Industrial and Systems Engineering Review, 3(1), 2015 ISSN (Online): 2329-0188

Military Resource Allocation as a Set Covering Problem

Alexander Bates, Zach Bell, Alexander Mountford, and Paul Evangelista

Department of Systems Engineering
United States Military Academy
West Point, NY, USA

Corresponding Author's Email: paul.evangelista@usma.edu

Author Note: Alexander Bates, Zach Bell, and Alex Mountford graduated from the United States Military Academy (USMA) in 2014 where they majored in Systems Engineering. Currently all second lieutenants in the U.S. Army, they worked on this research during their final year at USMA under the advisement of Paul Evangelista. Paul Evangelista is a lieutenant colonel in the U.S. Army and currently serving as an Academy Professor in the Department of Systems Engineering and the Director of the Operations Research Center at USMA.

Abstract: Fixed location resource allocation modeled as a set covering problem is a classic integer program. This framework has been widely applied to emergency service resource allocation, and this paper extends this approach to military resource allocation. Using military air medical evacuation resource allocation in Afghanistan as a proof of concept, a methodology is presented that could easily extend to other operational environments and other military resource allocation problems. Unique contributions include clustering of enemy activity reports to support demand signal analysis and consideration of set covering requirements for varying demand signal density.

Keywords: Set covering problem, k-means clustering, military resource allocation

1. Introduction

Resource allocation in military environments is a consistent problem for military planners. According to U.S. Army doctrine, planning is considered the “art and science of understanding a situation, envisioning a desired future, and laying out effective ways of bringing that future about” (U.S. Army, 2012). The military community stands to gain advantages by embracing systematic methods to guide resource allocation decisions in support of combat operations, both in terms of operational effectiveness and efficiency. Significant resource allocation decisions in operational environments include intelligence, surveillance, and reconnaissance (ISR) asset mission tasks; locations for quick response forces (QRF); locations for fires assets; and locations for medical evacuation assets. This paper will focus on the locations for air medical evacuation (MEDEVAC) assets; however, we assert that this optimization framework could be used for any of these other resource allocation decisions as well. Specifically, this research applies a set covering problem approach to allocate air MEDEVAC assets. Demand signals for the MEDEVAC covering problem have been derived from actual enemy activity reports. The analysis in this paper serves as a proof of concept and should not be considered as an actual solution to the current MEDEVAC demand in Afghanistan. Although this research uses actual data, there are several elements of the problem that need to be considered in real time before a solution could be operational. These elements include unit availability, Forward Operating Base (FOB) availability, and general aviation limitations and support requirements.

1.1 Related Work

There is a large body of literature related to facility location and set covering problems, particularly for emergency service location problems. Toregas et. al. (1971) wrote the seminal work and many others have extended this work (Church and ReVelle (1974), Marianov and ReVelle (1996), Daskin and Stern (1981), and Gendreau et. al. (1997)). Using a historical analysis framework, ReVelle and Rosing (2000) apply set covering to the stationing of Roman Legions across Europe in the third century A.D. In this work they also motivate other military applications with an emphasis on military unit deployment. The research on covering problems generally divides into two main types: the location set covering problem (LSCP) and the maximal covering location problem (MCLP). The MCLP, pioneered by Church and ReVelle (1974), maximizes demand site coverage within a desired service distance by locating a fixed number of service facilities. The MCLP does not require full coverage of all demand points, but rather it seeks the maximal coverage with a fixed number of service facilities. The LSCP aims to “locate the least number of facilities that are required to cover all demand points”, as defined by

Toregas et. al. (1971). Since all the demand points need to be covered in LSCP, regardless of their population, remoteness, and demand quantity, the resources required for facilities could be excessive. In order to maintain a tractable problem, the research in this paper uses k-means clustering to prevent unreasonable or excessive requirements. Our research has been identified as a location set covering problem. We seek to locate the least number of facilities that are required to fulfill all demand points.

This research aims to answer questions of set covering from a military perspective, specifically MEDEVACs. Medical facility allocation problems are commonly addressed with set covering models. For example, in the work "The Location of Emergency Service Facilities," the authors present three different formulations for large scale emergency medical service scenarios. Their research emphasized several assumptions to achieve a more tractable problem structure. We adapted their assumptions in our MEDEVAC problem. The first adapted assumption is that user demands can be represented as occurring at a finite set of points and that the potential locations for service facilities are also a finite set of points. The second adapted assumption is that the minimum distance or minimum response time between any user-node/service-facility pair is known.

