ESTIMATING THE PERFORMANCE OF EMS LOCATION MODELS VIA SIMULATION

by

YASIR TUNÇER

Submitted to the Graduate School of Engineering and Natural Sciences
in partial fulfillment of
the requirements for the degree of
Master of Science

Sabancı University
August 2011

ESTIMATING THE PERFORMANCE OF EMS LOCATION MODELS VIA SIMULATION

APPROVED BY:

Assoc. Prof. Dr. Tonguç Ünlüyurt ....
(Thesis Supervisor)

Assoc. Prof. Dr. Bülent Çatay ....

Assist. Prof. Dr. Fatma Tevhide Altekin ....

Assist. Prof. Dr. Güvenç Şahin ....

Assoc. Prof. Dr. Burçin Bozkaya ....

DATE OF APPROVAL: ....

ACKNOWLEDGEMENTS

First, I would like to my thesis supervisor Assoc. Prof. Dr. Tonguç Ünlüyurt for his sincerity and excellent personality. Throughout my study, his guidance and exceptional patience helped me to endure this path. I also would like to thank 112 Command Center employers for their kind and helpful approach to this study. I am also grateful for my family and friends for their understanding attitude and their help during my graduate education.

ESTIMATING THE PERFORMANCE OF EMS LOCATION MODELS VIA SIMULATION

Yasir Tunçer

Industrial Engineering, Master of Science Thesis, 2011

Thesis Supervisor: Assoc. Prof. Dr. Tonguç Unluyurt

Keywords: set covering location models, deterministic models, simulation, EMS, Hypercube model, busy probability, performance metric

Abstract

In this thesis, we address the problem of evaluating deterministic EMS location models via simulation. For deterministic set covering location models, the performance of the model is typically determined by an objective function representing a certain type of coverage. After determining the location of EMS stations by deterministic models, we propose to conduct a simulation analysis to evaluate the performance by estimating the "real" coverage of the population. We compare 4 different deterministic models, Maximum Coverage Location Model (MCLM), Single Period Backup Double Covering Model (SPBDCM) which is a variant of MCLM, Maximum Service Restricted Set Covering Location Model (MSRSCLM) and finally Centralized Final Ratio Model (CFRM). By using optimization tools, we find the location of ambulances for each model by sıolving the mathematical models and then we simulate each setting for 2 different policies under the same parameters. The models' overall performance is firstly tested on Istanbul data and then followed up with extensive experimental study with different problem size, different layout and different arrival rates with two distinct policies. The tested policies include first come first serve with zero line capacity and lost call approach, and dispatching the closest idle station whether the call origin is served at the moment of the call or not. The study is then extended by a myopic heuristic, which basically tries to improve the performance of the system by relocating ambulances.

ACİL YARDIM SİSTEMİ YERLEŞTİRME MODELLERİNİN BENZETİMLE BAŞARIMININ HESAPLANMASI

Yasir Tunçer

Endüstri Mühendisliği, Yüksek Lisans Tezi, 2011

Tez Danışmanı: Doç. Dr. Tonguç Ünlüyurt

Anahtar Kelimeler: küme kaplama yerleştirme modelleri, gerekirci modeller, benzetim, Acil Yardım Sistemleri, Hiperküp modeli, meşgul olasılığı, performans metriği

Özet

Bu tezde gerekirci Hızır Acil Sistemleri yerleştirme modellerinin benzetimle değerlendirilmesini konu alıyoruz. Gerekirci küme kaplama modelleri için, modelin performansı genellikle belirli bir kaplama türünü gösteren amaç işlevi tarafından belirlenir. Acil Yardım Sistemi istasyonlarının yerleri gerekirci modeller tarafından belirlendikten sonra, nüfusun "gerçek" kaplamasını hesaplamak için bir benzetim analizinin yapılmasını öne sürüyoruz. 4 farklı gerekirci modeli, EnBüyük Kapsama Modeli, Tek Dönemli Yedek Çift Kapsama Modeli, EnFazla Servis Kısıtlı Küme Kapsama Yerleştirme Modeli, ve Özekli Son Oran Modelini karşılaştıryoruz. Eniyileme araçlarını kullanarak, her bir model için matematiksel modellerden ambulansların yerlerini buluyoruz ve aynı parametreleri kullanarak 2 farklı kuralla her düzeni benzetimliyoruz. Modellerin toplam performansı öncelikle Istanbul verisi üzerinde test edildi ve farklı problem büyüklüğü, farklı yerleştirme, ve farklı varış hızlarıyla iki farklı kuralı içeren kapsamlı deneysel çalışmayla devam edildi. sıfır kuyruk kapasiteli ilk giren ilk çıkar ve kayıp çağrı yaklaşımı ile çağrı kökeni çağrı anında servis alsa da almasa da en yakın boşta olan istasyonun dağıtımı test edilen kurallardır, Çalışma daha sonrasında ambulansların yeniden konumlandırmasıyla sistemin performansını arttırmaya çalışan bir miyop bulgusalla genişletildi.

Table of Contents

List of Figures

List of Tables

CHAPTER 1

INTRODUCTION

The location planning for Emergency Service Systems (ESS) or Emergency Medical Systems (EMS) is very crucial and their importance is increasing every day. Since highly populated cities have high traffic volume and irregular urbanization, locating the ambulances effectively for EMS can decrease the number of disabilities, and fatalities drastically. However, even though the importance of locating ambulances is not negligible, we cannot solely determine the effectiveness of an EMS based on covered population. In order to realistically calculate the performance of an EMS, the overall busy probabilities and the population that cannot be served on average should also be considered.

The purpose of this thesis is to propose a simulation process to measure the performance of an EMS. Even though there have been lots of studies on location determination for EMS, the studies that focus on the performance of the system is fewer than expected. Brotcorne *et al.* (2003) and Goldberg (2004) give good review of the studies on location planning of EMS. The general approach for location planning of EMS is developing a mathematical model that reflects the fundamental assumptions that can vary substantially. One of the main models for EMS is Maximal Covering Location Model (MCLM) is proposed by Church and ReVelle (1974) where the single coverage of a demand point is assumed to be enough. Further on, double coverage for a demand point is claimed to be more realistic and effective; by providing a back up ambulance when the first ambulance is dispatched for another call the coverage of demand points can be maintained. Gendreau *et al.* (1997) stated Double Coverage Model (DCM) and this study is revised from the study of Hogan and ReVelle (1986) which stresses Backup Coverage Model (BCM). However, none of these studies provide an effective performance measure for their approach besides the objective function values of the deterministic mathematical models. In this regard, Larson (1974) provides an approximation for the busy probabilities of the demand points by introducing Hypercube Model. Furthermore, the model is quite effective, but calculating the steady state equations for large scale problems is a major drawback. Galvão and Morabito (2008) state that in order to be able to evaluate the performance of the system there should be some performance measures like overall busy probabilities that are usually obtained by queuing models or simulation. Moreover, some of the assumptions that are stated by Larson (1974) are contradictive to real life problems. In order to cope with these conditions, we embrace a simpler approach, simulation. After the locations of the ambulances are determined from mathematical models, we simulate the obtained settings under different parameters in order to estimate the busy probabilities for each demand point.

The proposed simulation process, assumes exponentially distributed interarrival times in other words Poison arrival process and for service time the process in our initial approach is assumed to be exponentially distributed as well. However, during late phases of our study, the service time is assumed to be compilation of sub processes which are mainly exponentially distributed with an additional distance dependent sub process. There are two policies that are tested throughout this study. First policy assumes that if a call is originated from a covered node, then the closest idle ambulance will be dispatched. For the first policy if a call is originated from an uncovered node then the call is assumed to be lost. However for the second policy, the closest idle ambulance is dispatched whether the node is covered or not. Even though there is no idle ambulance present at the coverage vicinity of the call origin, the policy still sends the closest idle ambulance to dispatch.

There are 4 different set covering location models that are evaluated in our study. The mathematical models are solved optimally in order to find the location of the ambulances. Maximal Coverage Location Model, Single Period Backup Double Covering Model, Maximum Service Restricted Set Covering Location Model, and finally we propose another mathematical model Centralized Final Ratio Model for comparison. Each model is firstly tested for Istanbul data which is a large scale problem and then the experimental study is conducted for each setting. For each model, if a demand point is covered then the whole population of the node is assumed be covered also. Each model is simulated under the same parameters for both policies in order to compare the efficiency of models to each other. After the tests on Istanbul data are conducted, we generate random instances for testing the robustness of the results. Even though, the performance of the models for Istanbul data give a general opinion on performance measures, the experimental study with different layout, different problem size, different arrival rates for the dispatching policies illustrated the importance of a realization tool. Without an analytical tool for realization, objective function values, the assumptions are inconclusive. Experimental study shows that the performance of a location planning model can vary according to problem data set based on problem size, the distribution of demand points, the design of the layout even with different dispatching policies.

Contribution of the thesis can be summarized as

- Testing the necessity of an analytical tool for evaluating the performance of deterministic location planning models
- Testing for relation between layout design and performance of the model
- Proposing two new deterministic set covering location models in order to improve simulation performance

The organization of this thesis is as follows. The literature on location planning of EMS stations, performance measures for the set covering location models and other analytical tools to determine the performance of an EMS in Chapter 2. Chapter 3 is dedicated to proposed simulation methodology, benefits and drawbacks of the mathematical models. The results of the simulations and sensitivity analysis of the models for Istanbul data is in Chapter 4. The experimental study and results are conducted in Chapter 5. Chapter 6 is dedicated to myopic improvement heuristic and application of this heuristic for Istanbul data. Finally, Chapter 7 concludes the study with discussion of results and possible future research.

CHAPTER 2

LITERATURE REVIEW

Since the average life expectancy is significantly increased in the last century, providing an effective ESS or EMS became more important. The topic attracted many researchers in the optimization field to present efficient mathematical models and new approaches. There are numerous attempts to solve the problem and its variants by heuristics, meta-heuristics however, by improved computational power now we can solve set covering problems optimally in reasonable times.

The different approaches can be classified into two main categories when the natures of the models are considered, deterministic and stochastic models. Fundamentally deterministic Set Covering Model (SCM) is stated by Toregas *et al.* (1971) and the objective is acquiring minimum number of servers that are needed in order to provide full coverage.. Since SCM grasps the general nature of the problem, the various methods are revised from this model. However for deterministic case SCM does not consider every aspect like when an ambulance is departed for a call, other nodes covered only by this ambulance are no longer covered in SCM as stated by Brotcorne *et al.* (2003). One of the revised and well known version of SCM is Maximal Coverage Location Model (MCLM) which is proposed by Church and ReVelle (1974). MCLM aims to maximize the population or the number of demand nodes with a limited number of stations. MCLM considers that a demand node is covered if any station can serve to this demand node. Tandem Equipment Allocation Model (TEAM) on the other hand is a revised version of MCLM. TEAM is suggested by Schilling *et al.* (1979) which considers two type of service levels and for each service type the total number of stations is limited. SCM, MCLM, and TEAM focuses on single coverage of any demand point. Since single coverage of a demand node is sufficient, the major drawback of these models is when any of the ambulances dispatched for a call, the region that is covered by the dispatched ambulance will be no longer be covered during service time. This drawback drew attention to further research and multiple coverage models are proposed in the literature. Daskin and Stern (1981) stressed the Modified Maximal Covering Location Model (MMCLM) and besides the primary objective, maximizing the covered population; a secondary objective is introduced which maximizes the multiple coverage of the demand points. The variants of MMCLM like

Double Covering Model (DCM), on the other hand, approached the location planning problem from a different perspective and rather than using two different objectives, maximizing the weighted coverage of demand points is embraced. Also another variant of DCM is proposed by Hogan and ReVelle (1986) where the population that is covered twice is maximized with limited number of stations. For EMS problems Setzler (2009), Simpson (2009) provide good reviews and stress the importance of emergency response.

Even though there have been remarkable studies on this subject, all of the previous models assumed that the number of stations that can be allocated to any supply point is restricted with a binary decision variable. However, Gendeau *et al.* (1997) proposed Double Standard Model (DSM) which maximizes the double covered population with two different service time restriction where the number of stations that can be allocated on any supply point has an integer upper bound.

When stochastic approach is considered, as Basar (2008) suggested the variations are originated from the objective function and constraints. One of the oldest probabilistic location planning models is Maximal Expected Covering Location Problem (MEXCLP) proposed by Daskin (1983). Daskin assigns an equal busy probability to all vehicles where this probability is based on the frequency of the calls and service time needed for each call. The restrictions are fundamentally the service provided on a day and the total number of vehicles where the objective function is to maximize the expected coverage of demand points. An extension of MEXCLP, TIMEXCLP is developed by Repede and Bernardo (1994) where the variation in travel speed throughout the day is explicitly considered. TIMEXCLP is then applied for Louisville, Kentucky data and authors proposed a simulation process to provide an assessment of the proposed solution. Extending the research with a simulation module showed that even though the objective function and constraints are important for an effective planning, the performance of the system is not solely based on these criteria. Batta *et al.* (1989) extended MEXCLP and each term in the objective function is multiplied with a correction factor which includes that ambulances do not operate independently and formed the adjusted MEXCLP, AMEXCLP. ReVelle and Hogan (1989) suggested two new models for stochastic location planning problems and formulated Maximum Availability Location Problem (MALP). For MALP I, authors suggested the busy fraction proposed by Daskin (1983) should be same for all potential candidate sites but for MALP II this condition is relaxed.

Whether the approach is deterministic or stochastic, given the plan of ambulance locations and demand levels, there are analytical tools to estimate the busy probabilities. As Brotcorne *et al.* (2003) suggested busy probabilities can be estimated with methods like Hypercube model, iterative optimization algorithm or simulation. One of the most effective methods for evaluating busy probabilities for server to customer systems is Hypercube model. Hypercube model is initially proposed by Larson (1974) and then extended by Jarvis (1975) who relaxed the service time distribution to be a general distribution rather than exponential. Swersey (1994) also applied Hypercube model and approximation of the methodology for different instances. However, even though Hypercube model is accurate for calculating busy probabilities, the equilibrium of each demand point is determined with respect to steady state equations. Galvão and Morabito (2008) suggest that in order to solve the hypercube model, an EMS with N servers, will require solution of 2^N linear equation. For large scale problems this number of linear equations cannot be solved simultaneously in reasonabletimes. Furthermore, the model is based on spatial queuing theory and Markovian property of the system and if the dispatching rule conflicts with these properties then the model becomes obsolete. For example if the dispatching rule requires the control of the previous state of the problem then the Markovian property would be lost. Examples of applications of the hypercube model in urban EMS in the United States can be found in studies by Larson and Odoni (1981), Brandeau and Larson (1986), Burwell et al. (1993) and Sacks and Grief (1994). Ingolfsson *et al*. (2008) studied the location planning with delays, Erkut *et al.* (2008) on the other hand emphasized the importance of location planning by considering survival ratios.

However, in order to deal with large scale problems and any dispatching rule, a generalized simulation approach can be embraced. Savas (1969) studied the performance of the New York EMS and applied a simulation methodology in order to deal with large scale problem size issue. Iannoni *et al.* (2009) compared the effectiveness of hypercube model with respect to data obtained from discrete simulation and concluded that the results between the methods are almost negligible but run times of methods are significantly different from each other. Morabito *et al.* (2007) suggests that iterative methods like Gauss-Siedel can solve the linear system with good performance where the problem size requires hundreds of thousands equations. Furthermore, if the exact solution is not a strict restriction, approximation of hypercube model can achieve good results in significant times. The hypercube approximation suggested by Jarvis (1985) can solve the problem in polynomial times and only requires *N* linear equations for problem size *N*. Still the requirement of Markovian property should not be omitted for hypercube model. The application of discrete time simulation on the other hand is relatively easy and can be obtained in reasonable computational times. Wu (2009) studied simulation on static deployment of ambulances.

CHAPTER 3

MATHEMATICAL MODELS AND SIMULATION MODULE

In this section of our study, we give basic notions and background information for each evaluated model. We also describe the proposed methodology and dispatching policies for discrete time simulation in the second part of this chapter.

3.1 Mathematical Models

3.1.1. Maximal Coverage Location Model

In the proposed single period model, the objective is to maximize the total weighted coverage of the population where the coverage criteria is determined as reaching from node *i* to node *j* with respect to predetermined average velocity within t_1 time units. The model is known as Maximal Coverage Location Model (MCLM) and the complexity of the MCLM is stated as NP-hard by Berman and Krass (2002). The model assumes that if a node is covered with a single facility then the whole population of the demand point is also covered. The total number of ambulances that can be assigned is limited; the model is as follows:

Notation:

M set of all demand points

N set of all possible candidate supply points

K maximum number of facilities that will be opened

t_1 time value that an ambulance need to reach in order to cover successfully (in minutes)

p_i population of node i where $i \epsilon M$

$$a_{ji} = \begin{cases} 1, if\ an\ ambulance\ at\ node\ j\ can\ cover\ node\ i \\ 0, otherwise \end{cases}$$

Decision Variables:

$$x_j = \begin{cases} 1, if\ a\ station\ is\ located\ at\ node\ j\ \ j\epsilon N \\ 0, otherwise \end{cases}$$

$$y_i = \begin{cases} 1, if\ demand\ point\ i\ is\ covered\ \ i\epsilon M \\ 0, otherwise \end{cases}$$

MCLM

Maximize $$z = \sum_{i\epsilon M} y_i * p_i \qquad (3.1)$$

Subject to $$\sum_{j\epsilon N} x_j = K \qquad (3.2)$$

$$\sum_{j\epsilon N} a_{ji} * x_j \geq y_i \qquad \forall i\epsilon M \tag{3.3}$$

$$x_j \in \{0,1\} \tag{3.4}$$

$$y_i \in \{0,1\} \tag{3.5}$$

The objective of the model is to maximize the total coverage of population. Constraint (3.2) enforces that the total number of ambulances that can be allocated is limited. Constraint (3.3) makes sure that if an ambulance is allocated at a candidate supply node then any demand node that can be reached within t_1 time units from this node should be assumed as covered. Constraints (3.4) and (3.5) show that all allocation and coverage decision variables are binary. Moreover, it should be noted that $K \geq 1$ in order to have a positive coverage.

