system, which is either expanded, contracted, or reconfigured. A good example in analyzing an existing fire department and its fleet of rescue vehicles and fire trucks can be found in Pérez et al. (2016). Their motivation was to improve reliable coverage of emergency response taking into account different types of response as well as vehicle busyness in the city of Santiago, Chile. Another interesting application of covering to improve fire response can be found in the work of Aktaş et al. (2013). They analyzed the services of fire department locations in the City of Istanbul, Turkey to address many types of demand, including historic structures built during the Roman and Ottoman empires that have recently been the target of radical extremists.

Time is a major factor which has been virtually ignored in this chapter. The main reason for this is to keep our presentation as simple as possible. Demand for EMS can vary considerably over the hours of a week. Church et al. (2001) analyzed the temporal EMS demand for Kansas City (Missouri and Kansas) and presented a scheduling model that optimizes crew shifts and ambulance deployment over a week, so that periods of peak demand are well served while ensuring that periods of low demand are not overcommitted with resources. Repede and Bernardo (1994) have extended the MCLP to handle multiple time periods and Brotcorne et al. (2003) have developed a dynamic relocation model that repositions ambulances each time a call for service is made. Their model uses what they term the "double standard" where cover is maximized within one distance and required within a range that is more distant. The basic element is to reposition as few units as possible with the least overall distance of repositioning so that the double standard of coverage is met. This creates a system that is best positioned to answer the next call. Another example is that of Rajagopalan et al. (2008) in which they develop a multiperiod set covering model that they use to dynamically redeploy ambulances. Repositioning of emergency vehicles when some are busy has been a subject of interest since the classic work of Kolesar and Walker (1974) and has been found to help a system increase response coverage in real time.

Another type of extension is where facilities are co-located at a given site. Co-location helps to keep the number of stations or dispatching posts as small as possible. The idea is that co-location should be pursued whenever it is possible to accomplish without seriously reducing coverage or increasing response times. Several of the models that are presented in the previous section have been formulated with the option of allowing co-location whenever it is beneficial. Presumably, co-location may reduce the need for repositioning. It should be recognized that in less dense suburban areas, co-location may not be a viable option due to the fact that demand is spread over a larger area and facilities may need to be spread far apart, whereas in the dense urban cores co-location may be achieved without degrading coverage values.

It is also important to note that may emergency services have a goal of maximizing coverage and minimizing average response times. This means that models such as the *p*-median problem have also been used to deploy emergency resources. A good example of how that can be accomplished can be found in Weaver and Church (1985) where service is accounted for in terms of the closest vehicle as well as more distant vehicles to a demand. Thus, average response is a function of vehicle average availability. Because many coverage models are special cases of *p*-median equivalents (see Church and Weaver 1986), there exist many avenues in which coverage and average response can be handled together. There is also the possibility of using the probability chain approach developed by O'Hanley et al. (2013) to handle differing probabilities of availability in optimizing both coverage and average response distance.

4.8 Summary and Concluding Comments

The models described in this chapter represent a rather large literature devoted to emergency services and safety planning, including fire, EMS, and police as well as sensor network design. Many of these models feature covering as the primary objective. It is only natural that such models have been extended and tested in many ways. For example Borras and Pastor (2002) have analyzed several forms of probabilistic covering models. They found that the model of Ball and Lin (1993) almost always achieves the required reliability level. Their tests involved the use of a simulation model. They used their analysis to suggest a further extension of the Ball and Lin (1993) approach to achieve the same desired levels of reliability while using fewer vehicles. Sorensen and Church (2010) also used simulation to confirm whether local reliability is achieve within the context of a MALP′ approach. They also demonstrated that, in general, using local reliability estimates in an expected coverage model achieves higher levels of coverage than using local reliability estimates in MALP′. They suggested that an expanded MEXCLP approach better meets the operations standards of many ambulance systems. Altogether, simulation models like that of Erkut and Polat (1992) and Heller et al. (1989) are important tools in validating the results of optimization models in emergency management, but are useful in their own right.

The constructs that are reviewed in this chapter all involve simplifications of a highly complex system of congestion. The main reason for this is that the probabilistic nature of a complex design system is difficult to fully capture in a simpler deterministic model. However, such deterministic equivalents represent the only realistic approach to finding what are called “optimal” solutions. Unfortunately, few if any comparisons have been made between a heuristic solution process to a more realistic version of an actual system and the optimal solutions obtained from a deterministic equivalent model. It is time to demonstrate which pathway has been the most fruitful to date in addressing this important problem.