1.2 Organization of this Paper

Section 2 of this paper describes the methodology applied to the MEDEVAC resource allocation problem. The example in this paper is based on real world data, and this section discusses important aspects of the manipulation and modeling of this data. Section 2 also includes the specific set covering mathematical formulation applied. Section 3 includes the results of the analysis as well as an interpretation of the results. Section 4, the discussion, includes the broader interpretation of the results and possible extensions of the work. Section 5 includes the conclusion and recommended future work.

2. Methodology

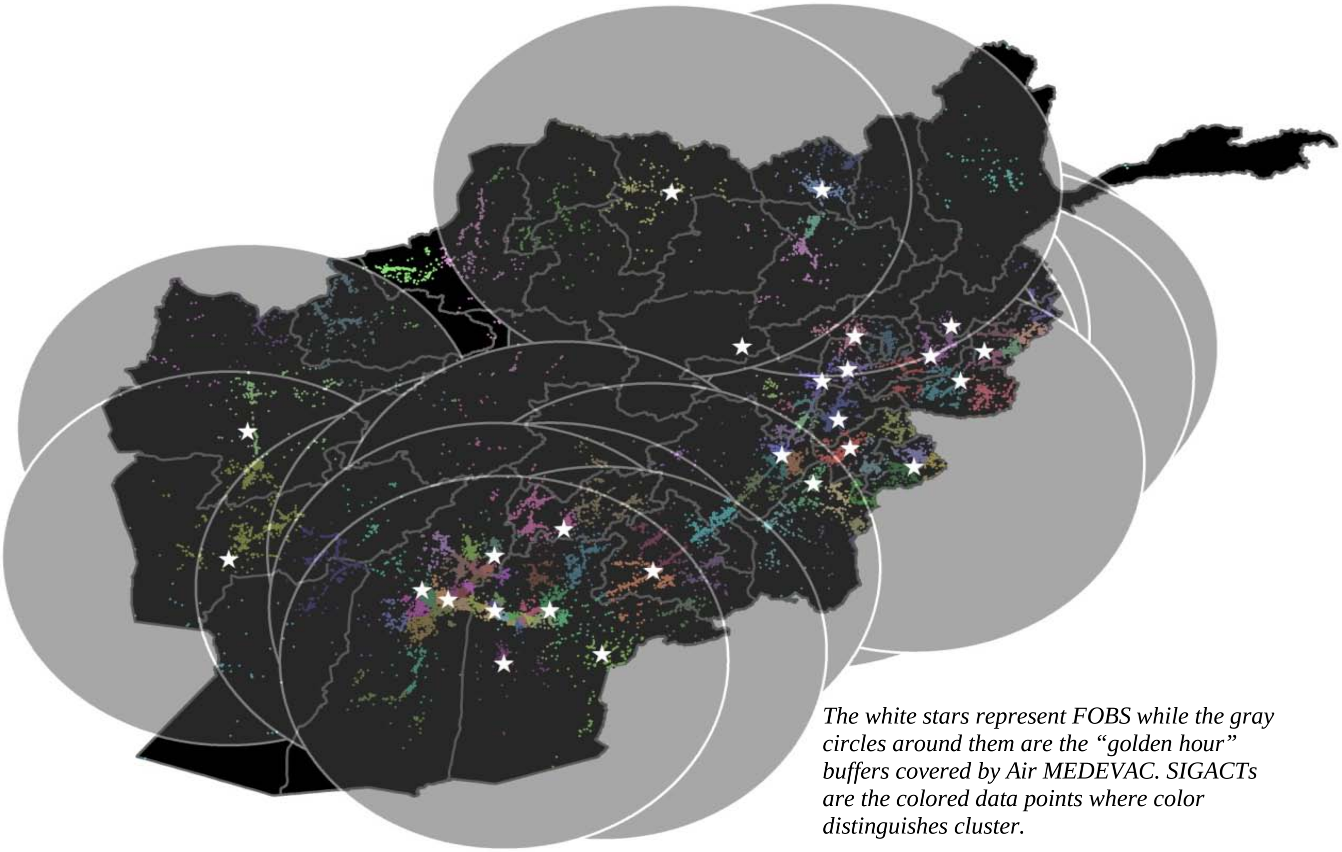
2.1 Demand Clusters, Service Locations, and the Golden Hour Criteria