3.1.2 Single Period Backup Double Covering Model

Secondly, the proposed model Single Period Backup Double Covering Model which is fundamentally a variant of MCLM is evaluated. In this model, the objective function is similar to MCLM but due to the nature of coverage decision variable the nodes should be covered two times within the determined t_1 and t_2 time units respectively. The main aim in this model is to propose a back up station in order to provide another alternative where the closest ambulance to the nodes is busy. Basically any demand point should be covered two times within t_1 and t_2 time units ($t_1 \leq t_2$) in order to be assumed as double covered. The model was originally proposed by Çatay *et. al* (2007) as follows:

Notation:

M set of all demand points

N set of all possible candidate supply points

K maximum number of facilities that will be opened

V average velocity that an ambulance will cover in an hour

t_1 time value that an ambulance need to reach in order to cover successfully (in minutes)

t_2 time value that an ambulance need to reach in order to cover second time successfully (in minutes)

p_i population of node i where $i\epsilon M$

$$a_{ji} = \begin{cases} 1, \text{if an ambulance at node } j \text{ can cover node } i \text{ in } t_1 \text{ time units with } V \\ 0, \text{otherwise} \end{cases}$$

$b_{ji} = \begin{cases} 1, \text{if an ambulance at node } j \text{ can cover node } i \text{ in } t_2 \text{ time units with } V \\ 0, \text{otherwise} \end{cases}$

Decision Variables:

$x_j = \begin{cases} 1, \text{if a station is located at node } j \quad j \epsilon N \\ 0, \text{otherwise} \end{cases}$

$y_i = \begin{cases} 1, \text{if demand point } i \text{ is double covered} \quad i \epsilon M \\ 0, \text{otherwise} \end{cases}$

SPBDCM

Maximize $$z = \sum_{i \epsilon M} y_i * p_i \qquad (3.6)$$

Subject to $$\sum_{j \epsilon N} x_j = K \qquad (3.7)$$

$$\sum_{j \epsilon N} a_{ji} * x_j \geq y_i \qquad \forall i \epsilon M \qquad (3.8)$$

$$\sum_{j \epsilon N} b_{ji} * x_j \geq 2 y_i \qquad \forall i \epsilon M \qquad (3.9)$$

$$x_j \in \{0,1\} \qquad (3.10)$$

$$y_i \in \{0,1\} \qquad (3.11)$$

The objective of the model is to maximize the total double covered population where the whole population is assumed to be covered like MCLM if demand node is covered. Constraint (3.7) ensures that the total number of ambulances that can be allocated should be equal to K where in order to get a positive coverage $K \geq 2$. Constraint (3.8) enforces that any demand point should be covered within t_1 time units in order to provide multiple coverage. Since $t_1 \leq t_2$ it is straightforward that each node will be covered second time by the same station within t_2 time units. Constraint (3.9) on the other hand imposes that each demand point should be covered by at least two different ambulances within t_2 time units. Constraints (3.10) and (3.11) require that the decision variables are binary.

This model is also stated as NP-hard and the main difference of SPBDCM and MCLM basically originates from the definition of decision variable y_i's by Basar (2008). Also intuitively when both models are simulated under the same parameters SPBDCM should have better performance measures since when the closest ambulance to a demand point is busy; another back-up station is allocated to provide the required service.

3.1.3. Maximum Service Restricted Set Covering Location Model

Thirdly, we introduce Maximum Service Restricted Set Covering Location Model (MSRSCLM) which is a revised version of MCLM as well. In this model, we allow both allocation and coverage decision variables to assume integer values rather than binary values. After the discussion with 112 Command Center about their allocation policy of ambulances, they informed us that according to the emergency system regulations an ambulance should serve at most 50000 people. However if the total population is too high, it is almost impossible to reach the levels obtained by MCLM and SPBDCM. So we propose a different constraint in order to cope with this condition. By adapting the idea of maximum population restriction to MCLM, we develop a new model MSRSCLM. In this order though, since we cannot control which ambulance serves to which calls, we cannot introduce an assignment variable to the model. To cope with this conflict, simply by adding the coverage constraint for each demand point, we relocate the ambulances. The new constraint imposes that each nodes' weighted coverage should be smaller than a constant value which is derived as the average population that an ambulance should serve. The constant is calculated as $(\sum_{i \epsilon M} p_i / K)$ and the new constraint is formed as $\sum_{j \epsilon L} a_{jk} * x_j * p_k \leq \sum_{i \epsilon M} p_i / K)$ $\forall k \epsilon M$ which is essentially the weighted coverage for each demand node should be smaller than the average value. The approach is counter intuitive because covering the highly populated demand nodes should be rewarded rather than being penalized. However, with this setting since the introduced coverage decision variables' range is non-negative integers, the objective function will force to cover the coverable demand points as much as possible.

Notation:

C average population value that an ambulance should cover $(\sum_{i \epsilon M} p_i / K))$

Decision Variables:

x_j ={number of ambulances allocated at node i, $i \,\epsilon N$

y_i ={number of coverages for node j, $j \epsilon M$

Model 1

Maximize $$z = \sum_{i \epsilon M} y_i * p_i \tag{3.12}$$

Subject to $$\sum_{j \epsilon N} x_j = K \tag{3.13}$$

$$\sum_{j\epsilon N} a_{ji} * x_j \geq y_i \qquad \forall i\epsilon M \qquad (3.14)$$

$$\sum_{j\epsilon N} a_{ji} * x_j * p_i \leq C \qquad \forall i\epsilon M \qquad (3.15)$$

$$x_j \in N^+ \qquad (3.16)$$

$$y_i \in N^+ \qquad (3.17)$$

We calculate that an ambulance should serve at most "C" unit of population where "C" is derived as average "fair" population that an ambulance should serve. The total population of the problem is divided by the total number of ambulances and constraint (3.15) is included in this model. Since the number of times that a node is covered for highly populated nodes should be rewarded, we relax the coverage decision variable to be an integer rather than binary. The main drawback of this model is that if a node has higher population than C value the model will force not to cover this node since left hand side of (3.15) will be higher than "C" value.

The objective function (3.12) of the proposed model is different than MCLM where the aim is to maximize weighted number of covered locations. Constraint (3.13) enforces that total number of ambulances that can be allocated should be equal to K. Constraint (3.14) controls the coverage decision of each node and constraints (3.16) and (3.17) determine the range of the decision variables. The problem is a revised version of MCLM, so this problem is also NP-hard..

The proof of complexity is straight-forward, consider a special case of MSRSCLM where "C" value is sufficiently large that for each demand point k the constraint is satisfied. Then the model reduces to MCLM which is as Berman and Krass (2002) stated, NP-hard.

3.1.4. Centralized Final Ratio Model

Finally we introduce the Centralized Final Ratio Model which considers that the number of times a node is covered should be proportional to the population of each demand point. The rational for this approach is the fact that the probability that any demand point will be the origin of a call should be proportional to its population. The ratio of a demand point's population to the total population of the problem is the likelihood of the demand point will be the origin of the next call. With this condition, we can assume that the expected number of calls for each demand point can be calculated. However the average number of calls for any demand point cannot be implemented in a mathematical model. The number of coverage required for a demand node should be smaller than or greater than some ratio in order to form a constraint. In this regard, we propose a maximum coverage number for each demand point since the objective function of the model is a maximization function; we do not want to accumulate all of the ambulances to the highest populated demand point. With this setting the problem turns into a revised version of a Knapsack problem and MCLM. The determination of coverage number for each node is not as simple as it seems. If the constraint turns out to be very tight then the setting will lead to a sub optimal result. On the other hand if the constraint is too loose then objective function will force the allocation of ambulances to cover highly populated demand nodes as much as possible. The trade-off between these decisions raises the question of how to determine the maximum coverage number for each demand node in order to obtain the optimal result.

We propose three ratios that should be included for the determination process. In order to simplify the approach of the forming process of the constraint, we propose some of the abbreviations and notations beforehand as follows:

p_k population of node k, $k \in M$

Q_s total population of the system; $\sum_{k\epsilon M} p_k$

π_j virtual weight of node j (centralized weight) ; $\pi_{j=\sum_{k\epsilon L} a_{jk} * p_k}\ \forall j \in M$

φ_j centralized region for node j (the nodes can be covered by node j)

W_s total centralized population of the system; $W_s = \sum_{k\epsilon M} \pi_k$

$\overline{W_s}$ average of centralized system

u_j ratio of π_j to W_s ; π_j / W_s

z_j ratio of π_j to Q_s ; π_j / Q_s

$$a_{jk} = \begin{cases} 1, \text{if an ambulance at node } j \text{ can cover node } k\ (d_{jk} \leq 3333.3 \text{ meters}) \\ 0, \text{otherwise} \end{cases}$$

Since the number of calls on average is proportional to the weight of the node, the expected number of calls for each node is determined. The ratio of p_k to Q_s gives the average percentage of the total calls originated from node k. By multiplying these ratios with the total expected number of calls, we can determine the average number of calls for each node. However, if a node is expected to originate 5 calls we cannot simply determine the number of coverage needed for this node as 1 or 2. In order to lower the busy probabilities, we need to consider the dispatching policies. If a call is originated from a covered node than the closest non-busy ambulance will be dispatched for policy

1; where for policy 2 whether the node is covered or not an ambulance is dispatched within 60 km range. So the expected number of calls for each node does not give a conclusive result for coverage constraint. To cope with this conflict, we propose a centralized weight calculation for each node π_j, for this purpose. As Goldberg (2004) stated aggregating the weights of demand points should represent the average workload of the region more precisely. The π_j value for each node is determined as $\pi_{j=\sum_{k\epsilon L} a_{jk} * p_k \forall j \in M}$ where a_{jk} value shows the coverage vicinity of the fixed node. Basically if an ambulance is allocated to a fixed node, we determine which nodes will be serviced from this ambulance when there is no other ambulance is allocated in this setting. After π_j for each node is determined, we recalculate the average number of calls that will be originated from φ_j by $z_j * E(\,total\ number\ of\ calls)$. Then we find the average required ambulance for each φ_j by dividing each $z_j * E(\,total\ number\ of\ calls)$ to average service rate. By taking the average of the new calculated ratios, we find the average workload of the new system as ratio1. Since most of the time the nodes will be counted multiple times, we simply cannot estimate the exact number of ambulances required to fulfill the demand for each φ_j. However, we propose an alternative approach to estimate the number of coverage needed for each node. On average the service rate of an ambulance is calculated and if expected service time for each ambulance is around 45 minutes, we derived that an ambulance can expectedly serve to 32 calls a day. The expected number of calls for φ_j is divided by the average centralized service rate to find ratio2. In order to calculate average centralized service rate, we calculate the u_j of each φ_j by dividing the π_j to W_s. The ratio gives the virtual effect of each φ_j to overall centralized system. Then the $\overline{W_s}$ value is calculated in order to assign that the virtual service rate of an ambulance for average φ_j as 32 calls a day. With this respect each ratio of u_j to $\overline{W_s}$ gives the inverse of service rate multiplying coefficient of an ambulance for φ_j. For example if φ_j's ratio to $\overline{W_s}$ is 2 then the virtual service rate for φ_j is determined as 16; or if the ratio is 0.1 then the average service rate is assumed to be 320. After the finalized service rates are calculated, expected number of calls for each φ_j is divided to virtual service rates which finalize ratio2. Finally in order to tighten the bound for right hand side of the constraint we added *K*-1/*K* where *K* represents the total number of ambulances for example if *K* value is determined as 100 then 0.99 is added to each ratio, this ratio is named as ratio3. The reason for adding 0.99

to each ratio is the following; if the ratio is smaller than 0.01 which basically is lower than the average workload of an ambulance; we should not cover these nodes more than they need. In order to eliminate the difference between 2.001 and 2.2 ratios where basically both of them will be covered at most twice, we add this ratio to cover φ_j with ratio 2.2 at most 3 times. The pseudo-code of the preprocessing of the centralized final ratio model is given in figure 1.

Step 1: Find the coverage matrix by deriving a_{jk} values from distance matrix

Step 2: Calculate the π_j values for each node with the help of coverage matrix

Step 2.a: Calculate the percentage of expected number of calls for each φ_j by dividing π_j of each node to Q_s of problem as z_j.

Step 2.b: Multiply z_j with expected number of calls and find the average number of calls expected from φ_j and divide with these numbers to average service rate.

Step 2.c: Find the average required number of ambulances from 2.b in order to find the average workload of the new system. Then take the average of the values obtained in step 2.b for all φ_j and set this average as ratio1.

Step 3: Find the u_j value of each φ_j with respect W_s

Step 3.a: Find the inverse multiplying coefficient for each φ_j by dividing the u_j value of each φ_j to $\overline{W_s}$.

Step 3.b: Set the average virtual service rate of new problem as default (number of calls that an ambulance can serve) where $u_j = 1$.

Step 4: Divide each u_j to $\overline{W_s}$ in order to calculate virtual service rate inverse multiplying coefficient for each centralized region.

Step 5: Find ratio2 as expected number of calls from each φ_j divided by virtual service rate.

Step 6: Find the ratio3 by (K-1)/K where K corresponds to total number of ambulances that can be allocated.

Step 7: Set the left hand side of the coverage constraint as $\sum_{j\epsilon M} a_{jk} * x_j$ for number of coverage for each node.

Step 8: Set the right hand side of the coverage constraint as summation of ratio 1,2 and 3. The sign of the constraint is smaller than equal to and this restriction should be applied for each demand node k.

Figure 1 - Pseudo code of forming the coverage constraint

Notation:

π_j virtual weight of node j (centralized weight) ; $\pi_{j=\sum_{k\epsilon L} a_{jk} * p_k} \; \forall j \in M$

ω_j ratio1 + ratio2 + ratio3

$$a_{jk} = \begin{cases} 1, \text{if an ambulance can cover demand node } k \text{ from point } j \\ 0, \text{otherwise} \end{cases}$$

Decision Variables:

x_j = {number of stations located at node j, $j\epsilon$N

y_i = {number of coverages for node i, $i\epsilon$M

Maximize	$z = \sum_{i\epsilon M} y_i * \pi_i$		(3.18)
Subject to	$\sum_{j\epsilon N} a_{ji} * x_j \geq y_i$	$\forall i\epsilon M$	(3.19)
	$\sum_{j\epsilon N} x_j = K$		(3.20)
	$\sum_{j\epsilon N} a_{ji} * x_j \leq \omega_i$	$\forall i\epsilon M$	(3.21)
	$x_j \in N^+$		(3.22)
	$y_i \in N^+$		(3.23)

The objective function of this model is to maximize the multiple coverage of aggregated weight. Constraint (3.19) enforces that the number of times that any demand point "*k*" is covered is related with the number of ambulances that can serve to this node. Total number of ambulances that can be allocated for the problem is restricted to *K* with constraint (3.20). For each demand point, the maximum number of times that a node can be covered should be smaller than the obtained ratio for that node is satisfied with constraint (3.21). Constraints (3.22) and (3.23) ensure that coverage and location variables can take non negative integer values.

3.2 Simulation Module

The suggested deterministic models in the literature have the following major drawback. Typically, these models have a decision variable that indicates whether or not a region is covered. They do not take into consideration what happens when all vehicles at a certain station are busy when an emergency call arrives. Usually the deterministic models aim to maximize a certain type of total coverage of the population by restricting the total number of ambulances which are fundamentally revised versions of SCLP. However, the underlying objective of these models should be maximizing the service level of the system rather than maximum coverage. The service level of the system could be evaluated by different methods like queuing models such as Hypercube Model presented by Larson (1974) or with simulation models. Since the problem size is too large for Istanbul or any other highly populated city, the Markov property should be taken into account and calculating the transition probabilities and states of the problem will become computationally hard. For example, in our study Istanbul has 867 districts with distinct weights and Hypercube model will require 2^{867} steady state equations in order to calculate the busy probabilities where in real life the dispatching rules may conflict with the Markov property.

Since Hypercube Model is computationally hard, we focus on other alternatives like simulation. By manipulating the VBA feature of MS Excel, we implement a simulation model for the selected models, in order to estimate busy probabilities and evaluate if the maximization of covered population is enough. The MCLM and SPBDCM are the main focus of the study and by solving their mathematical models optimally; we determine the locations of ambulances. The distance matrix for Istanbul is taken from the study of Basar (2008) and the population of each district is requested from Istanbul Metropolitan Municipality which was last obtained in year 2008. After the results of the models are obtained, the general setting of the simulation model is developed with 2 different policies.

In the simulation models, the first policy is a first come first serve based policy by allowing lost calls. Basically if a call is originated from a district, the closest covering non-busy ambulance is dispatched, however if the node is not covered at the moment of the call then this call is assumed to be lost. The nodes have basically uniform discrete probability distribution for originating a call where the population of each node corresponds to their weight. The interarrival times between calls and their distribution assumed to be exponential. The call is created by generating a random variable and by controlling the inverse of cumulative function value of each node, the origin of the call is determined. The service time of a dispatched ambulance also assumed to be exponential in our initial setting.

On the other hand, the second policy forces the system to dispatch the closest ambulance whether the node is covered or not, even though intuitively the second policy should lower the non serviced time, there are two main drawbacks like dispatching critical ambulances which cover highly populated districts and serving more calls compared to policy 1. By dispatching these important ambulances, it can lead to higher non serviced time in the long run.