References

Aktaş E, Özaydın Ö, Bozkaya B, Ülengin F, Önsel Ş (2013) Optimizing fire station locations for the Istanbul metropolitan municipality. Interfaces 43(3):240–255

Altınel İK, Aras N, Güney E, Ersoy C (2008) Binary integer programming formulation and heuristics for differentiated coverage in heterogeneous sensor networks. Comput Netw 52 (12):2419–2431

Ball MO, Lin FL (1993) A reliability model applied to emergency service vehicle location. Oper Res 41 (1):18–36

Batta R, Dolan JM, Krishnamurthy NN (1989) The maximal expected covering location problem: revisited. Transp Sci 23(4):277–287

Baron O, Berman O, Kim S, Krass D (2009) Ensuring feasibility in location problems with stochastic demands and congestion. IIE Trans 41(5):467–481

Borras F, Pastor JT (2002) The ex-post evaluation of the minimum local reliability level: an enhanced probabilistic location set covering model. Ann Oper Res 111(1):51–74
Brotcorne L, Laporte G, Semet F (2003) Ambulance location and relocation models. Eur J Oper Res 147(3):451–463
Chapman S, White J (1974) Probabilistic formulation of the emergency service facilities location problems, presented at ORSA/TIMS National Meeting, San Juan, Puerto Rico (Reprint series no. 7407)
Church RL, Gerrard RA (2003) The multi-level location set covering model. Geogr Anal 35:277–289
Church RL, Weaver JR (1986) Theoretical links between median and coverage location problems. Ann Oper Res 6(1):1–19
Church R, Sorensen P, Corrigan W (2001) Manpower deployment in emergency services. Fire Technol 37(3):219–234
Daskin M (1982) Application of an expected covering model to emergency medical service system design. Decis Sci 13:416–439
Daskin M (1983) A maximum expected covering location model; formulation, properties, and heuristic solution. Transp Sci 17:48–70
Erkut E, Polat S (1992) A simulation model for an urban fire fighting system. Omega 20(4):535–542
Goldberg J, Dietrich R, Chen JM, Mitwasi MG, Valenzuela T, Criss E (1990) Validating and applying a model for locating emergency medical vehicles in Tuczon, AZ. Eur J Oper Res 49 (3):308–324
Heller M, Cohon JL, Revelle CS (1989) The use of simulation in validating a multiobjective EMS location model. Ann Oper Res 18(1):303–322
Kolesar P, Walker WE (1974) An algorithm for the dynamic relocation of fire companies. Oper Res 22(2):249–274
Lei TL, Tong D, Church RL (2014) Designing robust coverage systems: a maximal covering model with geographically varying failure probabilities. Ann Assoc Am Geogr 104(5):922–938
Marianov V, Revelle C (1994) The queuing probabilistic location set covering problem and some extensions. Socio-Econ Plan Sci 28(3):167–178
Marianov V, ReVelle C (1996) The queueing maximal availability location problem: a model for the siting of emergency vehicles. Eur J Oper Res 93(1):110–120
Marianov V, Serra D (1998) Probabilistic, maximal covering location—allocation models for congested systems. J Reg Sci 38(3):401–424
Murray AT (2015) Fire station siting. In: Eiselt HA, Marianov V (eds) Applications of location analysis. Springer, Berlin, pp 293–306
Murray AT, Church RL (1992) The reliability of α-reliability. Presented at ORSA/TIMS Joint National Meeting, San Francisco, California, USA, November 1–4, 1992
O'Hanley JR, Scaparra MP, García S (2013) Probability chains: a general linearization technique for modeling reliability in facility location and related problems. Eur J Oper Res 230(1):63–75
Pérez J, Maldonado S, Marianov V (2016) A reconfiguration of fire station and fleet locations for the Santiago Fire Department. Int J Prod Res 54(11):3170–3186
Plane DR, Hendrick TE (1977) Mathematical programming and the location of fire companies for the Denver fire department. Oper Res 25(4):563–578
Rajagopalan HK, Saydam C, Xiao J (2008) A multiperiod set covering location model for dynamic redeployment of ambulances. Comput Oper Res 35(3):814–826
Repede JF, Bernardo JJ (1994) Developing and validating a decision support system for locating emergency medical vehicles in Louisville, Kentucky. Eur J Oper Res 75(3):567–581
ReVelle C, Hogan K (1988) A reliability-constrained siting model with local estimates of busy fractions. Environ Plan B: Plan Des 15(2):143–152
ReVelle C, Hogan K (1989) The maximum availability location problem. Transp Sci 23(3):192–200
Sorensen P, Church R (2010) Integrating expected coverage and local reliability for emergency medical services location problems. Socio-Econ Plan Sci 44(1):8–18
Toregas C (1970) A covering formulation for the location of public service facilities. Masters Thesis, Cornel University, Ithaca

Toregas C, Revelle C (1973) Binary logic solutions to a class of location problem. Geogr Anal 5 (2):145–155

Toregas C, Swain R, ReVelle C, Bergman L (1971) The location of emergency service facilities. Oper Res 19(6):1363–1373

Weaver JR, Church RL (1985) A median location model with nonclosest facility service. Transp Sci 19(1):58–74