Set covering problems require demand locations and service providing locations. Given k demand locations and n possible service providing locations, the LSCP seeks to find the fewest number of service providing locations from the set of n capable of covering k demand locations. Significant activity (SIGACT) reports from Afghanistan created a demand signal for the set covering problem methodology employed in this paper, serving as a proxy for potential MEDEVAC locations. The SIGACTs used in this research came from the unclassified INternational Distributed Unified Reporting Environment (INDURE) database and the period covered included October of 2012 to September 2013. Each SIGACT typically represents an incident in which enemy engaged, or attempted to engage, coalition forces. SIGACTS in this data are divided into two main categories: explosive hazard or enemy action. For every SIGACT, in addition to a brief description and date of the event, there is also a latitude and longitude representing location.

Forward Operating Base (FOB) sites represented service providing locations. Two references provided the information necessary for identifying FOB locations: GlobalSecurity.org and the Afghanistan Order of Battle (Wesley, 2013). GlobalSecurity.org maintains a comprehensive list of FOB locations, and the Afghanistan Order of Battle contains a detailed history over time of unit locations by FOB. Each FOB was ultimately assigned a latitude and longitude location, the final list for this research included 26 FOBs.

INDURE records contain over 47,000 reports for the one year time period applied for this research. K-means clustering was used to reduce the total number of demand locations, or clusters, to $k = 100$. Selection of the number of clusters, k, was somewhat arbitrary and remains worthy of further study. The general rule for selecting k was "larger is better" while maintaining a reasonable solving time frame (less than 10 minutes). After forming clusters, it was necessary to determine whether or not a FOB could provide service to the cluster. In order to ensure that the FOB could cover all locations in a given cluster, distance to the farthest point in the cluster was used to represent the FOB to cluster distance. If this distance was less than the coverage distance threshold, then $c_{ij} = 1$, indicating that FOB i could reach cluster j, otherwise $c_{ij} = 0$. The final FOB to cluster coverage matrix was 26 x 100, a manageable size for solving a set covering problem on a laptop using an Excel solver.

It is worth noting the operational consideration regarding FOB to cluster coverage. The binary indicator c_{ij} indicates whether or not a MEDEVAC helicopter from FOB i can reach a casualty in cluster j within the "golden hour" criteria. The "golden hour" is attributed to R Adams Cowley, a historic trauma physician (Sloan, 1992). The "golden hour" concept asserts that trauma patients that receive emergency medical care in a trauma center within one hour of sustaining their injury have higher chances of survival (Lerner and Moscati, 2001). This concept is widely embraced in the military. The "golden hour" criteria used in this paper establishes that MEDEVAC helicopters must reach casualties within one hour. To model this concept, the range that a UH-60M Blackhawk helicopter could reach in one hour needed to be determined. To do this, the industry standard of 5 minutes to dust-off and the time it takes to get the helicopter in the air, was considered. Then, the remaining 55 minutes and the cruising speed of the helicopter, 280 km/h, could be used to calculate the range. We rounded down to 250 km to account for possible deviations in the air routes that would result in a slightly less-than-linear path to the casualty. For this reason, the covering radius surrounding every FOB was 250 km.



The white stars represent FOBS while the gray circles around them are the "golden hour" buffers covered by Air MEDEVAC. SIGACTs are the colored data points where color distinguishes cluster.

Figure 1. Model of Afghanistan with FOBs, "golden hour" buffers, and SIGACT clusters

2.1 Formulation

$$\text{min} \qquad \sum_{i=1}^{n} x_i \tag{1}$$

$$\text{S.T.} \qquad \sum_{i=1}^{n} c_{ij} x_{ij} \geq m_j \tag{2}$$

x_i = number of MEDEVAC units placed at FOB i, constrained to non-negative integers, this variable is the decision variable for the formulation.

c_{ij} = this parameter indicates coverage of cluster j by FOB i and is binary. c_{ij} = 1 indicates that FOB i covers cluster j, 0 otherwise.

m_j = minimum number of MEDEVAC helicopters required to cover cluster j
If cluster j has less than or 30 total SIGACTs on average per month then $m_j = 1$
If cluster j has greater than 30 total SIGACTs on average per month then $m_j = 2$
If cluster j has greater than 60 total SIGACTs on average per month then $m_j = 3$
If cluster j has greater than 90 total SIGACTs on average per month then $m_j = 4$