The pseudo-code of the policy 1 and policy 2 for the initial setting is shown in figure 2. Even though some of the steps are reevaluated and changed, the general approach stays the same.

Step 1: Find the locations of the ambulances by solving the mathematical model

Step 2: Generate an exponentially distributed random variable for the next interarrival time

Step 2.a: Generate another random variable and by using inverse function to find the origin of the call by checking cumulative function of weights.

Step 2.b (policy 1): Check whether the node is covered or not if the node is covered, apply a greedy algorithm to find the closest non-busy ambulance location, otherwise go to step 2

Step 2.b (policy 2): Apply a greedy algorithm to find the closest non-busy ambulance location maximum distance that an ambulance can be assigned should be smaller than 60km

Step 3: After determining the location of the ambulance to be assigned, update the coverage matrix and mark the ambulance as busy

Step 3.a: Create another exponentially distributed random variable for service time

Step 3.b: Store the return time and location of ambulance

Step 4: Check the master clock and decide the next events type as either arrival or return

Step 5: If the event type is return update the coverage matrix and busy conditions of the ambulances

Step 5.b: If the event type is arrival apply from step2 through step4

Step 6: Run the previous steps from step 2 to step 5b until the master clock completes a full day simulation.

Step 7: Calculate the overall non served time of each node and ambulance for 24 hours.

Step 8: Apply from step1 through step7 for a total of 10 runs and calculate the average and standard deviation values of the non serviced ratio for each node.

Figure 2 - The algorithm of policy 1 (if the call is not covered, the call is lost) and policy 2 (dispatch the closest idle ambulance even if the node is not covered)

The general approach and underlying methodology of policy 2 is similar to policy 1 but when the call origin is not covered then another non-busy closest ambulance is assigned to the call. However the maximum distance that an ambulance can cover should be set in this policy because of the outlier nodes of Istanbul. For example, if a call comes from a district in Şile and the closest non-busy ambulance is in Sarıyer then it should not be assigned. The maximum distance that greedy algorithm will make the search is within the determined distance of 60 km since most of the time when an uncovered node originates a call the search within reasonable distance like 15 km turns as empty.

Exponentially distributed random variables like interarrival times and service times are derived by inverse transformation technique. Basically a random number between 0 and 1 generated in order to represent a uniform distribution. By taking -ln function of this random variate and dividing to mean of the exponential distribution we can generate new exponentially distributed random variables.

Some of the studies in the literature also consider that the service process can be expressed as the sum of other sub processes. As Savas (1969) suggested, the sequence of events during a call is composed of sub processes. Decomposition of the service process into sub processes like initialization & dispatching, first response, service to closest hospital and finally the travel time to original location gave a more realistic model of the overall process. A close observation of the whole dispatching, service and arrival processes in 112 Command Center made it clear that the decomposition of the service process is crucial. Improvement in any of these operations will reduce the overall service time, hence it will lead to more successful service levels. All of the sub processes except the first initialization process is exponentially distributed in our study. Whenever a call reaches to Emergency Command Center, headquarter determines a non-busy ambulance closest to incident and orders a dispatch; the initialization & dispatching process is based on a distance related function. The average travel speed of an ambulance is taken as constant throughout the day and assumed as 40km/hour. First aid and emergency response operation (first response operation) consists of a general check-up of the injured and questioning the accident's basic details which will require 5 minutes on average. After the injured is loaded to ambulance, the ambulance departs for closest hospital and this process also takes about 15 minutes on average. Finally after the injured is transferred successfully to a hospital, our model assumes that the ambulance is still busy until it returns to its original location. Therefore a traveling back process is included and the distribution of the process is exponential with mean 15 minutes.

CHAPTER 4

SIMULATION OF MODELS FOR ISTANBUL DATA

After the policies for the simulation process are determined, we determined the simulation parameters and mathematical model parameters for Istanbul data. Firstly, we conduct sensitivity analysis for MCLM and SPBDCM in order to determine the total number of ambulances and time restrictions. We acquired the current number of ambulances from Istanbul Emergency Response Unit (112 Command Center) as 117. On the other hand, for time variables we took 5 minutes for single coverage and for double back up coverage the time values are determined 5 and 8 minutes from Başar (2008) as default.

4.1 Sensitivity Analysis

4.1.1 MCLM Sensitivity Analysis

In our study, we analyzed several scenarios where the amount of resources is limited. When the available resources are high enough to yield fully covered scenarios, the location planning of ambulances would be simple. Furthermore, in order to represent the effect of uncovered nodes on simulation performance, we applied sensitivity analysis where the overall coverage is in reasonable percentages. When we use 5 minutes and 117 ambulances, the overall coverage for MCLM is found as 99.20% of the whole population. By incremental changes on each parameter one by one 100% coverage is reached. Even though the changes are relatively small 100% coverage could not be reached until total number of ambulances is 160 and time parameter is 8 minutes. The main reason for this condition is basically originated from the distribution of Istanbul population. To cover the rural areas and remote districts with low populations required extra ambulances.

The sensitivity analysis led to some decisions like dividing Istanbul's population into urban and rural districts then simulate with these conditions. Some of the studies in US take this condition and apply different parameters for each region separately. Grossman *et al.* (1997) stated that there is a significant difference in the response times between rural and urban areas. Furthermore, some of the studies originated from this conclusion consider only the rural areas and try to increase the overall efficiency of the system especially in circumstances which require fast response times like trauma and heart related incidents.

Rather than dividing the population, we neglected the remaining uncovered population since it is relatively small which is only 0.797% of the whole population for 117 ambulance and 5 minutes reaching time setting. Also, when the total number of expected calls in a day is considered, only 6 calls are missed with these parameters. Since the average increase of coverage with sensitivity analysis is 2.11E-05%, we concluded that our setting could represent reality with lower total number of ambulances. Figure 3 shows the incremental change in objective function value with the sensitivity analysis.

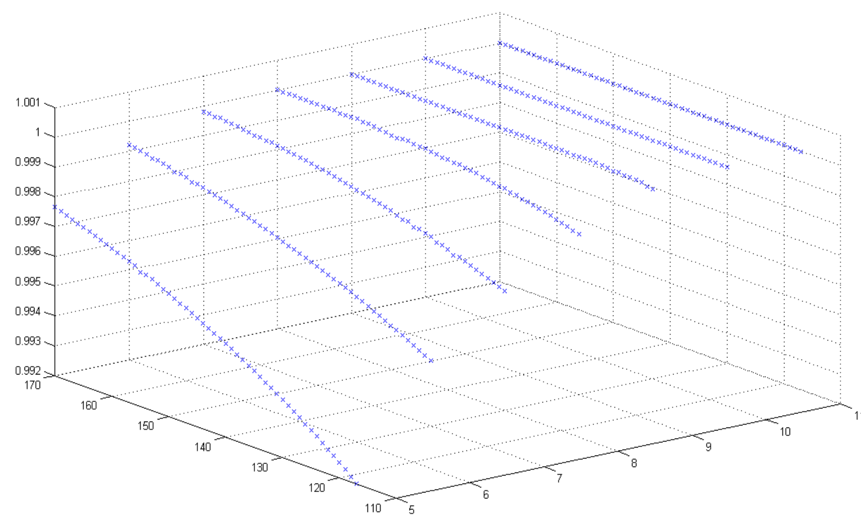

Figure 3 - The change in coverage percentage of MCLM with respect to ambulance and time variables

After setting the reaching time variable as 5 minutes and total number of ambulances as 100, the testing parameters of simulation is finalized. The detailed sensitivity analysis for single coverage is given in Appendix A for further illustration. Table 1 shows the default, maximum and final values for the analysis.

	t_1	# of ambulances	total coverage	Overall Percentage
Default	5	117	11,707,095	99.20%
Finalized	5	100	11,665,058	98.84%
Setting 1	8	160	11,801,244	100%
Setting 2	11	117	11,801,244	100%

Table 1 - Sensitivity Analysis of MCLM for Istanbul Data

In setting 1, we calculate the minimum reaching time parameter when the upper bound for the ambulances is set to 170. For smaller values of time variable like 5, 6 or 7 minutes the total coverage of the problem is not satisfied even with 170 ambulances. Also for setting 2, when we keep the ambulance number as constant, the minimum required reach time for full coverage is 11 minutes.

4.1.2 SPBDCM Sensitivity Analysis

The location and coverage decision variables for SPBDCM are binary however by relaxing this restriction and allowing location decision variable to take integer values, we can find upper bounds for the model. By relaxing the location decision variable as integer we find a better setting for the model. However, the nature of the model forces that the maximum number of ambulances that can be allocated to a node is 2. Since t_1 and t_2 time units ($t_1 \leq t_2$) restriction if any node is covered within t_1 time units, another allocated ambulance will also cover these nodes within t_2 time units. This will force the relaxed version of the model to assign at most 2 distinct ambulances to each possible supply node. Even though the objective function value of the relaxed model is better than SPBDCM, it is mandatory to test both under the same simulation parameters in order to conclude that relaxed version of this model has better performance measures.

Furthermore, the applied sensitivity analysis for MCLM should also be applied to SPBDCM in order to find the appropriate parameters for time and number of ambulances. As default the setting from Başar (2008) is applied for time variables and the total number of ambulances is also assumed 117 like MCLM. In this sensitivity analysis, the minimum difference between time units is fixed as 3 minutes and the maximum value of t_2 is set to 16 minutes where maximum value of t_1 is 11 minutes. The range of the number of ambulances is similar to MCLM sensitivity analysis and varies between 117 and 170.

Even though full coverage is reached in MCLM, SPBDCM cannot reach this level with these parameters. The incremental changes with respect to each variable and weighted coverage values are given in APPENDIX A. The average increase of the coverage percentage when t_1 and total number of ambulances are fixed is 0.101%. When t_1 and t_2 parameters are fixed the average increase in the total coverage percentage is 4.67E-05%. Finally, when t_2 and the total number of ambulances parameters are fixed the average increase is 0.1%. When we compare the average magnitudes, we can conclude that the effects of t_1 and t_2 parameters to coverage percentage are similar and more effective than the changes in the total number of ambulances.

Setting number of ambulances to 100, 117 and 170 the sensitivity analysis is derived when time parameters' minimum and maximum values are considered separately. In table 2, we show the changes in OFV with respect to total number of ambulances when t_1 and t_2 values are set their minimum.

t_1	t_2	# of ambulances	OFV	coverage percentage
5	8	100	11,052,331	93.65%
5	8	117	11,564,816	97.99%
5	8	170	11,601,292	98.30%

Table 2 - OFV and percentage for minimum t_1 and t_2 values

Even though the changes in OFV between 117 ambulances and 170 ambulances for fixed minimum t_1 and t_2 values is relatively small, we can observe that if the total number of ambulances is set to 100, there is a significant decrease in OFV. We conducted the same analysis for minimum t_1 and maximum t_2 values in table 3. In table 4, for maximum t_1 and minimum t_2 values the effect of total number of ambulances on OFV is shown. Finally, in table 5 for maximum t_1 and t_2 values the analysis is conducted.

t_1	t_2	# of ambulances	OFV	coverage percentage
5	16	100	11,060,375	93.72%
5	16	117	11,697,163	99.11%
5	16	170	11,769,195	99.72%

Table 3 - OFV and percentage for minimum t_1 and maximum t_2 values

t_1	t_2	# of ambulances	OFV	coverage percentage
11	16	100	11,524,985	97.65%
11	16	117	11,792,683	99.96%
11	16	170	11,796,881	99.99%

Table 4 - OFV and percentage for maximum t1 and minimum t2 values

t_1	t_2	# of ambulances	OFV	coverage percentage
11	14	100	11,487,948	97.34%
11	14	117	11,789,367	99.89%
11	14	170	11,790,223	99.90%

Table 5 - OFV and percentage for maximum t_1 and t_2 values

4.2. Simulation Results for Istanbul Data

In order to clarify the terms and their meanings that will be used throughout the study, we give an example on a simple instance. In this instance, there are 2 ambulances that will be located and 5 demand points. Demand point 1 can cover demand node 2, demand point 4 can only be covered by demand point 3 besides itself. Demand point 5 on the other hand can only be covered by itself. The population of each demand node is 300, 250, 280, 400, 100 for demand points 1 through 5 respectively. After the mathematical model of MCLM is solved there are 4 optimal results. Basically one of the ambulances is located either in demand node 1 or 2 for covering these points and the other ambulance is located either in demand node 3 or 4 in order to cover these demand points. Demand point 5 is not covered in initial setting. The obtained result for location decision variables that the illustration will be evaluated is {1,0,1,0,0} and thus the coverage decision variables are {1,1,1,1,0} for demand points 1 through 5 respectively. The day is assumed to be 24 time units, which yields 10 events to represent the general behavior of the system during a whole day simulation which can be shown in table 6.

					Coverage of the Nodes				
Time	Event Type	Next Interrarrival Time	Call Origin	Service + Return Time	1	2	3	4	5
0	N/A	2	N/A	N/A	1	1	1	1	0
2	Arrival	4	Node 2	8	0	1	1	1	0
6	Arrival	5	Node 1	N/A	0	1	1	1	0
10	Return	N/A	N/A	N/A	1	1	1	1	0
11	Arrival	3	Node 4	7	1	1	0	1	0
14	Arrival	3	Node 1	8	1	0	0	1	0
17	Arrival	4	Node 5	No available ambulance	1	0	0	1	0
18	Return	N/A	N/A	N/A	1	0	1	1	0
21	Arrival	5	Node 3	5	1	0	1	0	0
22	Return	1	N/A	N/A	1	1	1	0	0
24	Completion	N/A	N/A	N/A	1	1	1	0	0

Table 6 - Illustration of simulation process for policy 1 on dummy instance

The time variable keeps the master clock of the system and the simulation starts at time 0. Next interarrival time is generated at time 0 as 2 thus the next event will be at time 2 as an arrival process. When an arrival process occurs, immediately the origin of the call is determined; in this case node 2 is the origin. Since node 2 is covered at that time, the closest covering ambulance is dispatched to answer the call. The closest ambulance is located at node 1 and the generated service time is 8 for this instance. When the ambulance located at node 1 is dispatched to serve the call originated from node 2, the coverage condition of the nodes is updated. Node 1 is no longer covered during the service time of the assigned ambulance. Furthermore if there would be more nodes that are covered solely by ambulance at node 1 then all of the nodes will be uncovered during service time. The next process is again an arrival process since the return times of the busy ambulances is later than the next arrival time. At time 6 the call is originated from node 1 however node 1 is uncovered at that time so the call is lost and coverage conditions stayed the same. At time 10 the ambulance originally located at node 1 completes its service and returns to its location and coverage conditions are updated. At time 11 the call is originated from node 4 and this node is covered by the ambulance located at node 3. The ambulance is dispatched for the call and its service time is generated as 7, which yields node 3 to be uncovered during 7 time units. At time 14 the call is originated from node 1 and since the closest covering ambulance is located at node 1, node 2 is uncovered for 8 time units. At time 17 there are no available ambulances to be dispatched so the call is lost whether the call origin is covered or not. At time 18 the ambulance located at node 3 completes its service; hence the coverage conditions are updated. At time 21 the call is originated from node 3 which is covered by the ambulance located at the same location, the coverage of node 4 is updated for this event. At time 22 the simulation is terminated because there is not another event until time is 26. The return of ambulance located at node 1 is the final event and the coverage conditions are updated respectively.

Node 1 is not able to be served for 8 time units like node 2. Node 3 is not served for 7 time units where node 4 is only not served for 3 time units. However node 5 is not covered in initial setting so it is straightforward that for policy 1, it will be never covered during simulation. Non serviced ratio of the nodes are calculated as (total non served duration) / (simulation duration) which yields 33.33%, 33.33%, 29.17%, 12.50%, 100% for nodes 1 through 5 respectively. The total population of the setting is 1330 people, however due to the initial setting and dispatches; the system cannot serve to whole population. The population that is never covered is determined as non covered population and derived as population of node 5. Furthermore the population of node 5, 100 is also the lower bound for this system. Since the simulation includes time variable, non served population hour term is introduced. Basically non served population hour includes the non served ratio the node and the population of the node variables. For example for node 1, the non served ratio is 0.33 and the population is 300 which yields 100 people hour non served per day. For overall system 414.175 people hour can be derived by including each node. However the system has 1330 people as total population and if all the nodes are covered during the simulation at any time, it is expected to serve 1330 people hour. Furthermore, due to policy 1 dispatch rule, 100 people is never served thus determined as lower bound for non serviced population hour, named as non covered population hour. When the behavior of the system is analyzed, non serviced percentage (NSP) term is introduced and total non served population hour is divided to whole population hour hence for this system calculated as 31.14%. Throughout the study NSP is interchangeably used with performance of the system.

For policy 1, non serviced percentage and non served population hour values of the evaluated models for Istanbul data are given in table 7. On the other hand table 8 shows the performance metric of the models for policy 2 for Istanbul data.

	SPBDCM	MCLM	MSRSCLM	CFRM
Total Population	11,801,244	11,801,244	11,801,244	11,801,244
Non Covered Population Hour	264,570	136,186	1,514,965	1,583,065
Total Non Served Population Hour	3,501,017	3,782,759	2,580,020	1,747,931
Non Served Population Hour due to Simulation	3,236,447	3,646,572	1,065,055	164,866
Non Served Percentage (NSP)	29.67%	32.05%	21.86%	14.81%
Lower bound for NSP	2.24%	1.15%	12.84%	13.41%

Table 7 - Performance of Models for Istanbul Data for policy 1

	SPBDCM	MCLM	MSRSCLM	CFRM
Total Population	11,801,244	11,801,244	11,801,244	11,801,244
Non Covered Population Hour	264,570	136,186	1,514,965	1,583,065
Total Non Served Population Hour	5,085,586	5,620,874	3,248,974	1,904,210
Non Served Population Hour due to Simulation	4,821,016	5,484,688	1,734,009	321,145
Non Served Percentage (NSP)	43.09%	47.63%	27.53%	16.14%
NSP % difference wrt policy 1	13.43%	15.58%	14.69%	2.72%

Table 8 - Performance of Models for Istanbul Data for policy 2

4.2.1 Maximum Coverage Location Model

4.2.1.1 Policy 1

For single coverage model or MCLM the parameters are set as 100 for total number of ambulances, 5 minutes for reaching time. After the mathematical model is solved optimally, the location of the ambulances and the initial coverage of each node are determined. Since during the simulation, the ambulances return to their originally assigned location; some of the nodes in Istanbul are never covered. However, for this setting the coverage percentage is 98.84% which in terms of weighted coverage is quite satisfying. Even though the weighted coverage level is satisfactory, there are 178 nodes that are never covered. The remaining 689 nodes will originate almost 791 calls on average; only 9 calls on average will be lost by default setting. Furthermore, the probability that one of the 178 nodes that are never covered will be the origin of a call is at most 0.024%.

After 10 runs of a whole day simulation with divided service time and time dependent arrival rates, policy 1 for MCLM gives on average 3,782,758.8 people that are not served in a day. When the lower bound for the service level is considered, MCLM policy 1 can lead to at most 136,186 non served population hour. The main reason that MCLM policy 1 gives high NSP score is due to the setting of the model, whenever a call is originated from an initially covered node, the closest ambulance serves to this node is dispatched and leaves the remaining nodes uncovered.

4.2.1.2 Policy 2

For MCLM, policy 2 is expected to have higher NSP score, since the total number of calls that will be accepted for policy 2 is higher than policy 1. Also when policy 1 and policy 2 are compared, we have not considered any penalty for lost calls for policy 1. Fundamentally, policy 2 is a lower bound for each model where the penalty of a missed call varies with respect to dispatched ambulance and the population that is covered by the dispatched ambulance during its service time. In this regard, the simulation for MCLM policy 2 is conducted for 10 day simulation like policy 1.

4.2.2 SPBDCM

4.2.2.1 Policy 1

For SPBDCM, on the other hand, the locations of the ambulances are determined in such a way that if a call is originated from any node that is covered then there is another back up ambulance to cover this node during the dispatched ambulance service time. However, with 100 ambulances and 5 and 8 minutes reaching time, the total double coverage of the Istanbul data is 11,536,674 people. With these numbers we can calculate the lower bound for SPBDCM policy 1 by simply subtracting the serviced population from the whole population. Furthermore the lower bound for SPBDCM policy is 264,570 non served population hour, is still worse than the lower bound for MCLM policy 1 which is 136,186.4 non served population hour. However we cannot simply determine the lower bounds for each model as a performance metric because the objective functions of the models are not the same. For MCLM the objective function is to maximize the single covered population whereas for SPBDCM anode is assumed to be covered if the node is covered by two different ambulances. The obtained lower bounds for models are recalculated as if a node can be serviced by any number of ambulances then it is assumed to be covered, in other words each lower bound and the coverage of nodes during simulation is based on single coverage.

4.2.2.2 Policy 2

Even though, for MCLM policy 1 the NSP scores are higher than SPBDCM policy 1 we cannot derive that this condition will also hold for policy 2. Since there is a back up covering ambulance for any covered node in SPBDCM, the results can vary with respect to each dispatched ambulance. For example if a call is originated from a highly populated region of Istanbul and one of the covering ambulance for this region is still in service, the dispatch of back up ambulance for an out of reach dispatch will yield high NSP score. However, since this situation can also occur for MCLM, we had to simulate both of the models under the same parameters. After the simulation of SPBDCM, the results obtained from table 8 are still better than MCLM policy 2 and also gap between policy 1 and policy 2 for SPBDCM is lower than MCLM. In this regard, we can conclude that locating ambulances with respect to SPBDCM will give better results than MCLM.

After MCLM and SPBDCM have been simulated for Istanbul data, the NSP scores seemed higher than expected. In order to lower NSP scores, we propose new models where the location of the ambulances should consider the dispatching policies. The intuition behind the proposed models are based on two different approaches, maximum serviceable population for an ambulance without implementing any assignment constraint and centralized regions by considering the overall possible number of calls that an ambulance could be sent for dispatch. The mathematical models of the proposed approaches are stated in chapter 3 in detail.

4.2.3 Maximum Service Restricted Set Covering Location Model (MSRSCLM)

4.2.3.1 Policy 1

Even though the lower bound for MSRSCLM is worse than both MCLM and SPBDCM with 1,514,965 non served population hour, the overall performance of this setting is better than both of the models. Since highly populated districts are never covered for policy 1, during the simulation remaining nodes' non served ratio affected the overall performance for this model. The overall non served population hour for MSRSCLM is 2,580,020.3 on average after 10 days of simulation. The reason for the improved result for this model is rather than focusing on maximum coverage of highly populated districts, the ambulances are allocated to the remaining nodes more adequately in order to reduce their non served ratio. However, this model will be crippled if the outlier population values are observed more frequently. The distribution of population for Istanbul data can be fit to skewed normal distribution but for the distributions with skewness value that is not close to 0, the model will allocate the ambulances to districts with average population more frequently. Furthermore, since the number of coverage for each district is limited in order to satisfy the restriction, the model will allocate the remaining ambulances unnecessarily which will yield higher NSP score for this setting.

When MSRSCLM is compared to SPBDCM with respect to lower bounds when the population value is considered, MSRSCLM is worse by 1,250,395 people per day which corresponds to 472.61% worse than SPBDCM. However, when the

the overall NSP score of the system is in comparison with MSRSCLM is better by 920,996.26 people per day which corresponds to new model is better by 26.3% than SPBDCM. The % difference between MSRSCLM and MCLM with respect to never served population is larger than MSRSCLM vs SPBDCM, 1,378,778.6 people per day. Even though lower bound for MSRSCLM is worse than MCLM by 1012.42%, the overall performance of the new model is better by 31.8%. The comparisons conclude that MSRSCLM works better than both SPBDCM and MCLM when policy 1 for each model is considered. After each models' simulations are completed the ranking is SPBDCM is better than MCLM by 7.45% and MSRSCLM is better than SPBDCM by 26.3%.

4.2.3.2 Policy 2

The overall performance of MSRSCLM policy 2 is better than both MCLM and SPBDCM policy 1 with 27.53% overall NSP. The immediate question arose from this result is how come MSRSCLM can have better performance when the total number of calls that an ambulance is dispatched is higher than the average calls for policy 1. The answer again lies in the population distribution of Istanbul, most of the time the demand points with outlier population values are not accumulated in one region. There is always a relatively less busy ambulance allocated outside these regions which can be dispatched for the calls originated from highly populated districts.

4.2.4 Centralized Final Ratio Model (CFRM)

When we observed that MSRSCLM works better than the proposed models by literature, we realized in order to reduce overall NSP score of the system, the number of coverage for each node should be proportional to their weights. Since the expected number of calls from any demand point is proportional to the population of the node, covering more than twice seemed more realistic and intuitively would work better than previous models. However, since we cannot control the dispatching ambulance for each call, we cannot assign any of the ambulances to specific nodes. Furthermore, the total number of expected calls that an ambulance could possibly cover should be included into this approach. In this regard, we calculated the covering nodes for each node which will lead to possible workload of an ambulance. After each nodes' neighbors are found, the centralized weight for each node is calculated which basically corresponds to the overall population of the node and its neighbors. In order to be counted as a neighbor of the node the distance between nodes should be smaller than the coverage distance that is 3333.3 meters; derived from 5 minutes of reach in time and 40km/hour for average velocity. The construction of the new regions and included ratios was given in detail with the mathematical model in chapter 3.

4.2.4.1 Policy 1

The locations of ambulances obtained from CFRM initially cover 86.59% of the whole population when single coverage is concerned. The lower bound for CFRM is obtained from this ratio as 1,583,065 people non served per day that will be never served during policy 1. However when we apply the policy 1 simulation for CFRM, overall NSP score on average is estimated as 14.81% which means 1,747,930.7 people not served per day. The average overall NSP score of the new setting is better than MSRSCLM by 32.25% and is actually better than SPBDCM by %50 and better than MCLM by 53.8%. The comparison of CFRM with the rest of the models' policy 1 is as follows.

We can observe that from table 7, for CFRM the lower bound of the model is the worst among alternatives but overall performance for policy 1 is the best among them. The centralized population approach could not serve only 164,865.7 people hour due to simulation. Furthermore, we can derive that if the total number of ambulances for each policy is higher than 100 then CFRM performance will increase since the lower bound for policy 1 will be lowered drastically.

4.2.4.2 Policy 2

For the overall performance of MSRSCLM policy 2, we expressed that it is better than both MCLM and SPBDCM policy 1 overall performance. For CFRM policy 2 on the other hand is even better than MSRSCLM policy 1 in regard of NSP scores. With 1,904,209.9 not served population hour value, CFRM sets a new level for performance metric. 16.13% NSP of CFRM is far better than the alternatives. Also the % difference between policy 1 and policy 2 for CFRM is lower than all alternatives with only 2.72% which we can conclude that the setting for CFRM is more robust when it is compared to the other models.

4. 3. Sensitivity Analysis of SPBDCM and Relaxed SPBDCM

After the simulation for proposed models for both of the policies have been conducted, the relaxed version of SPBDCM for location decision variables is solved by mathematical model. The nature of SPBDCM enforces that maximum number of ambulances that can be allocated to any supply node to at most 2. However with the new setting the performance for policy 1 and policy 2 should be reevaluated. Even though the relaxed version of SPBDCM performs better, during the experimental study the relaxed version is not included. For Istanbul data, policy 1 and policy 2 comparisons with respect to original version of SPBDCM is as provided in table 9.

	SPBDCM	Relaxed SPBDCM
Total Population	11,801,244	11,801,244
Non Covered Population Hour	264,570	215,594
Total Non Served Population Hour	3,501,017	3,472,208
Non Served Population Hour due to Simulation	3,236,447	3,256,614
Non Served Percentage (NSP)	29.67%	29.42%
Lower bound for NSP	2.24%	1.83%

Table 9 - Comparison of SPBDCM and relaxed version for policy 1

For policy 1 overall NSP score of the relaxed version is slightly better after 10 day simulation however this difference is not conclusive for any argument. Furthermore, the lower bound for relaxed version is better than the original model which means that objective function value of relaxed version is improved when single coverage is concerned.

For policy 2, the performance of the relaxed version is still better but like policy 1 it is not conclusive to state relaxed version of SPBDCM works better with the same parameters. We can also observe from table 10 that gap between policy 1 and policy 2 for each setting is close to each other.

	SPBDCM	Relaxed SPBDCM
Total Population	11,801,244	11,801,244
Non Covered Population Hour	264,570	215,594
Total Non Served Population Hour	5,085,586	4,949,309
Non Served Population Hour due to Simulation	4,821,016	4,733,715
Non Served Percentage (NSP)	43.09%	41.94%
NSP % difference with respect to policy 1	13.43%	12.52%

Table 10 - Comparison of SPBDCM and Relaxed Version for Policy 2

CHAPTER 5

EXPERIMENTAL STUDY

Even though the models for Istanbul data give a general opinion about performance of various models, in order to test whether the performance of the models follow a similar pattern, we test the models on randomly generated data. It turns out, the results quite depend on the characteristics of the data. We randomly generate 240 instances with different problem size with respect to number of demand points, different layouts and different location and coverage values from each model. The arrival rate for Istanbul data is proportioned and a different type of arrival rate is also studied where the arrival rate is time dependent. The simulations are conducted for two different policies and for each parameter combination we have 10 independent 24 hours simulation runs.

5.1 Random Data Generation

For the experimental study, 4 different layouts have been implemented where each layout have their unique distribution of demand points. For each layout we propose a 50x50 km square region which is divided into various number of sub-regions according to layout design. For layout 1, we propose 4 different sub-regions which are basically dividing the whole region into 4 equal square regions. Each sub-region is populated with respect to determined problem size as 200, 300, 400 nodes. The coordinates of demand points have a uniform distribution and each demand point is populated with an exponentially distributed random variable with mean 10000. For each problem size, 5 instance have been generated thus a total of 15 instance have been created for each layout. For layout 1 each sub-region contains 25% of the demand points.

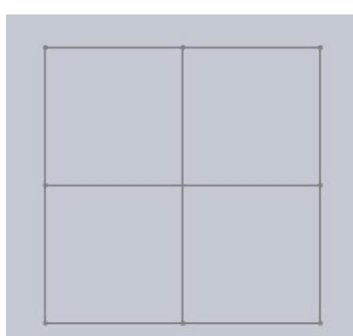

Figure 4 - The design of layout 1

For the second layout, we propose a similar layout to Istanbul data where the region is again divided into 4 sub-region however there is a 2*50 km area in the middle which is a representation of the Bosphorus. In this layout also the distribution of the nodes are different, sub-region 1 and 3 which are located on lower regions each contains the %35 of the whole demand points where sub-region 2 and 4 each contains %15. Figure 5 shows the design of layout 2.

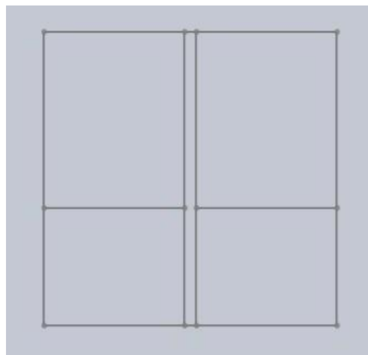

Figure 5 - Design of layout 2

For layout 3 we use one of the layouts proposed by Basar (2008) for our experimental study. The distribution of demand points in this setting is %25 in one corner of 50x50 region, another %25 in the opposite corner and remaining demand points are scattered throughout remaining area. The areas of corner sub-regions differ from Basar (2008). Figure 6 shows the design of layout 3.

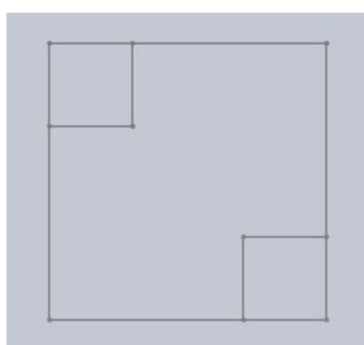

Figure 6 - Design of layout 3

For layout 4 we again use one of the layouts proposed by Basar (2008). The layout is divided into 4 squares where length of each squares' edge is proportional to inner squares' edge length. The middle square has 12.5 km edge length, the second layer square has 25 km edge length, the third layer square has 37.5 km edge length and final square is set to 50x50 km region. The sub-regions are determined as the remaining areas between each square and each region contains %25 of the demand points. The design of the layout 4 is given in figure 7.

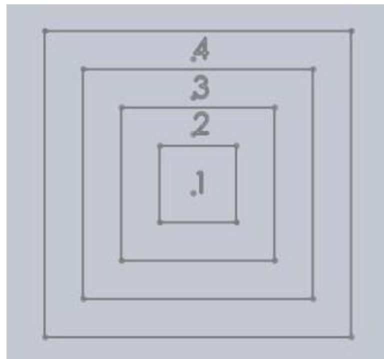

Figure 7 - The design of layout4 and corresponding sub-regions (1 represents region 1)

5.2 Results

For each model the NSP scores, the lower bounds and % difference these values have been calculated. We conduct our analysis for arrival rates proportional to expected population of the problem. For Istanbul data the overall expected number of calls is 800 and since the expected population of a demand point for experimental study is 10000 we established the overall expected number of calls for problem size 200 will be almost 20% of Istanbul data. Problem size 300 will have 240 expected calls and problem size 400 will be simulated with twice the arrival rate of problem size 200. Furthermore, the default arrival rate for each problem size is multiplied with 2 in order to create a higher arrival rate process. Moreover, the number of ambulances that can be allocated for each problem size is also proportional with each other, not proportional to Istanbul data though. The total number of ambulances is determined as 30, 45, and 60 for problem size 200, 300, and 400 respectively. For each model the comparisons consist of four different setting, policy 1 with low arrival rate, policy 1 with high arrival rate, policy 2 with low arrival rate and finally policy 2 with high arrival rate. For each setting, we compare the effectiveness of the model with respect to layout and size. Moreover, in order to find any relation between model and layout with respect to policy, we compare the layout averages' to each other for different problem size.

5.2.1 MCLM

Since the distribution of demand points' coordinates is uniform, the performance of MCLM improves compared to Istanbul data. The general performance of MCLM for policy 1 with respect to different population sizes is shown in table 11.