The parameter m_j represents MEDEVAC coverage demand based on SIGACT density. Areas with a higher density of SIGACTs are assumed more likely to require concurrent MEDEVAC requests. It is understood that m_j is not a perfect representation and an estimation of MEDEVAC coverage demand. This value requires subject matter expertise with recent operational knowledge. This model also assumes that an increase in SIGACTs in a particular cluster results in an increase in MEDEVAC demand, assuming that all SIGACTs generate the same number of casualties. Some SIGACTs are more likely to produce a casualty than others. An explosive event SIGACT compared to a SIGACT reporting a removal of unexploded ordnance should be expected to create more casualties. Ideally, MEDEVAC demand would be based upon actual data if available. However, it is possible that MEDEVAC planning will be necessary in new locations. This is a very likely planning scenario. If this is the case, some type of proxy or estimate for MEDEVAC demand would be necessary.

3. Results

The solution calls for a total of 15 MEDEVAC units spread across eight FOBs. Bagram Airfield and Camp Holland both call for 3 units. Camps Alamein and Marmal along with Herat Airfield call for 2 MEDEVAC units. Finally, FOBs Kunduz, Salerno, and Walton each call for a single MEDEVAC unit. It is interesting to note that the units needed to cover the entire country of Afghanistan are placed on 8 of the total 26 bases.

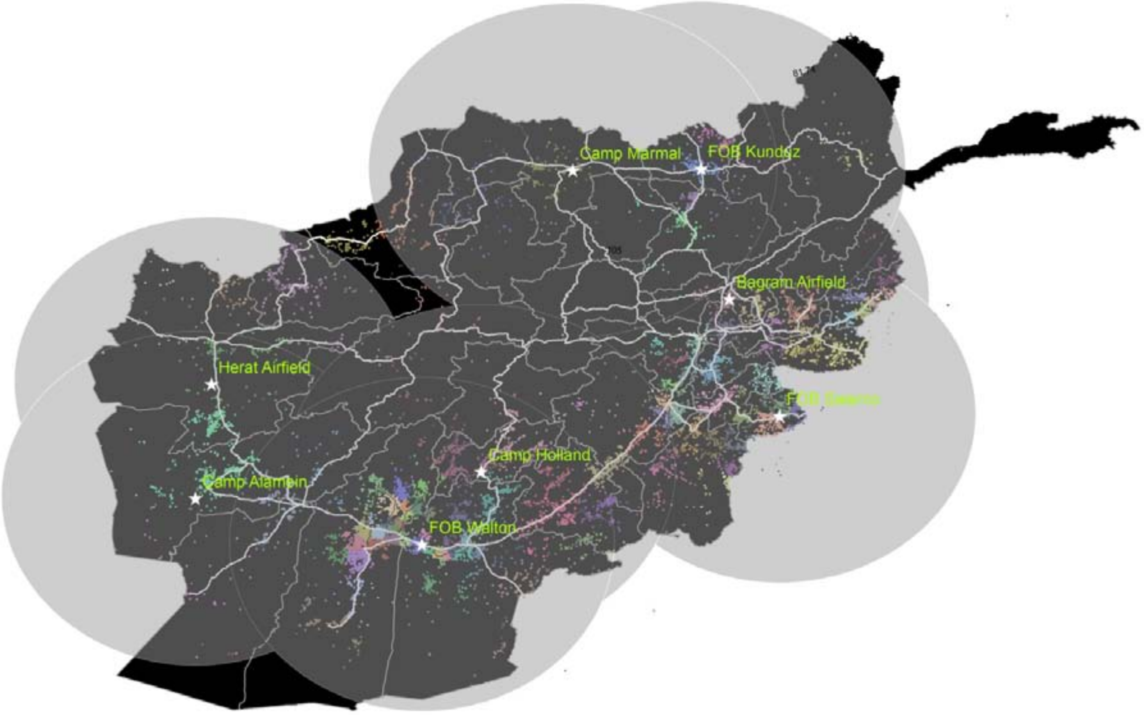


Figure 2. Set covering solution for Afghanistan MEDEVAC problem

The logic of the solution can be seen when examining each FOB included in the solution. Some of the nodes are connected to an extremely large number of clusters, up to 43 out of 100 clusters, such as FOBs Salerno and Walton, Camp