		MCLM							
		layout 1		layout 2		layout 3		layout 4	
		Low	High	Low	High	Low	High	Low	High
Size 200	instance1	28.84%	35.44%	26.59%	33.53%	23.67%	32.96%	24.65%	33.58%
	instance2	30.31%	36.12%	24.10%	30.44%	23.84%	34.24%	27.13%	34.38%
	instance3	30.41%	36.32%	23.74%	33.42%	23.14%	31.24%	24.13%	31.70%
	instance4	28.04%	34.98%	27.96%	34.99%	24.61%	34.53%	24.27%	33.03%
	instance5	27.43%	34.85%	27.66%	32.92%	21.58%	30.37%	24.32%	32.87%
	On average	29.01%	35.54%	26.01%	33.06%	23.37%	32.67%	24.90%	33.11%
Size 300	instance1	16.59%	19.36%	37.24%	44.17%	20.40%	26.69%	19.33%	25.01%
	instance2	43.23%	50.31%	12.76%	15.92%	14.10%	18.01%	39.72%	51.60%
	instance3	19.45%	24.17%	9.79%	12.54%	8.99%	11.27%	17.92%	22.23%
	instance4	40.84%	48.26%	21.12%	26.96%	37.25%	45.45%	22.21%	29.73%
	instance5	21.71%	25.40%	12.11%	14.48%	15.32%	19.13%	19.55%	24.50%
	On average	28.36%	33.50%	18.61%	22.81%	19.21%	24.11%	23.75%	30.61%
Size 400	instance1	13.29%	17.47%	21.27%	29.33%	14.25%	18.81%	17.04%	25.59%
	instance2	16.70%	21.29%	5.48%	7.96%	8.17%	12.51%	10.16%	14.89%
	instance3	5.04%	7.28%	20.01%	28.55%	13.01%	19.04%	26.04%	34.47%
	instance4	8.10%	10.87%	10.00%	16.20%	12.66%	15.90%	4.84%	7.01%
	instance5	25.28%	31.92%	10.08%	14.67%	22.24%	34.08%	13.73%	21.80%
	On average	13.68%	17.76%	13.37%	19.34%	14.07%	20.07%	14.36%	20.75%

Table 11 - Performance of MCLM for policy 1

When the layouts and their averages for each population size are examined, we can conclude that the increase in problem size is negatively related with the average NSP score of the layout. For example even though the average NSP score for layout 1 for problem size 300 instance 4 is higher than the overall average for layout 1 problem size 200, the average of layout 1 for size 300 is smaller than the average of layout 1 for size 200. However, the comparison among layouts is not conclusive. The best performing layout for different problem sizes vary. For problem size 200, layout 3 has the lowest average but for problem size 300 and 400, the average of layout 2 is better than layout 3.

When we compare the performance of MCLM with respect to layouts with low arrival rate, it does not provide a definite conclusion; since except size 400, where layout 1 average is the worst among all. Furthermore, we can observe that for high arrival rate policy 1 when the size of the problem increases, the average of the layout decreases which can be interpreted as negative relation between problem size and average of the layout. Moreover, some instances can affect the overall NSP score of the system significantly. Even though the instances are created with the same parameters, some of them do not reflect the general behavior exactly. For example for layout 4 when the problem size is 300, except the instance 2 the average is around 25% but with 51% NSP score the average of the layout 4 for size 300 increased to 30%. For size 300, instance 2 and instance 4 the overall NSP score of the system have outlier values for layout 1. Table 12 shows the performance of MCLM for policy 2.

		MCLM							
		layout 1		layout 2		layout 3		layout 4	
		Low	High	Low	High	Low	High	Low	High
Size 200	instance1	33.73%	45.78%	29.66%	42.82%	26.57%	40.47%	27.62%	45.55%
	instance2	33.62%	46.60%	28.31%	41.81%	27.18%	43.34%	30.99%	44.56%
	instance3	33.88%	46.74%	28.34%	42.38%	27.82%	44.62%	26.08%	41.32%
	instance4	30.97%	42.82%	32.18%	42.42%	28.54%	43.64%	26.37%	40.79%
	instance5	31.22%	42.60%	29.91%	42.21%	24.45%	40.77%	28.02%	43.08%
	On average	32.68%	44.91%	29.68%	42.33%	26.91%	42.57%	27.82%	43.06%
Size 300	instance1	19.32%	34.97%	16.23%	30.35%	15.25%	28.93%	17.39%	32.21%
	instance2	21.45%	39.20%	16.91%	30.75%	17.44%	31.97%	17.87%	30.58%
	instance3	18.38%	34.89%	16.36%	30.48%	16.62%	28.09%	16.32%	30.15%
	instance4	18.28%	31.67%	16.88%	30.43%	16.45%	30.62%	18.06%	35.61%
	instance5	19.24%	36.23%	15.32%	30.61%	18.29%	31.24%	14.52%	30.85%
	On average	19.33%	35.39%	16.34%	30.52%	16.81%	30.17%	16.83%	31.88%
Size 400	instance1	10.40%	17.89%	9.31%	18.04%	11.19%	20.48%	11.65%	18.18%
	instance2	12.95%	20.09%	9.32%	18.04%	11.27%	22.05%	11.87%	22.28%
	instance3	9.68%	17.59%	10.73%	19.09%	11.32%	21.08%	10.49%	19.43%
	instance4	9.72%	18.39%	9.09%	17.83%	11.74%	21.62%	10.97%	20.04%
	instance5	11.86%	19.57%	9.93%	19.93%	11.54%	20.95%	11.48%	19.97%
	On average	10.92%	18.71%	9.68%	18.58%	11.41%	21.24%	11.29%	19.98%

Table 12 - Performance of MCLM for policy 2

When we consider policy 2 for MCLM, intuitively from Istanbul data it is expected to have higher NSP scores for the same instances. However this is not the case, even though the system responds to more calls than policy 1, policy 2 with low arrival rate has lower NSP scores for most of the instances except when the problem size is 200. For Istanbul data when the performance of each model is observed, policy 2 has higher NSP scores than policy 1 although this condition does not hold for every instance. The reason behind this contradiction lies again in the distribution of demand points coordinates. For most of the instances the out of reach dispatch of an ambulance is not used as frequent as Istanbul data during the simulations. Basically policy 2 acts like policy 1 in most of the simulation duration.

For low arrival rate policy 2, we can conclude that the negative relation between layout averages and problem size is still valid. However for the outlier scenarios, policy 2 has less variant NSP scores when we compare with policy 1. Although for problem size 300 and layout 1, the instances 2 and 4 have outlier values for policy 1, policy 2 does not respond to these instances as much as policy 1.

For high arrival rate policy 2, as expected the NSP scores increased- compared to low arrival rate policy 2. Furthermore, when we compare the results obtained from high arrival rate policy 2, the distinct difference between policy 1 and policy 2 for low arrival rate diminishes. Even though the negative relation between layout average and problem size is maintained, for some of the instances policy 2 works better but for most of the instances policy 1 keeps its advantage. The robustness of policy 2 is sustained for high arrival rate. We can observe that for problem size 300 the instances 2 and 4 for layout 1 have outlier values for policy 1 but we cannot claim that these instances are also outlier for policy 2.

5.2.2 SPBDCM

SPBDCM like MCLM is firstly evaluated for each policy in order to observe if there are any relations. Between problem sizes and the average of the layouts there is negative relation for MCLM. However, even most of the layouts' averages decrease when the problem size increases; there is an exception for layout 1 as it can be seen from table 13.

		SPBCDM							
		layout 1		layout 2		layout 3		layout 4	
		Low	High	Low	High	Low	High	Low	High
Size 200	instance1	41.60%	46.46%	34.12%	40.30%	30.79%	36.26%	31.10%	37.01%
	instance2	39.02%	42.60%	34.09%	39.11%	30.32%	36.47%	32.85%	38.81%
	instance3	43.13%	46.69%	31.56%	37.73%	28.77%	34.99%	30.55%	38.32%
	instance4	39.41%	44.39%	34.62%	39.40%	33.73%	39.37%	30.82%	36.13%
	instance5	38.14%	43.10%	35.41%	40.48%	30.85%	35.19%	30.29%	35.58%
	On average	40.26%	44.65%	33.96%	39.41%	30.89%	36.46%	31.12%	37.17%
Size 300	instance1	21.82%	23.99%	45.96%	51.50%	29.97%	34.35%	25.28%	29.84%
	instance2	68.78%	73.64%	16.50%	19.53%	18.88%	21.94%	51.04%	61.34%
	instance3	30.33%	31.84%	12.87%	14.87%	11.32%	13.78%	24.24%	28.29%
	instance4	55.34%	61.94%	26.45%	31.68%	46.30%	52.04%	30.03%	33.65%
	instance5	32.40%	35.75%	15.94%	18.45%	19.83%	23.41%	24.27%	29.13%
	On average	41.74%	45.43%	23.54%	27.21%	25.26%	29.11%	30.97%	36.45%
Size 400	instance1	19.76%	24.20%	24.54%	30.73%	17.48%	20.72%	23.65%	29.02%
	instance2	27.75%	31.72%	7.37%	9.85%	10.85%	13.85%	12.33%	17.41%
	instance3	9.89%	11.82%	30.05%	36.66%	19.51%	25.38%	29.96%	39.73%
	instance4	13.63%	16.13%	17.24%	21.81%	15.41%	18.49%	6.78%	8.77%
	instance5	37.06%	42.53%	12.98%	16.46%	34.94%	44.23%	23.24%	29.62%
	On average	10.92%	25.28%	18.44%	23.10%	19.64%	24.53%	19.19%	24.91%

Table 13 - Performance of SPBDCM for policy 1

For problem size 300, instances 2 and 4 are known to be outliers but SPBDCM is very sensitive to these outlier instances. The NSP score of the system increases drastically compared to other instances of the same layout and problem size. Overall NSP score of these instances with 68% and 55% values are almost 2.5 times of the other instances' average. Although the other instances' scores decrease the layout average, layout 1's average for problem size 300 is still higher than problem size 200.

For high arrival rate policy 1, the conclusions for MCLM are still valid for SPBDCM. Furthermore, the outlier instances crippled the score of the model for high arrival rate. Even though layout 2 has the best layout average for problem size 300 and 400, layout 3's average is better than layout 2 for problem size 200. Like low arrival rate the layouts' averages decrease with respect to increase in problem size except layout 1 for problem size 300. The main reason for this contradiction is again originated from the outlier instances for problem size 300.

The policy 2 with low arrival rate however, still maintains its robustness for SPBCDM. The outlier instances of layout 1 for problem size 300 are coped with out of reach dispatches. Except problem size 200, policy 2 with low arrival rate has better performance than policy 1 with low arrival rate for the model. When most of the instances are considered the performance of policy 2 is better than policy 1. Table 14 shows the NSP scores for SPBDCM with respect to layout and problem size for policy 2.

		SPBCDM							
		layout 1		layout 2		layout 3		layout 4	
		Low	High	Low	High	Low	High	Low	High
Size 200	instance1	45.39%	54.80%	37.38%	48.20%	34.87%	45.30%	34.62%	45.59%
	instance2	41.55%	53.64%	38.68%	47.52%	33.55%	46.35%	34.33%	49.40%
	instance3	47.00%	56.30%	34.43%	45.18%	32.31%	46.72%	34.46%	45.60%
	instance4	44.27%	54.69%	38.24%	49.55%	36.79%	49.63%	34.79%	47.28%
	instance5	42.07%	51.19%	40.26%	50.74%	34.89%	44.31%	33.18%	46.38%
	On average	44.06%	54.12%	37.80%	48.24%	34.48%	46.46%	34.28%	46.85%
Size 300	instance1	26.86%	34.97%	22.15%	30.35%	21.03%	28.93%	23.24%	32.21%
	instance2	31.67%	39.20%	20.63%	30.75%	21.76%	31.97%	20.82%	30.58%
	instance3	27.22%	34.89%	21.75%	30.48%	18.85%	28.09%	21.37%	30.15%
	instance4	24.73%	31.67%	21.70%	30.43%	21.05%	30.62%	24.12%	35.61%
	instance5	28.69%	36.23%	20.42%	30.61%	21.20%	31.24%	21.47%	30.85%
	On average	27.84%	35.39%	21.33%	30.52%	20.78%	30.17%	22.21%	31.88%
Size 400	instance1	13.83%	20.23%	11.84%	18.82%	14.65%	21.05%	13.90%	21.60%
	instance2	17.30%	23.94%	12.21%	19.57%	12.60%	21.44%	13.17%	21.11%
	instance3	14.25%	19.90%	12.06%	20.00%	12.48%	20.66%	13.35%	21.17%
	instance4	13.62%	20.04%	11.97%	19.84%	13.50%	21.41%	13.25%	21.36%
	instance5	16.21%	21.99%	11.52%	19.60%	15.53%	24.58%	13.68%	21.55%
	On average	10.92%	21.22%	11.92%	19.57%	13.75%	21.83%	13.47%	21.36%

Table 14 - Performance of SPBDCM for policy 2

For policy 2 with low arrival rate, the negative relation between layout average and problem size can be observed easily. The performance of SPBDCM with respect to layout averages is again not definitive where for problem size 200 and 300 layout 3 has the best scores but for problem size 400 layout 2 has a better score than layout 3.

Except the increased NSP scores for each instance, policy 2 with high arrival rate represents all the characteristics of policy 2 with low arrival rate. The negative relation between problem size and layout average, the performance with respect to layout averages and most of the instances policy 2 with high arrival rate works better than policy 1 with high arrival rate are the main conclusions.

5.2.3 MSRSCLM

Among other alternatives, MSRSCLM is the most sensitive model with respect to demand points' coordinates distribution. Any outlier design can affect the performance of the model hugely. Due to this sensitivity, the negative relation between population size and layout average diminishes. Furthermore, there is still no indication of relation between model and any layout design. In table 15, the results for policy 1 is given.

		MSRSCLM							
		layout 1		layout 2		layout 3		layout 4	
		Low	High	Low	High	Low	High	Low	High
Size 200	instance1	48.19%	50.70%	47.06%	49.14%	42.54%	44.43%	45.63%	49.03%
	instance2	50.27%	52.86%	47.19%	49.31%	41.16%	46.86%	41.00%	44.91%
	instance3	49.47%	51.45%	43.19%	46.15%	41.50%	45.68%	39.80%	45.47%
	instance4	51.62%	54.27%	45.77%	47.67%	41.38%	43.86%	34.99%	38.98%
	instance5	46.88%	48.59%	49.41%	52.72%	44.83%	46.56%	41.48%	45.28%
	On average	49.29%	51.57%	46.52%	49.00%	42.28%	45.48%	40.58%	44.73%
Size 300	instance1	29.82%	30.60%	70.25%	73.54%	42.66%	44.44%	37.69%	38.79%
	instance2	86.86%	89.04%	24.47%	25.66%	26.55%	27.68%	71.25%	75.22%
	instance3	37.31%	38.89%	17.04%	18.26%	16.27%	17.48%	30.69%	32.29%
	instance4	77.35%	80.51%	41.13%	43.04%	66.54%	70.93%	41.91%	44.69%
	instance5	44.71%	45.83%	22.84%	24.07%	26.81%	27.62%	34.86%	36.09%
	On average	55.21%	56.97%	35.15%	36.91%	35.76%	37.63%	43.28%	45.41%
Size 400	instance1	39.42%	41.31%	51.50%	53.56%	33.27%	33.77%	46.81%	48.79%
	instance2	51.43%	53.59%	15.78%	16.33%	21.16%	22.06%	29.82%	31.09%
	instance3	19.00%	19.51%	51.69%	53.52%	40.81%	41.87%	72.18%	75.61%
	instance4	21.60%	22.71%	31.51%	32.83%	26.99%	28.30%	14.63%	15.17%
	instance5	55.56%	57.59%	26.68%	28.05%	60.20%	63.28%	44.33%	45.77%
	On average	10.92%	38.94%	35.44%	36.86%	36.49%	37.85%	41.55%	43.29%

Table 14 - Performance of MSRSCLM for policy 1 with low arrival rate

Likely for high arrival rate for policy 1, performance of MSRSCLM is decreased with the outlier instances and led the relation between problem size and layout average to vanish. For some of the instances, the overall NSP score raises the question how can this model work better than MCLM or SPBDCM for Istanbul data. The answer lies in the sensitivity of the model, for some instances like the highly populated demand nodes are dense in some regions, mathematical model of MSRSCLM assigns the ambulances out of coverage vicinity of these regions. Hence the overall NSP score is so high for policy 1.

In Table 16, for instance 2 of layout 1 with problem size 300 the performance of MSRSCLM is very disappointing with 89%. Even though there are some instances which have high NSP scores, there are also counter examples of these results like problem size 400 layout 1 instance 3 which is almost 50% of the layout average. However, even though this instance's score is better than others, it is still worse than score of this instance for both MCLM and SPBDCM.

		MSRSCLM							
		layout 1		layout 2		layout 3		layout 4	
		Low	High	Low	High	Low	High	Low	High
Size 200	instance1	50.53%	59.41%	50.40%	56.96%	46.02%	52.78%	49.25%	57.35%
	instance2	51.76%	58.81%	49.74%	57.25%	46.70%	59.04%	44.00%	52.81%
	instance3	52.94%	61.19%	44.53%	53.54%	45.11%	52.32%	44.23%	52.48%
	instance4	55.33%	62.81%	47.65%	56.97%	44.92%	51.54%	37.36%	45.21%
	instance5	48.48%	54.81%	53.15%	62.32%	46.45%	51.71%	45.19%	53.00%
	On average	51.81%	59.41%	49.09%	57.41%	45.84%	53.48%	44.01%	52.17%
Size 300	instance1	41.04%	46.26%	36.35%	42.31%	39.77%	44.37%	40.72%	44.88%
	instance2	45.46%	49.20%	33.69%	38.34%	37.45%	42.66%	34.30%	38.65%
	instance3	35.85%	40.87%	35.47%	41.50%	30.32%	36.39%	34.89%	40.90%
	instance4	34.29%	39.72%	39.96%	46.69%	34.64%	39.85%	36.57%	40.99%
	instance5	44.48%	49.54%	31.81%	38.36%	36.08%	40.64%	36.02%	41.05%
	On average	40.22%	45.12%	35.46%	41.44%	35.65%	40.78%	36.50%	41.30%
Size 400	instance1	32.22%	35.34%	29.81%	32.99%	35.59%	38.73%	30.70%	33.28%
	instance2	37.69%	41.97%	28.53%	32.13%	27.96%	31.73%	31.24%	35.78%
	instance3	34.03%	37.40%	25.94%	29.77%	31.04%	34.57%	34.02%	38.14%
	instance4	25.92%	30.18%	27.55%	30.70%	29.71%	32.93%	28.33%	31.61%
	instance5	32.94%	37.14%	25.74%	29.90%	27.07%	32.00%	26.39%	29.84%
	On average	10.92%	36.41%	27.51%	31.10%	30.28%	33.99%	30.14%	33.73%

Table 16 - Performance of MSRSCLM for policy 2 with low arrival rate

In Table 16, the results of MSRSCLM for policy 2 with respect to layout and problem size are compared. Despite the low performance of MSRSCLM for policy 1, with out of reach dispatches the outlier instances have been neutralized by policy 2. With the help of this neutralization, the negative relation between problem size and layout average is validated again. However, we cannot derive the conclusion about the performance of policy 2 is better than policy 1. Nevertheless for most of the instances policy 2 works better than policy 1. Equivalently the characteristics of policy 2 are held for high arrival rate.

5.2.4 CFRM

Even though, CFRM works best among other settings for Istanbul data, the uniformity of demand points decreases the performance of this model. Despite the uniformity, the robustness of the model should still yield better results than alternatives most of the time. The distinct effect of policy 2 should be reduced in this model since the construction of model includes the expected calls per region which reduces the ratio of out of reach dispatches. Furthermore, the negative relation between layout average and problem size is conserved despite of the arrival rate and applied policy. Still there is no indication of relation among layouts and model's performance which concludes that the performance of model is not based on the distribution of demand points. In Table 17, the NSP score for each instance for policy 1 is given.

		CFRM							
		layout 1		layout 2		layout 3		layout 4	
		Low	High	Low	High	Low	High	Low	High
Size 200	instance1	31.55%	37.18%	30.79%	37.40%	27.81%	34.37%	26.87%	33.74%
	instance2	32.05%	37.29%	27.14%	34.29%	26.08%	33.14%	28.74%	35.20%
	instance3	34.14%	38.48%	26.44%	33.13%	28.36%	34.93%	27.93%	34.80%
	instance4	31.15%	37.62%	31.88%	37.68%	25.80%	32.18%	30.03%	35.86%
	instance5	30.81%	37.25%	30.01%	36.72%	23.04%	31.61%	27.67%	36.16%
	On average	31.94%	37.56%	29.25%	35.84%	26.22%	33.25%	28.25%	35.15%
Size 300	instance1	15.50%	18.13%	38.58%	47.22%	22.05%	27.70%	18.59%	22.74%
	instance2	41.66%	49.07%	13.11%	15.90%	13.31%	16.94%	35.94%	45.19%
	instance3	17.96%	21.97%	10.83%	13.47%	9.11%	12.00%	15.84%	20.47%
	instance4	41.54%	49.51%	21.09%	26.59%	36.49%	44.17%	20.56%	26.66%
	instance5	18.71%	23.09%	11.72%	14.68%	15.49%	20.32%	21.58%	25.74%
	On average	27.07%	32.36%	19.07%	23.57%	19.29%	24.22%	22.50%	28.16%
Size 400	instance1	13.31%	18.49%	27.48%	35.11%	13.92%	18.93%	20.58%	27.74%
	instance2	17.55%	21.64%	5.87%	8.32%	12.44%	14.72%	11.06%	16.01%
	instance3	5.90%	7.94%	25.03%	32.82%	15.06%	22.06%	28.47%	37.24%
	instance4	8.49%	10.64%	13.11%	20.07%	12.67%	16.71%	5.08%	7.49%
	instance5	21.98%	27.62%	12.60%	16.59%	28.42%	36.56%	19.45%	24.78%
	On average	10.92%	17.27%	16.82%	22.58%	16.50%	21.80%	16.93%	22.65%

Table 17 - Performance of CFRM for policy 1 with low arrival rate

In Table 18, the results of simulation under policy 2 can be seen. Even though layout 2 average for problem size 200 and 300 is better than the other layouts, for problem size 400 the performance of CFRM with respect to layout averages is changed. There is still no relation between layout design and model's performance for either policy. The negative relation between layout average and problem size can be observed as expected due to the robustness of CFRM. The other characteristics of the experimental study are also valid for this setting.

		CFRM							
		layout 1		layout 2		layout 3		layout 4	
		Low	High	Low	High	Low	High	Low	High
Size 200	instance1	35.12%	49.04%	33.00%	47.10%	31.78%	43.63%	28.59%	45.20%
	instance2	35.09%	50.23%	29.57%	42.80%	28.96%	42.37%	30.07%	44.69%
	instance3	37.80%	47.69%	29.85%	41.82%	30.85%	47.58%	31.58%	44.11%
	instance4	33.77%	47.24%	34.88%	47.70%	29.87%	42.12%	32.79%	45.75%
	instance5	34.11%	45.86%	33.41%	45.89%	24.56%	42.62%	31.79%	47.99%
	On average	35.18%	48.01%	32.14%	45.06%	29.20%	43.66%	30.96%	45.55%
Size 300	instance1	24.91%	34.34%	21.13%	29.19%	17.46%	28.32%	21.63%	31.10%
	instance2	24.41%	33.86%	17.92%	26.75%	21.56%	30.99%	22.17%	30.35%
	instance3	20.86%	31.10%	21.06%	29.53%	20.18%	27.16%	19.07%	28.75%
	instance4	19.74%	28.65%	21.91%	30.44%	18.06%	30.49%	21.11%	30.12%
	instance5	23.40%	31.87%	20.36%	28.17%	19.83%	30.47%	22.24%	30.78%
	On average	22.67%	31.96%	20.48%	28.82%	19.42%	29.48%	21.24%	30.22%
Size 400	instance1	14.83%	21.33%	14.91%	23.22%	14.58%	23.57%	16.15%	24.42%
	instance2	17.92%	24.99%	12.79%	21.68%	16.69%	25.84%	14.62%	23.72%
	instance3	12.58%	18.64%	15.94%	24.30%	15.74%	25.99%	14.08%	23.68%
	instance4	12.09%	20.18%	13.00%	21.27%	17.96%	26.02%	11.83%	20.04%
	instance5	15.14%	21.99%	14.90%	21.84%	15.88%	23.05%	15.95%	23.58%
	On average	10.92%	21.42%	14.31%	22.46%	16.17%	24.89%	14.53%	23.09%

Table 15 - Performance of CFRM for policy 2 with low arrival rate

5.2.5 COMPARISONS OF MODELS

In this section of our study, we compare the overall score of each model with respect to population size, different policies and different arrival rates. Even though each simulation consists of 10 day simulation, we cannot derive that each score represents the exact behavior of the system definitely. However, we can interpret the general performance of the models according to these scores. For problem size 200 the scores of models for policy 1 is given in table 19.

	MCLM		SPBDCM		MSRSCLM		CFRM	
	Low	High	Low	High	Low	High	Low	High
instance1_layout1	33.73%	35.44%	45.39%	46.46%	50.53%	50.70%	35.12%	37.18%
instance1_layout2	29.66%	33.53%	37.38%	40.30%	50.40%	49.14%	33.00%	37.40%
instance1_layout3	26.57%	32.96%	34.87%	36.26%	46.02%	44.43%	31.78%	34.37%
instance1_layout4	27.62%	33.58%	34.62%	37.01%	49.25%	49.03%	28.59%	33.74%
instance2_layout1	33.62%	36.12%	41.55%	42.60%	51.76%	52.85%	35.09%	37.29%
instance2_layout2	28.30%	30.44%	38.68%	39.11%	49.74%	49.31%	29.57%	34.29%
instance2_layout3	27.18%	34.24%	33.55%	36.47%	46.70%	46.86%	28.96%	33.14%
instance2_layout4	30.99%	34.38%	34.33%	38.81%	44.00%	44.91%	30.07%	35.20%
instance3_layout1	33.88%	36.32%	47.00%	46.69%	52.94%	51.45%	37.80%	38.48%
instance3_layout2	28.34%	33.42%	34.43%	37.73%	44.53%	46.15%	29.85%	33.13%
instance3_layout3	27.82%	31.24%	32.31%	34.99%	45.11%	45.68%	30.85%	34.93%
instance3_layout4	26.08%	31.70%	34.46%	38.32%	44.23%	45.47%	31.58%	34.80%
instance4_layout1	30.97%	34.98%	44.27%	44.39%	55.33%	54.27%	33.77%	37.62%
instance4_layout2	32.18%	34.99%	38.23%	39.40%	47.65%	47.67%	34.88%	37.68%
instance4_layout3	28.54%	34.53%	36.79%	39.37%	44.92%	43.86%	29.87%	32.18%
instance4_layout4	26.37%	33.03%	34.79%	36.13%	37.36%	38.98%	32.79%	35.86%
instance5_layout1	31.22%	34.85%	42.07%	43.10%	48.48%	48.59%	34.11%	37.25%
instance5_layout2	29.91%	32.92%	40.26%	40.48%	53.15%	52.72%	33.41%	36.72%
instance5_layout3	24.45%	30.37%	34.89%	35.19%	46.45%	46.56%	24.56%	31.61%
instance5_layout4	28.02%	32.87%	33.18%	35.58%	45.19%	45.28%	31.79%	36.16%

Table 19 - Comparison of models for policy 1 for problem size 200

When we compare the NSP scores of the models, it can be seen from table 19 that for most of the instances MCLM works best. The general ranking is MCLM is best, CFRM's performance is close to MCLM, SPBDCM is 3^{rd} and MSRSCLM is 4^{th}. The ranking does not change except layout 4 instance 2 for low arrival rate; CFRM's score is better than the score of MCLM for this instance. The change in ranking with respect to Istanbul data is originated from the same reason throughout the experimental study, the distribution of demand points' coordinates. The average scores of the models are 29.27%, 37.65%, 47.68%, and 31.87% for MCLM, SPBDCM, MSRSCLM and CFRM relatively.

For high arrival rate policy 1 when the problem size is 200, the ranking of the models does not change at all. MCLM works best, followed by CFRM, SPBDCM, and MSRSCLM. Even though for high arrival rate, the arrival rate is twice of low arrival rate, the average of the NSP scores for each model increased around 5%. The most increase can be observed for CFRM with 6.5%.

For low arrival rate policy 2 when the problem size is 200, the ranking of the models change more frequently. Also when each models' performances are considered with respect to policies, policy 1 works better than policy 2 for every instance except problem size 200 with low arrival rate.

	MCLM		SPBDCM		MSRSCLM		CFRM	
	Low	High	Low	High	Low	High	Low	High
instance1_layout1	35.44%	45.78%	46.46%	54.80%	50.70%	59.41%	37.18%	49.04%
instance1_layout2	33.53%	42.82%	40.30%	48.20%	49.14%	56.96%	37.40%	47.10%
instance1_layout3	32.96%	40.47%	36.26%	45.30%	44.43%	52.78%	34.37%	43.63%
instance1_layout4	33.58%	45.55%	37.01%	45.59%	49.03%	57.35%	33.74%	45.20%
instance2_layout1	36.12%	46.60%	42.60%	53.64%	52.85%	58.81%	37.29%	50.23%
instance2_layout2	30.44%	41.81%	39.11%	47.52%	49.31%	57.25%	34.29%	42.80%
instance2_layout3	34.24%	43.34%	36.47%	46.35%	46.86%	59.04%	33.14%	42.37%
instance2_layout4	34.38%	44.56%	38.81%	49.40%	44.91%	52.81%	35.20%	44.69%
instance3_layout1	36.32%	46.74%	46.69%	56.30%	51.45%	61.19%	38.48%	47.69%
instance3_layout2	33.42%	42.38%	37.73%	45.18%	46.15%	53.54%	33.13%	41.82%
instance3_layout3	31.24%	44.62%	34.99%	46.72%	45.68%	52.32%	34.93%	47.58%
instance3_layout4	31.70%	41.32%	38.32%	45.60%	45.47%	52.48%	34.80%	44.11%
instance4_layout1	34.98%	42.82%	44.39%	54.69%	54.27%	62.81%	37.62%	47.24%
instance4_layout2	34.99%	42.42%	39.40%	49.55%	47.67%	56.97%	37.68%	47.70%
instance4_layout3	34.53%	43.64%	39.37%	49.63%	43.86%	51.54%	32.18%	42.12%
instance4_layout4	33.03%	40.79%	36.13%	47.28%	38.98%	45.21%	35.86%	45.75%
instance5_layout1	34.85%	42.60%	43.10%	51.19%	48.59%	54.81%	37.25%	45.86%
instance5_layout2	32.92%	42.21%	40.48%	50.74%	52.72%	62.32%	36.72%	45.89%
instance5_layout3	30.37%	40.77%	35.19%	44.31%	46.56%	51.71%	31.61%	42.62%
instance5_layout4	32.87%	43.08%	35.58%	46.38%	45.28%	53.00%	36.16%	47.99%

Table 20 - Comparison of models for policy 2 for problem size 200

The ranking in general is similar to policy 1, MCLM is best and followed by CFRM, SPBDCM and MSRSCLM. Although the general ranking is similar, the performances for instance 5 of layout 4 is quite interesting because SPBDCM works better than CFRM but still is behind MCLM. CFRM on the other hand works best in instance 3 of layout 2 and instance 4 of layout 3.

For high arrival rate, policy 2 works worse than policy 1 for each instance. Although the NSP scores are increased for high arrival rate, the frequency of change in the ranking also increased with respect to low arrival rate. For problem size 200 with high arrival rate, the performance of the models for policy 2 on the other hand the general ranking is similar to previous results. However CFRM beats MCLM in 3 instances; instance 1 of layout 4, instance 2 of layout 3, and finally instance 4 of layout 3. Moreover SPBDCM is better than CFRM for instance 5 of layout 4 but still behind the score of MCLM.

When problem size 300 is analyzed, with low arrival rate policy 1 performance of the models is very interesting. The ranking of the models in general is changed with this setting; CFRM works better than MCLM in 60% of the instances and for the averages of models CFRM is also better than MCLM. However the remaining ranking stays same SPBDCM is 3rd and MSRSCLM is in 4th order.

	MCLM		SPBDCM		MSRSCLM		CFRM	
	Low	High	Low	High	Low	High	Low	High
instance1_layout1	16.59%	19.36%	21.82%	23.99%	29.82%	30.60%	15.50%	18.13%
instance1_layout2	37.24%	44.17%	45.96%	51.50%	70.25%	73.54%	38.58%	47.22%
instance1_layout3	20.40%	26.69%	29.97%	34.35%	42.66%	44.44%	22.05%	27.70%
instance1_layout4	19.33%	25.01%	25.28%	29.84%	37.69%	38.79%	18.59%	22.73%
instance2_layout1	43.23%	50.31%	68.78%	73.64%	86.86%	89.04%	41.66%	49.07%
instance2_layout2	12.76%	15.92%	16.50%	19.53%	24.47%	25.66%	13.11%	15.90%
instance2_layout3	14.10%	18.01%	18.88%	21.94%	26.55%	27.68%	13.31%	16.94%
instance2_layout4	39.72%	51.60%	51.04%	61.34%	71.25%	75.22%	35.94%	45.19%
instance3_layout1	19.45%	24.17%	30.33%	31.84%	37.31%	38.89%	17.96%	21.97%
instance3_layout2	9.79%	12.54%	12.87%	14.87%	17.04%	18.26%	10.83%	13.47%
instance3_layout3	8.99%	11.27%	11.32%	13.78%	16.27%	17.48%	9.11%	12.00%
instance3_layout4	17.92%	22.23%	24.24%	28.29%	30.69%	32.29%	15.84%	20.47%
instance4_layout1	40.84%	48.26%	55.34%	61.94%	77.35%	80.51%	41.54%	49.51%
instance4_layout2	21.12%	26.96%	26.45%	31.68%	41.13%	43.04%	21.09%	26.59%
instance4_layout3	37.25%	45.45%	46.30%	52.04%	66.54%	70.93%	36.49%	44.17%
instance4_layout4	22.21%	29.73%	30.03%	33.65%	41.91%	44.69%	20.56%	26.66%
instance5_layout1	21.71%	25.40%	32.40%	35.75%	44.71%	45.83%	18.71%	23.09%
instance5_layout2	12.11%	14.48%	15.94%	18.45%	22.84%	24.07%	11.72%	14.68%
instance5_layout3	15.32%	19.13%	19.83%	23.41%	26.81%	27.62%	15.49%	20.32%
instance5_layout4	19.55%	24.50%	24.27%	29.13%	34.86%	36.09%	21.58%	25.74%

Table 21 - Comparison of models for policy 1 for problem size 300

As shown in Table 21 CFRM is worse than MCLM in only 8 instances, which shows that the uniformity of the demand points' coordinates is also not conclusive for determining the performance of a model without simulation. For high arrival rate for policy 1 when problem size is 300, CFRM still works better than MCLM in 60% of the instances. Also the instances that CFRM is better are not same with the instances that CFRM's score is better than the score of MCLM for low arrival rate. However, 10 of the instances that CFRM is better than MCLM are same for both low and high arrival rates. The ranking is similar to low arrival rate though for the remaining models. The average of CFRM is slightly better than MCLM but the difference is too small to state that CFRM works better, only 0.68% difference compared to overall NSP scores. The results obtained from policy 1 problem size 300 prove that even though MCLM works better for most of the cases, in order to determine a model's performance we have to use some analytical tool to evaluate.

For problem size 300, when policy 2 is evaluated however the dominance of CFRM to MCLM is lost in this setting. Furthermore SPBDCM's score is better than the score of CFRM in 4 instances. The general ranking is MCLM is the best; CFRM is 2nd and followed by SPBDCM and MSRSCLM.