Holland, and Bagram Airfield. The distribution of MEDEVAC units to these nodes makes perfect sense. By placing MEDEVAC units at these nodes, MEDEVAC units are capable of accessing a large battle space. The other four nodes, Camps Alamein and Marmal, Herat Airfield, and FOB Kunduz, are not connected to a large amount of clusters. The largest number of clusters that one of them is connected to is 7 out of 100 clusters. Because they are not well connected nodes a different reasoning is presented for their selection. These nodes are connected to some of the outlying clusters that don't have many connections yet still require MEDEVAC attention. Because every cluster must be capable of being reached with the number of MEDEVACs required given the number of SIGACTs, these nodes are selected based on their connection to these outlying clusters. Readers will also notice that there is a strip of terrain in northwestern Afghanistan that is beyond the reach of any existing FOB. The model was manually adjusted to account for this gap, allowing the nearest FOBs to cover these demand locations while realizing the distances does exceed the golden hour threshold.

Sensitivity analysis for *k*, the number of clusters, is a subject worthy of further research. Changing the number of clusters provides some insight into model behavior and general trends. The number of clusters within the model was both increased to 150 clusters and decreased to 50 clusters. Upon increasing the count of clusters to 150 the model returned an optimal result that reduced the total number of MEDEVAC units required to 14, a decrease of 1 unit. Additionally, there were 9 FOBs that served as MEDEVAC holding areas, an increase of 1 FOB. A much more drastic result was produced when the number of clusters of was reduced to 50. The number of required MEDEVAC units increased to 21, a gain of 6 units. Also, the number of FOBs serving as MEDEVAC holding areas increased to 11, an increase of 3 FOBs. These results are actually not surprising however when compared against the effect clusters have on the model. As the number of clusters decrease, the number of SIGACTs that occur within a single cluster increases. This increase also inflates demand for MEDEVACs within a particular cluster. Because the set of FOBs that cover a particular cluster does not drastically change when the cluster's area is increased, the same FOBs then have to produce a greater number of MEDEVAC units based on the assumptions of the model. The opposite holds true when cluster size is decreased, but the results of the sensitivity testing indicate that increasing the number of clusters does not have a drastic impact on the number of MEDEVAC units required.

4. Discussion

When examining these results it is important to keep in mind that the results were produced using the data set of only one year. While using an entire year's worth of data seemingly provides sufficient fidelity, it should also be understood that this data only represents one year of a war that is currently running over a decade long; therefore, the SIGACTs within a cluster might vary for a multitude of reasons. If allied forces decide to reduce to presence in an area, SIGACTs will decrease due to a lack of targets and fewer MEDEVAC units will be needed to assist those clusters. On the opposite spectrum, if allied forces significantly increase troop presence in an area two results are possible. SIGACTs might go down in the region because insurgents would want to move away from the more powerful force. However, after over a decade of conflict in Afghanistan, there is little evidence to support this theory. Therefore, the more likely scenario is that SIGACTs could increase within the cluster due to the increased number of targets for insurgents. The variation in courses of action that the enemy could take in addition to the older, unclassified SIGACT data prevents this model from becoming explanatory, but rather it can be utilized in a predictive method with up-to-date data to try and help plan for future operations.

This model could easily be extended to assist other military applications and provide insight past what is provided in this paper. The area that is the easiest extension is with a quick reaction force or a QRF. QRFs are the primary reaction force to assist friendly units operating within a certain distance. A QRF operates on similar premise as a MEDEVAC unit but it often is restricted to ground operations. Therefore the reach of a singular QRF may not be as much but there are likely more units available to assist in the operations. Another extension in the military that this model could be extended to is the concept of fires. Fires as a military concept incorporates a large variety of units ranging from an aerial UAV engagement team, close air support, and artillery. Each of these unit types require correct placement for the optimal use of their capabilities. Thus to enhance their capabilities this model could be used to determine where each unit should be placed to achieve maximum impact on the battle space.

5. Conclusion And Future Work

The location set covering methodology discussed in this paper has been specifically tailored to MEDEVAC coverage, however we believe it can expand to a wide variety of problems and theaters. As previously discussed, the set covering problem formulation has been broadly applied to emergency services applications. Similarities between emergency services and military operations include unique resource allocation decisions, geographic areas of responsibility, and timely responsiveness. Military applications for set covering problems abound. Deployment problems (ReVelle and Rosing, 2000), selection of operating base locations, and resource allocation are amongst the types of military problems that could be addressed with set covering approaches.