	MCLM		SPBDCM		MSRSCLM		CFRM	
	Low	High	Low	High	Low	High	Low	High
instance1_layout1	19.32%	28.14%	26.86%	34.97%	41.04%	46.26%	24.91%	34.34%
instance1_layout2	16.23%	24.63%	22.15%	30.35%	36.35%	42.31%	21.13%	29.19%
instance1_layout3	15.25%	24.38%	21.03%	28.93%	39.77%	44.37%	17.46%	28.32%
instance1_layout4	17.39%	29.18%	23.24%	32.21%	40.72%	44.88%	21.63%	31.10%
instance2_layout1	21.45%	30.19%	31.67%	39.20%	45.46%	49.20%	24.41%	33.86%
instance2_layout2	16.91%	27.13%	20.63%	30.75%	33.69%	38.34%	17.92%	26.75%
instance2_layout3	17.44%	27.61%	21.76%	31.97%	37.45%	42.66%	21.56%	30.99%
instance2_layout4	17.87%	28.18%	20.82%	30.58%	34.30%	38.65%	22.17%	30.35%
instance3_layout1	18.38%	28.14%	27.22%	34.89%	35.85%	40.87%	20.86%	31.10%
instance3_layout2	16.36%	27.34%	21.75%	30.48%	35.47%	41.50%	21.06%	29.53%
instance3_layout3	16.62%	26.19%	18.85%	28.09%	30.32%	36.39%	20.18%	27.16%
instance3_layout4	16.32%	27.64%	21.37%	30.15%	34.89%	40.90%	19.07%	28.75%
instance4_layout1	18.28%	27.30%	24.73%	31.67%	34.29%	39.72%	19.74%	28.65%
instance4_layout2	16.88%	26.73%	21.70%	30.43%	39.96%	46.69%	21.91%	30.44%
instance4_layout3	16.45%	28.49%	21.05%	30.62%	34.64%	39.85%	18.06%	30.49%
instance4_layout4	18.06%	29.23%	24.12%	35.61%	36.57%	40.99%	21.11%	30.12%
instance5_layout1	19.24%	28.61%	28.69%	36.23%	44.48%	49.54%	23.40%	31.87%
instance5_layout2	15.32%	26.53%	20.42%	30.61%	31.81%	38.36%	20.36%	28.17%
instance5_layout3	18.29%	28.80%	21.20%	31.24%	36.08%	40.64%	19.83%	30.47%
instance5_layout4	14.52%	24.47%	21.47%	30.85%	36.02%	41.05%	22.24%	30.78%

Table 16 - Comparison of models for policy 2 for problem size 300

The average NSP scores of each model with low arrival rate for policy 2 for problem size 300 are 17.32% for MCLM, 23.03% for SPBDCM, 36.95% for MSRSCLM and 20.95% for CFRM which can also be calculated from table 22. For high arrival rate on the other hand, policy 2 results of the models are quite close to each other especially for CFRM and SPBDCM for most of the instances. CFRM works best in only one instance, instance 2 of layout 2. Although the results are close and CFRM beats MCLM in one instance, the ranking is still the same.

Finally for problem size 400, the models are compared with respect to low and high arrival rate for both policy 1 and policy 2. Even though there are some instances that the order of MCLM, CFRM, SPBDCM and MSRSCLM is not maintained, generally the performances of the models are similar to previous settings. In Table 23, the NSP scores of each model for policy 1 is given in detail.

	MCLM		SPBDCM		MSRSCLM		CFRM	
	Low	High	Low	High	Low	High	Low	High
instance1_layout1	13.29%	17.47%	19.76%	24.20%	39.42%	41.31%	13.31%	18.49%
instance1_layout2	21.27%	29.33%	24.54%	30.73%	51.50%	53.56%	27.48%	35.11%
instance1_layout3	14.25%	18.81%	17.48%	20.72%	33.27%	33.77%	13.92%	18.93%
instance1_layout4	17.03%	25.59%	23.65%	29.02%	46.81%	48.79%	20.58%	27.74%
instance2_layout1	16.70%	21.29%	27.74%	31.72%	51.43%	53.59%	17.55%	21.64%
instance2_layout2	5.48%	7.96%	7.37%	9.85%	15.78%	16.33%	5.87%	8.32%
instance2_layout3	8.17%	12.51%	10.85%	13.85%	21.16%	22.06%	12.44%	14.72%
instance2_layout4	10.16%	14.89%	12.33%	17.41%	29.82%	31.09%	11.06%	16.01%
instance3_layout1	5.04%	7.28%	9.89%	11.82%	19.00%	19.51%	5.90%	7.94%
instance3_layout2	20.01%	28.55%	30.05%	36.65%	51.69%	53.52%	25.03%	32.82%
instance3_layout3	13.01%	19.04%	19.51%	25.38%	40.81%	41.87%	15.06%	22.06%
instance3_layout4	26.04%	34.47%	29.96%	39.73%	72.18%	75.61%	28.47%	37.24%
instance4_layout1	8.10%	10.87%	13.63%	16.13%	21.60%	22.71%	8.49%	10.64%
instance4_layout2	10.00%	16.20%	17.24%	21.81%	31.51%	32.83%	13.11%	20.07%
instance4_layout3	12.66%	15.90%	15.41%	18.49%	26.99%	28.30%	12.67%	16.71%
instance4_layout4	4.84%	7.01%	6.78%	8.77%	14.63%	15.17%	5.08%	7.49%
instance5_layout1	25.28%	31.92%	37.06%	42.53%	55.56%	57.59%	21.98%	27.62%
instance5_layout2	10.08%	14.67%	12.98%	16.46%	26.68%	28.05%	12.60%	16.59%
instance5_layout3	22.24%	34.08%	34.94%	44.23%	60.20%	63.28%	28.42%	36.56%
instance5_layout4	13.73%	21.80%	23.24%	29.62%	44.33%	45.77%	19.45%	24.78%

Table 23 - Comparison of models for policy 1 for problem size 400

CFRM works better than MCLM as best performing model for instance 5 of layout 1 and instance 1 of layout 3. Moreover, rather than CFRM, SPBDCM works as second best performing model for instance 1 of layout 2 and instance 2 of layout 3. Also for some instances MSRSCLM performs almost 4 times worse compared to other models. Even though there are exceptional instances, the averages of the models are compatible with the ranking. For high arrival rate there are also some instances where SPBCDM performs better than CFRM and for only one instance

CFRM works as best setting. All of the other conclusions are still valid for high arrival rate for policy 1.

For policy 2 with low arrival rate on the other hand is quite intriguing because not only the ranking of the models changed, the robustness of policy 2 exceeds the other evaluated problem size. For most of the instances, under low arrival rate, policy 2 works better than policy 1 for problem size 400.

	MCLM		SPBDCM		MSRSCLM		CFRM	
	Low	High	Low	High	Low	High	Low	High
instance1_layout1	10.40%	17.89%	13.83%	20.23%	32.22%	35.34%	14.83%	21.33%
instance1_layout2	9.31%	18.04%	11.84%	18.82%	29.81%	32.99%	14.91%	23.22%
instance1_layout3	11.19%	20.48%	14.65%	21.05%	35.59%	38.73%	14.58%	23.57%
instance1_layout4	11.65%	18.18%	13.90%	21.60%	30.70%	33.28%	16.15%	24.42%
instance2_layout1	12.95%	20.09%	17.30%	23.94%	37.69%	41.97%	17.92%	24.99%
instance2_layout2	9.32%	18.04%	12.21%	19.57%	28.53%	32.13%	12.79%	21.68%
instance2_layout3	11.27%	22.05%	12.60%	21.44%	27.96%	31.73%	16.69%	25.84%
instance2_layout4	11.87%	22.28%	13.17%	21.11%	31.24%	35.78%	14.62%	23.72%
instance3_layout1	9.68%	17.59%	14.25%	19.90%	34.03%	37.40%	12.58%	18.64%
instance3_layout2	10.73%	19.09%	12.06%	20.00%	25.94%	29.77%	15.94%	24.30%
instance3_layout3	11.32%	21.08%	12.48%	20.66%	31.04%	34.57%	15.74%	25.99%
instance3_layout4	10.49%	19.43%	13.35%	21.17%	34.02%	38.14%	14.08%	23.68%
instance4_layout1	9.72%	18.39%	13.62%	20.04%	25.92%	30.18%	12.09%	20.18%
instance4_layout2	9.09%	17.83%	11.97%	19.84%	27.55%	30.70%	13.00%	21.27%
instance4_layout3	11.74%	21.62%	13.50%	21.41%	29.71%	32.93%	17.96%	26.02%
instance4_layout4	10.97%	20.04%	13.25%	21.36%	28.33%	31.61%	11.83%	20.04%
instance5_layout1	11.86%	19.57%	16.21%	21.99%	32.94%	37.14%	15.14%	21.99%
instance5_layout2	9.93%	19.93%	11.52%	19.60%	25.74%	29.90%	14.90%	21.84%
instance5_layout3	11.54%	20.95%	15.53%	24.58%	27.07%	32.00%	15.88%	23.05%
instance5_layout4	11.48%	19.97%	13.68%	21.55%	26.39%	29.84%	15.95%	23.58%

Table 17 - Comparison of models for policy 2 for problem size 400

In table 24, SPBDCM works better than CFRM 80% of the time. Also the average values for each model suggest that the best performing models are in order of MCLM, SPBCDM, CFRM and finally MSRSCLM. Furthermore policy 2 works better than policy 1 on 61.25% of the instances. The least sensitive model to robustness of policy 2 is CFRM which is originated from the construction of model.

The results of high arrival rate for policy 2 when the problem size is 400 the deviation increases. On 20% of the instances SPBDCM works as best performing model however for 4 instances SPBDCM is beaten by CFRM, but on average

SPBDCM is in second place in regard of efficiency. Moreover, the results conflict with the general assumption on overall performances of the models where usually MCLM works best and for some cases CFRM transcends the performance of MCLM. Even though the average of SPBDCM is worse than MCLM, 80% of the instances proved otherwise.

When we compare all of the results throughout the experimental study, we can conclude that even though the uniformity of the demand points' coordinates affects the performance of the models, without simulation we cannot state a definitive argument on models' performance. The negative relation between problem size and layout average is strong when the model can be stated as robust. If the model is affected from the outlier instances, this can lead to diminish of the relation. Neither the layout design, nor the distribution of demand points' coordinates are related with the performance of the model. For some instances model's performance can perform better than the alternatives. Even though the general ranking throughout the experimental study is sustained, the counter examples that are observed disproved a generalization. However, we can conclude that the uniformity of demand points' coordinates clearly affects the performance of the model. On the other hand, even policy 2 responds to more calls than policy 1 this condition is not conclusive to assume policy 2's NSP scores should be higher than policy 1's for every instance. The performance of policy 2 is weakly related with the construction of the model. If during the construction of model, the performance of the system is based on out of reach dispatches like MSRSCLM, policy 2 can transcend the performance of policy 1. The simulation results, lower bound and gap between these values of each model for different policy, different arrival rate with respect to problem size are given in APPENDIX B for further research.

5.3 Statistical Testing

When we analyzed the performance of the models with respect to different layouts, problem sizes, arrival rates and policies we extended our study by providing confidence intervals. The confidence intervals are derived for the difference between each models performance and the hypothesis testing is applied in order to statistically argue that a model is working better than another or not. The comparison of the models can be observed from table 25 and 26 where 1 represents first model is working better, 0 represents inconclusive for comparison and -1 represents the second model is working better.

		MCLM vs SPBDCM		MCLM vs MSRSCLM		MCLM vs CFRM	
		Low	High	Low	High	Low	High
Size 200 Policy 1	Lower Bound	2.99%	0.62%	11.93%	7.46%	2.76%	-1.55%
	Upper Bound	13.77%	11.03%	24.90%	20.74%	9.60%	5.26%
	Testing	1	1	1	1	1	0
Size 200 Policy 2	Lower Bound	0.62%	0.02%	7.46%	4.80%	7.62%	-1.84%
	Upper Bound	11.03%	11.38%	20.74%	20.00%	16.33%	6.55%
	Testing	1	1	1	1	1	0
Size 300 Policy 1	Lower Bound	-2.43%	-2.66%	-0.08%	-1.45%	-3.46%	-4.75%
	Upper Bound	18.23%	16.23%	39.82%	34.40%	2.47%	3.39%
	Testing	0	0	0	0	0	0
Size 300 Policy 2	Lower Bound	1.53%	0.53%	13.31%	7.80%	0.33%	-0.72%
	Upper Bound	9.88%	8.56%	25.95%	21.62%	6.91%	6.07%
	Testing	1	1	1	1	1	0
Size 400 Policy 1	Lower Bound	-1.03%	-1.25%	3.98%	2.52%	-2.89%	-2.51%
	Upper Bound	12.74%	11.19%	43.72%	36.99%	7.00%	5.70%
	Testing	0	0	1	1	0	0
Size 400 Policy 2	Lower Bound	0.51%	-1.45%	13.01%	7.02%	1.58%	0.18%
	Upper Bound	4.93%	4.18%	25.58%	21.34%	6.53%	6.50%
	Testing	1	0	1	1	1	1

Table 25- Confidence levels for the models

The confidence levels, upper and lower bounds are calculated by assuming the data is normally distributed and by applying t testing we derived the lower bounds for the comparisons as $\mu - \sigma * \theta$ where θ is calculated from normal table for 95% two sided testing which yields 1.96 value. However, for upper bound $\mu + \sigma * \theta$ equation is applied and if the interval includes *0* then it is inconclusive however if the interval includes only positive values than the first model is working better, finally if the interval includes only negative values than the second model's performance is better.

		SPBDCM vs MSRSCLM		SPBDCM vs CFRM		MSRSCLM vs CFRM	
		Low	High	Low	High	Low	High
Size 200 Policy 1	Lower Bound	4.02%	2.80%	-8.77%	-9.14%	-19.29%	-18.43%
	Upper Bound	16.05%	13.75%	4.37%	1.21%	-5.19%	-6.06%
	Testing	1	1	0	0	-1	-1
Size 200 Policy 2	Lower Bound	2.80%	-0.22%	-0.07%	-8.83%	-8.72%	-18.24%
	Upper Bound	13.75%	13.61%	12.38%	2.14%	4.47%	-1.85%
	Testing	1	0	0	0	0	-1
Size 300 Policy 1	Lower Bound	0.05%	-1.13%	-20.01%	-18.34%	-41.10%	-35.84%
	Upper Bound	23.89%	20.50%	3.22%	3.40%	0.37%	1.53%
	Testing	1	0	0	0	0	0
Size 300 Policy 2	Lower Bound	8.88%	4.59%	-6.94%	-5.41%	-22.15%	-17.12%
	Upper Bound	18.96%	15.74%	2.77%	1.67%	-9.86%	-6.96%
	Testing	1	1	0	0	-1	-1
Size 400 Policy 1	Lower Bound	1.61%	0.77%	-11.53%	-11.48%	-40.40%	-35.03%
	Upper Bound	34.38%	28.79%	3.93%	4.71%	-3.19%	-1.30%
	Testing	1	1	0	0	-1	-1
Size 400 Policy 2	Lower Bound	11.10%	6.46%	-2.47%	-2.05%	-22.30%	-18.32%
	Upper Bound	22.05%	19.16%	5.14%	6.00%	-8.18%	-3.36%
	Testing	1	1	0	0	-1	-1

Table 26- Confidence levels for the models

CHAPTER 6
MYOPIC HEURISTIC

After the experimental study for generated instances is concluded, we proposed a heuristic to improve the performance of models by reassigning ambulances. The purpose for this heuristic is to evaluate the possible improvement that can be made on models' initial settings. The proposed heuristic is applied on Istanbul data under policy 1. In each iteration of the heuristic the difference between previous iteration's overall NSP score and current iteration's NSP score are compared. Since the improvement in each iteration can be derived, we compare these results for each model to observe the possible improvement that can be made. The initial setting for each model is obtained from their mathematical models, however by relocating the ambulances the system converges to a better setting. Even though the mathematical models are solved optimally, the simulations demonstrated that locating ambulances optimally for the problem cannot be concluded as effective location planning. With this heuristic we emphasize that initial setting of the models can be improved. The % difference between improved result and initial setting indicates that the bigger the improvement how much worse the initial setting is.

In each step of the algorithm, the score of each ambulance is calculated. The score of an ambulance is composed of two values, aggregated weights and average busy probability of the ambulance. The busy probability of the ambulance is conducted from simulation and after 10 runs, the average is calculated. Despite reassigning the ambulance with highest average busy probability value, we consider the population affected from this probability. For example if the ambulance is busy for 85% of the day and serving 120000 people, then reassigning of this ambulance to another location will be resulted 18000 more non served people per day on average. However if the ambulance's busy probability is 70% and serving only 20000 people then reassignment of this ambulance will affect the overall performance by an additional 6000 not served people. Even though the calculation of busy probability is easy, we cannot derive the population that is served form the ambulance exactly. Like Centralized Final Ratio Model (CFRM), we suggest aggregated population to deal with this issue. Since the dispatching policy checks for the closest non-busy ambulance with in coverage distance, we can suggest that the possible workload of an ambulance should be considered. For deriving the aggregated population for each node, we transform the distance matrix into binary values. The transformation process controls if the demand node can be covered from another node. Basically if the distance between node "*k*" and node "*m*" is smaller than 3333.3 meters, node "*k*" can cover node "*m*" and vice versa. The calculated distance is based on 5 minutes reaching time and 40 km/hour for average travel speed. After the transformation, the aggregated population for each node is calculated. Then the ambulance that will be reassigned is designated, and by using the average non serviced ratio of each node, the new location of the ambulance is determined. After the reassignment, the setting is simulated and the new performance of the system is calculated. In each iteration of heuristic a single ambulance is determined for relocation and system is simulated with the new values, for a total of 10 iterations. If the heuristic cannot improve the previous iteration's performance the algorithm stops.

In Table 25, the performance of MCLM, lower bound for each iteration and improvement between each iteration is shown.

	MCLM			
	Non Served Population Hour	NSP	Lower bound	Improvement
Iteration 0	3,782,759	32.05%	1.15%	N/A
Iteration 1	3,664,672	31.05%	2.17%	1.00%
Iteration 2	3,574,672	30.29%	3.36%	0.76%
Iteration 3	3,483,463	29.52%	4.42%	0.77%
Iteration 4	3,398,671	28.80%	4.11%	0.72%
Iteration 5	3,308,932	28.04%	3.40%	0.76%
Iteration 6	3,266,879	27.68%	4.03%	0.36%
Iteration 7	3,189,571	27.03%	5.15%	0.66%
Iteration 8	3,132,591	26.54%	7.16%	0.48%
Iteration 9	3,082,885	26.12%	8.91%	0.42%
Iteration 10	3,037,956	25.74%	12.51%	0.38%

Table 25 - Performance metric for MCLM for 10 iterations

For iteration 0, the initial result obtained for MCLM for Istanbul data is used. For iteration 1, the algorithm finds the least busy ambulance and assigns to a new location then simulates the new setting. Since MCLM forces the maximum number of ambulances on any given node as 1, reassignment of an ambulance leads to higher lower bound level. However, the reassignment of the ambulance as expected lowers the non served percentage and non served population hour variables. The improvement for each iteration is calculated with respect to overall population hour. In iteration 4, algorithm finds the least busy ambulance at a node where there are 2 ambulances located in iteration 3. Removing one of the ambulances and assigning to another location thus yields a lower lower bound value. Furthermore, after 7 iterations the performance of the system does not improve as much as the first few iterations and converges to approximately 3,000,000 non served people hour. In figure 8, the convergence of non served population hour parameter for MCLM is shown in detail.

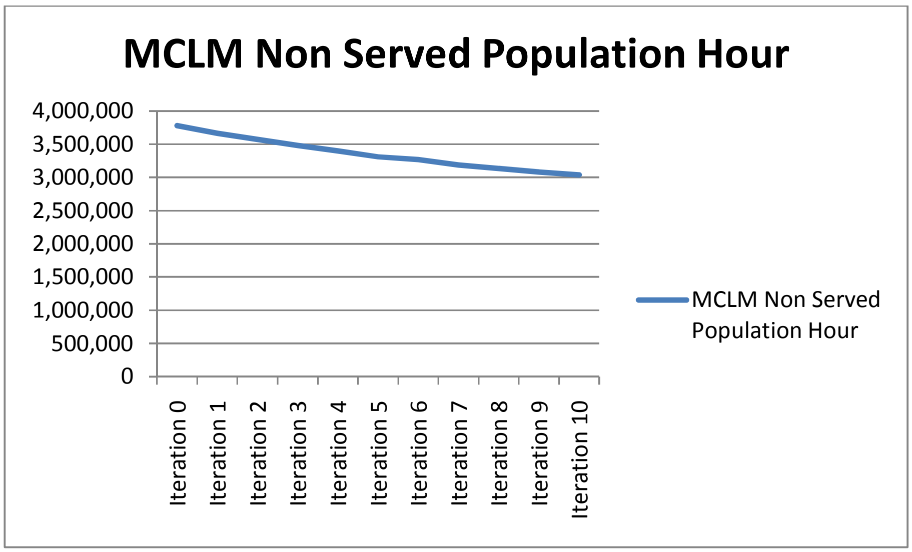

Figure 8 – Non served population hour for MCLM during iterations

Even though, the initial setting of SPBDCM for Istanbul data is better than MCLM, the overall improvement of MCLM transcends SPBDCM. Like MCLM, SPBDCM enforces the maximum number of ambulances that can be allocated to a supply node as 1, however there is a back up ambulance for each covered node in initial setting of SPBDCM. The property of maintaining a back up ambulance is neglected throughout the iterations. The relocation of an ambulance and determination of the least busy ambulance conditions are same for all models. In table 26 the performance metric of SPBDCM for each iteration of heuristic is given.

	SPBDCM			
	Non Served Population Hour	NSP	Lower bound	Improvement
Iteration 0	3,501,017	29.67%	2.24%	N/A
Iteration 1	3,434,087	29.10%	6.76%	0.57%
Iteration 2	3,378,967	28.63%	9.30%	0.47%
Iteration 3	3,314,679	28.09%	11.09%	0.54%
Iteration 4	3,265,789	27.67%	10.68%	0.41%
Iteration 5	3,199,743	27.11%	11.14%	0.56%
Iteration 6	3,122,841	26.46%	12.62%	0.65%
Iteration 7	3,089,076	26.18%	13.36%	0.29%
Iteration 8	3,064,367	25.97%	14.46%	0.21%
Iteration 9	3,032,061	25.69%	14.98%	0.27%
Iteration 10	3,005,432	25.47%	11.09%	0.23%

Table 18 - Changes in SPBDCM performance during iterations

Like MCLM, the system converges to 3,000,000 non served population hour but the final location of ambulances for each model differs. However, the coverage conditions of demand points for iteration 3 and 10 are same. Even though the same demand nodes are covered, the performance of the system is distinctly different. The reason behind this is due to the locations of the ambulances. NSP is derived from simulation and not related with the lower bound of the system. For both iteration 3 and iteration 10, the number of times that a node is covered for some demand nodes differs from each other. Some ambulances are assigned to same regions in order to lower NSP score of the model. The change in non served population hour for SPBDCM is given in figure 9.

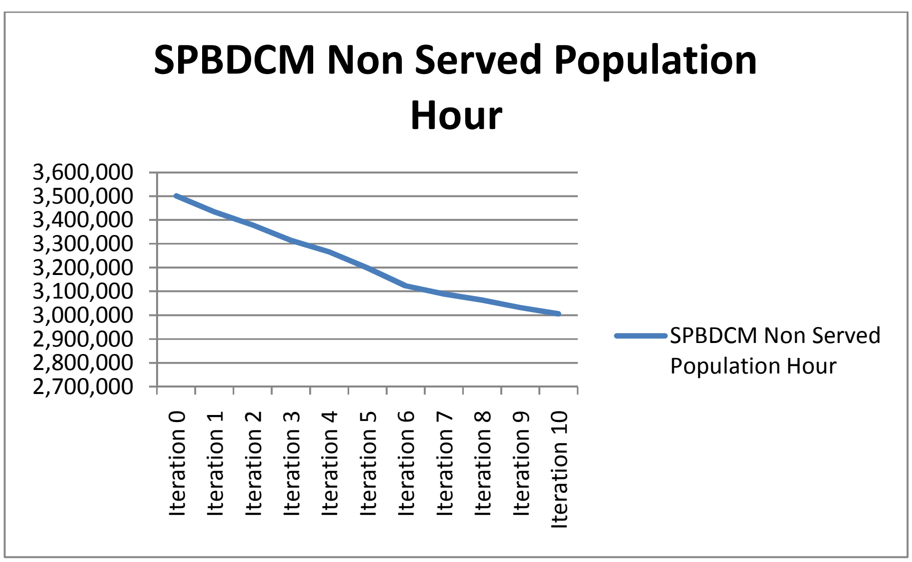

Figure 9 - Non served Population Hour for SPBDCM during 10 iterations

The initial setting of MSRSCLM allows multiple ambulances that can be allocated to same supply node. For the first iteration, there are 2 ambulances that have the same busy probability which makes both of these ambulances eligible for relocation. Both of these ambulances are located at the same node hence reassignment of one of these ambulances lowers the lower bound value of the system. The average improvement in each iteration is relatively small when it is compared to both MCLM and SPBDCM. However, the initial setting and the final setting of MSRSCLM after 10 iterations still perform better than both of the models. In Table 27, the performance of MSRSCLM is shown.

	MSRSCLM			
	Non Served Population Hour	NSP	Lower bound	Improvement
Iteration 0	2,580,020	21.86%	12.83%	N/A
Iteration 1	2,506,752	21.24%	11.67%	0.62%
Iteration 2	2,471,087	20.94%	11.35%	0.30%
Iteration 3	2,434,789	20.63%	10.65%	0.31%
Iteration 4	2,399,752	20.33%	11.32%	0.30%
Iteration 5	2,328,932	19.73%	11.15%	0.60%
Iteration 6	2,286,879	19.38%	10.33%	0.36%
Iteration 7	2,268,693	19.22%	12.14%	0.15%
Iteration 8	2,255,689	19.11%	11.04%	0.11%
Iteration 9	2,240,086	18.98%	12.30%	0.13%
Iteration 10	2,224,583	18.85%	10.55%	0.13%

Table 19 - Performance of MSRSCLM during iterations

For CFRM on the other hand, the improvements are shown only for 4 iterations. Since if the next iteration cannot improve the performance of the model is one of the stopping criteria, the algorithm is terminated. Since the heuristic works as a myopic algorithm, the obtained result is only a local optimum. Furthermore, the lower bound obtained in first iteration is interesting because the least busy ambulance covers some of the demand nodes solely by itself. The relocation of this ambulance lowers the NSP score but the reassigned location's aggregated population is lower than expected. Thus the immediate question arises from this condition is whether the model's constraint is loose or the total number of ambulances for the setting is not sufficient. Since the relocation of ambulance is determined based on 99 ambulances and an additional ambulance for assignment, the decision could vary for different number of ambulances. On the other hand, the maximum number of ambulances that can be assigned to supply nodes constraint of the model is probably not as efficient as it seems. In Table 28, the performance of CFRM for heuristic is given.

	CFRM			
	Non Served Population Hour	NSP	Lower bound	Improvement
Iteration 0	1,747,931	14.81%	13.41%	N/A
Iteration 1	1,720,487	14.58%	13.76%	0.23%
Iteration 2	1,704,918	14.45%	13.26%	0.13%
Iteration 3	1,689,083	14.31%	12.16%	0.13%
Iteration 4	1,678,438	14.22%	11.24%	0.09%
Iteration 5	No improvement	N/A	N/A	N/A

Table 28 - Performance of CFRM for heuristic

Although the improvements for each model is possible, with a meta heuristic the optimal NSP score of the problem data set can be reached. Without allowing assignments that can increase the NSP score, the heuristics might be trapped in local optimum. However, since the heuristic only allows one ambulance to be relocated at each iteration the initial setting of the model limits the maximum possible improvement. Hence the NSP scores converges for both MCLM, and SPBDCM to approximately 25.5%.

.

CHAPTER 7
CONCLUSION AND FUTURE RESEARCH

In this study, we presented two new models for deterministic location problems, CFRM and MSRSCLM. After the NSP scores of MCLM and SPBDCM for Istanbul data are evaluated, we developed CFRM in order to get better simulation scores. Even though there are numerous deterministic set covering location models in the literature, we stress the necessity of proposing new models for simulation purposes. Although, for experimental study CFRM performance was worse than MCLM for most of the cases, the robustness of CFRM is relatively better. By using simulation, we estimate the performance of deterministic models on different problem data sets. MCLM, SPBDCM, CFRM and MSRSCLM are evaluated with two dispatching policies. We emphasized the importance of an analytical tool for evaluating the performance of a deterministic location model. Even though the assumptions, the construction of constraints and objective function values are important, without a realization tool, the general behavior of the model cannot be estimated. The study shows that the model's performance can vary according to problem data set. Even for the same layout, same problem size, same distribution of demand points, the models can perform differently.

Further research on this topic may include developing new models for simulation purpose, testing the models for different demand points' coordinates distribution, and testing the performance of the models for different dispatching rules. Furthermore, the proposed myopic heuristic can be exchanged with a metaheuristic to evaluate the possible improvement that can be made for the models' initial settings. Assigning penalty costs for missed calls may be another extension for studied policy 1.

Consequently, the importance of a realization tool like queuing models or simulation is significant for EMS location planning. By omitting this aspect of the research, the "real" service provided by the EMS stations cannot be evaluated properly.

BIBLIOGRAPHY

Başar, A., 2008. Planning Emergency Medical Service Stations and Ambulances. MSc. Master Thesis, Sabancı University, Istanbul, Turkey.

Batta, R., Dolan, J.M., Krishnamurty, N.N., 1989. The Maximal Expected Covering Location Problem: Revisited. Transportation Science 23, 277–287.

Berman, O., Krass, D., 2002. The Generalized Maximal Covering Location Problem. Computers & Operations Research 29(6): 563-581.

Brandeau, M., Larson, R.C., 1986. Extending and Applying the Hypercube Queuing Model to Deploy Ambulances in Boston. TIMS Studies in the Management Sciences 2, 121–153.

Brotcorne L., Laporte G., Semet F., 2003. Ambulance Location and Relocation Models. European Journal of Operational Research 147, 451-63.

Burwell, T.H., Jarvis, J.P., McKnew, M.A., 1993. Modeling Co-Located Servers and Dispatch Ties in the Hypercube Model. Computers and Operations Research 20, 2, 113–119.

Church, R.L., ReVelle, C.S., 1974. The Maximal Covering Location Problem. Papers of the Regional Science Association 32, 101-118.

Çatay, B., Başar, A., Ünlüyurt, T., 2007. Istanbul'da Acil Yardım Istasyonları ve Araçlarının Planlanması. IBB Projem Istanbul Projesi Sonuç Raporu, Istanbul.

Daskin, M.S., Stern, E.H., 1981. A Hierarchical Objective Set Covering Model for Emergency Medical Service Vehicle Deployment. Transportation Science 15, 137-152.

Daskin, M.S., 1983. A Maximum Expected Location Model: Formulation, Properties and Heuristic Solution. Transportation Science 7, 48-70.

Erkut E., Ingolfsson A., Erdogan G., 2008. Ambulance Location for Maximum Survival. NAVAL RESEARCH LOGISTICS Volume: 55 Issue: 1, 42-58.

Galvão, R.D. , Morabito, R., 2008. Emergency Service Systems: The Use of the Hypercube Queueing Model in the Solution of Probabilistic Location Problems. International Transactions in Operational Research 15, 525-549.

Gendreau, M., Laporte, G., Semet, F., 1997. Solving an Ambulance Location Model by Tabu Search. Location Science 5, 75-88.

Goldberg, J.B., 2004. Operations Research Models for the Deployment of Emergency Services Vehicles. EMS Management Journal 1, 20-39.

Grossman, D.C.,Kim, A., Macdonald, S.C., Klein, P., Copass, M.K., Maier, R.V.,1997. Urban-rural Differences in Prehospital Care of Major Trauma. Journal of Trauma-Injury Infection & Critical Care: April 1997 - Volume 42 - Issue 4, 723-29.

Hogan, K., ReVelle, C.S., 1986. Concepts and Applications of Backup Coverage. Management Science 32, 1434-44.

Iannoni, A.P., Morabito, R., Saydam, C., 2009. An Optimization Approach for Ambulance Location and Districting of the Response Segments on Highways. European Journal of Operational Research 195, 528-42.

Ingolfsson A., Budge S., Erkut E., 2008. Optimal Ambulance Location with Random Delays and Travel Times. HEALTH CARE MANAGEMENT SCIENCE Volume: 11 Issue: 3, 262-274.

Jarvis, J.P., 1975. Approximating the Equilibrium Behavior of Multi-Server Loss Systems. Management Science 31, 235-39.

Larson, R.C., 1974. A Hypercube Queuing Model for Facility Location and Redistricting in Urban Emergency Services. Computers & Operational Research 1, 67-95.

Larson, R.C., Odoni, A.R., 1981. Urban Operations Research. Prentice Hall, New Jersey.

Morabito, R., Chiyoshi, F., Galvão, R.D., 2007. Non-Homogeneous Servers in Emergency Medical Systems: Practical Applications Using the Hypercube Queueing Model. Socio-Economic Planning Sciences 42, 255–70.

Repede, J.F., Bernardo, J.J., 1994. Developing and Validating a Decision Support System for Locating Emergency Medical Vehicles in Louisville, Kentucky. European Journal of Operational Research 75, 567–581.

ReVelle, C.S., Hogan, K., 1989. The Maximum Availability Location Problem. Transportation Science 23, 192-200.

Sacks, S.R., Grief, S., 1994. Orlando Police Department Uses OR/MS Methodology, New Software to Design Patrol districts. OR/MS Today, Baltimore, 30–32.

Savas, E.S. 1969. Simulation and Cost-effectiveness Analysis of New York's Emergency Ambulance Service. Management Science, 15(12)

Setzler H., Saydam C, Park S., 2009. EMS Call Volume Predictions: A Comparative Study. COMPUTERS & OPERATIONS RESEARCH Volume: 36 Issue: 6, 1843-1851.

Schilling, D.A., Elzinga, D.J., Cohon, J., Church, R.L., ReVelle, C.S., 1979. The TEAM/FLEET Models for Simultaneous Facility and Equipment Sitting. Transportation Science 13, 163-175.

Simpson N. C., Hancock P. G., 2009. Fifty Years of Operational Research and Emergency Response. JOURNAL OF THE OPERATIONAL RESEARCH SOCIETY Volume: 60, S126-S139.

Swersey, A.J., 1994. The Deployment of Police, Fire and Emergency Medical Units. OR & MS 6, 151-200.

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

Wu, C.H., Hwang K.P., 2009. Using a Discrete-event Simulation to Balance Ambulance Availability and Demand in Static Deployment Systems. ACADEMIC EMERGENCY MEDICINE Volume: 16 Issue: 12, 1359-1366

APPENDIX A

SENSITIVITY ANALYSIS RESULTS

Since the analysis includes too many data, the file that contains analysis is uploaded to internet. The analysis can be found at

http://people.sabanciuniv.edu/~tonguc/yasir/appendixA

APPENDIX B

EXPERIMENTAL STUDY DATA

Since the analysis conducted on 240 instances, the data is too large for appendix hence is uploaded to internet. The data can be found at

http://people.sabanciuniv.edu/~tonguc/yasir/appendixB