It is also worth noting the analytical tools used to compile this analysis. ArcMap from ESRI supported all of the GIS work; clustering was completed using R's "kmeans" function from the stats library (R Core Team 2013); and Microsoft Excel was used for the proximity calculations (Haversine function between all FOBs and SIGACTs) and as the solver engine. Each of these tools was essential to this work and has also shown their worth in many other military operations research applications.

Analysis of seasonality also presents opportunities for future work. Seasonality in Afghanistan is significant (Gons et. al., 2000), and adjusting models for expected changes in seasonality – both in terms of enemy behavior as well as seasonal effects on military equipment capabilities – remains an open area for research. Sensitivity analysis for the clustering approach used in this paper also warrants further work. Creating neighborhoods for optimization problems is common practice. Ultimately, optimization models require a finite solution space. Creating neighborhoods is a method to create a finite solution space and reduce solver processing requirements. Clustering is a method of creating neighborhoods, however several assumptions must be made to impellent clustering approaches. The clustering size and clustering method are two of the major assumptions, both of which deserve further analysis.

The most immediate recommendation for future work on this model is inclusion of the ground network to more accurately represent the operational environment and capabilities of forces deployed in Afghanistan. Adding the ground network to work concurrently with the air network will provide a more accurate set of options for resource allocation. For example, the current model assigns air MEDEVAC units according to all SIGACTs within a given radius. However, air MEDEVAC units are not necessary to service all medical evacuations.

6. References

Boyaci, B. and Geroliminis, N. (2012). Facility Location Problem for Emergency and On-Demand Transportation Systems. Proceedings of the 2012 Swiss Transport Research Conference.

Church, R. and ReVelle, C. (1974). The Maximal Covering Location Problem. *Papers in Regional Science*, 32: 101–118.

Daskin, M. and Stern, E. (1981). A Hierarchical Objective Set Covering Model for Emergency Medical Service Vehicle Deployment. *Transportation Science*. 15(2): 137–152.

Edlich, R.F., Wish, J.R., Britt, L.D., and Long, W.B. (2004). An Organized Approach to Trauma Care: Legacy of R Adams Cowley. *Journal of Long Term Effects of Medical Implants*. 14(6): 481-511.

Gendreau, M., Laporte, G., and Semet, F. (1997). Solving an Ambulance Location Model by Tabu Search. *Location Science*. 5(2): 75–88.

Gons, E., Schroden, J., McAlinden, R., Gaul, M., and VanPoppel, B. (2012). Challenges of Measuring Progress in Afghanistan Using Violence Trends: The Effects of Aggregation, Military Operations, Seasonality, Weather, and Other Causal Factors. *Defense and Security Analysis*. 28(2): 100-113

Jia, H., Ordóñez, F. and Dessouky, M. (2007). A Modeling Framework for Facility Location of Medical Services for Large-Scale Emergencies. *IIE Transactions*. 39(1); 41-55.

Lerner, E. B. and Moscati, R. M. (2001). The Golden Hour: Scientific Fact or Medical "Urban Legend"? *Academic Emergency Medicine*. 8: 758-760.

Marianov, V. and ReVelle, C.S. (1996). The Queueing Maximal Availability Location Problem: A Model for the Siting of Emergency Vehicles. *European Journal of Operational Research*. 93(1): 110–120.

Morgan, Wesley. (2013). Afghanistan – Order of Battle. Retrieved from the Institute for the Study of War, http://www.understandingwar.org/report/afghanistan-order-battle.

R Core Team. (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

ReVelle, C.S. and Rosing, K.E. (2000). Defendes Imperium Romanum: A Classical Problem in Military Strategy. *The American Mathematical Monthly*. 107(7): 585-594.

Sloan, H. (1992). R Adams Cowley, MD: 1917-1991. Annals of Thoracic Surgery. 53(6): 954.

Toregas, C., Swain, R., ReVelle, C.S. and Bergman, L. (1971). The Location of Emergency Service Facilities. *Operations Research*, 19: 1363-1373.

United States Army. (2012). *The Operations Process.* Washington, D.C.: Headquarters, Department of the Army.