AN APPLICATION OF INTEGRATED STAFFING AND SHIFT OPTIMIZATION FOR A CALL CENTER

by

Cihangir Ertaban

B.S., Management Engineering, İstanbul Technical University, 2007

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Industrial Engineering

Boğaziçi University

2011

ACKNOWLEDGEMENTS

After the efforts I spent for my graduate program and this thesis, and saw my dream come true, I first and foremost have to thank my family and especially my mother. She is the reason I survive. I owe her everything I have done, and everything I will do in my life. I dedicate this thesis to her with all my heart.

It gives me great honour in acknowledging the aid of my Professor Refik Güllü. Everytime I needed, he generously shared his infinite wisdom with me. For all his support, understanding approach and contribution, I am indebted to him. It definetely was not possible for me to complete this thesis without his help.

Lastly, I am greatful to my former managers Devrim Rodop, Mark Schweitzer and Ertekin Güzel for the support and encouragement they have given throughout my graduate education.

ABSTRACT

AN APPLICATION OF INTEGRATED STAFFING AND SHIFT OPTIMIZATION FOR A CALL CENTER

Inbound call center problems need decisions in four main areas; the number of incoming calls, necessary number of agents to handle these calls, shift starting times for the operations and the number of agents to be allocated to each shift. Call center models mainly cover these areas seperately. This study approaches to the problem with an integrated Markovian approximation, heuristic supported simulation and optimization model. Incoming demand data is adapted from a private Turkish Bank and an application of the model is conducted. Moreover, probabilistic lunch breaks are introduced to the optimization part of the model. Finally, the model can be used as an easy decision support tool for the call center managers.

ÖZET

BÜTÜNLEŞİK BİR ÇALIŞAN PLANLAMA VE VARDİYA ENİYİLEMESİ MODELİNİN BİR ÇAĞRI MERKEZİ UYGULAMASI

Çağrı merkezi problemlerinde dört ana alanda verilecek kararlar belirlenmelidir; gelen çağrı sayısı, bu çağrılara yanıt verecek çalışan sayısı, çalışanların mesai başlangıç saatleri ve bu mesailere kaç çalışanın atanacağı. Çağrı merkezi modelleri bu alanları içerir. Bu çalışmada probleme bütünleşik bir Markov yaklaşımı, sezgisel bir model ile desteklenmiş bir benzetim ve enyileme ile yaklaşılmıştır. Gelen çağrı sayısı için ise, özel bir Türk bankasının çağrı merkezi baz alınmış ve model uygulanmıştır. Buna ek olarak, olasılıksal öğle yemeği molaları modele eklenmiştir. Son olarak, bu model, çağrı merkezi yöneticileri için kolay kullanılabilir bir karar destek aracıdır.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS / ABBREVIATIONS

ACD	Automatic Call Distributor
CRM	Customer Relationship Masnagement
ED	Efficiency Driven
ICT	Information and Communication Technologies
IID	Independently and Identically Distributed
SIPP	Stationary, Independent, Period by Period
QD	Quality Driven
QED	Quality and Efficiency Driven

1. INTRODUCTION AND STUDY OBJECTIVES

Call center, also referred to as contact center, is an operational office in which employees use telephones to make large volumes of transactions (Wikipedia, 2011). The employees, who make transcations through calls are called *Agents*. Agents may reply to the customers who call the call center, or make new calls. The process in which customers call the contact center is named as *inbound* and the process in which agents make calls to the customers is named as *outbound*. There are inbound, outbound and a mix of inbound and outbound call centers in the industry. In this study, we will be focusing on a purely inbound call center.

1.1. The Need for Call Centers

In the modern business world, customer satisfaction is the main goal for the companies. In order to enhance customer satisfaction, customer contact is a key aspect. It is stated that customers, who perceive that they receive help and they are treated respectfully, are more likely to be satisfied with the service provider (Adams, 1963). Improved customer service through the call center will have a positive effect on the satisfaction of the customers who need to reach the company (Dickson et al., 2010). Customer satisfaction leads to growth and profitability for the companies, besides, Lywood et al (2008) argue that customer experience has a direct relationship with profitability. Customer satisfaction and customer contact being so important and effective for the companies, it is obvious to see that call centers, which is an easy, fast and clear solution for the companies, is vitally important. We will now provide a brief history and worldwide information of call centers.

Call centers have started in 1960's (Cagrimerkezidernegi.org, 2008) as a platform for complaints and requests of the customers. Through the past 50 years, with the changing business needs, customers requirements and developing communication and technologic tools, they have transformed to a knowledge center composing of edge databases, intelligent CRM applications including fast and approprite offerings, skill based call routing and web interactions (Riggs and Thyfault, 1999). These changes both lowered the cost per transaction of the call centers and improved service level with better and faster solutions.

Together with the advancements in Information and Communication Technologies (ICT), the direct effect on customer satisfaction and profitability explain the huge growth of call centers in the last two decades (Callaghan and Thompson, 2002), (Deery et al., 2002), (Avramidis et al., 2010), (Borst et al., 2004) (Gans et al., 2003). Being a low-cost, high effective service delivery and even sales opportunity, call centers have become the main interface between the companies and their customers (Lywood et al., 2008). Any company which make sales to the last customer, provide call centers as the way of delivering after sales service. Moreover, call centers are used for business-to-business sales. Eventually, it is easy to see why and how call centers have evolved to a large and international sector on its own.

1.2. Call Centers in the World and in Turkey

The industry is estimated to be over $340 billion in the world including 150,000 call centers, employing 10 million staff all over the world (Hurriyet.com.tr/ekonomi, 2011). 12% of call centers was estimated to be outsource in 2008 and 14% of call center services was provided offshore (Cagrimerkezleridernegi.org, , 2008). United States is the largest market and employer for call centers in the world, whereas India, Philippine and Egypt are successful offshore outsource examples that grow rapidly by providing government initiated lower cost opportunities which is an important advantage against developed economies.

Although not remarkable as India or Philippines, Turkey also has a growing call center industry. In 2011, number of employees in call centers are estimated to be around 40,000 in Turkey. Investments have reached 200 million TL in the last 5 years. Sharp growth is even more boosted with the governmental incentive legalized in April 14, 2011 which include tax reductions, investment location allocation, insurance and finance support if the investment amount is over 1 million TL and investment is in Region 4 (Region 4 consist of Trabzon, Ordu, Giresun, Rize, Artvin, Gümüşhane, Malatya, Elazığ, Bingöl, Tunceli, Erzurum, Erzincan, Bayburt, Ağrı, Kars, Iğdır, Ardahan, Van, Muş, Bitlis, Hakkari, Şanlıurfa, Diyarbakır, Mardin, Batman, Şırnak, Siirt). Thus, call center industry is defined and legalized by government with the industry and product code: 6420.0.99. Legalization of the industry is a very important milestone for the sector whose expectation is 85,000 employees in 2015 (Hurriyet.com.tr/ekonomi, 2011b).

1.3. Improving Efficiency

From the business perspective, since call center operations have reached large amounts of budgets and gained vital importance for customer satisfaction, outsourcing has become a very reasonable approach. The industry needs special equipments, well trained employees, complicated planning and demand management. Moreover, although call centers are fast and efficient ways for customer satisfaction, they also might lead to negative results and responses in customer attitude for the company, in case service level is lower than expectations. Thus, service level versus costs trade-off for the companies require careful analysis before decision making. Outsourcing opportunities help the companies keep their focus on their core business and let call center operations be managed by experienced and specialized companies.

Increase in budgets, outsourcing and tough competition brings the necessity of cutting costs. The growth in size and scope of the call centers create profitability opportunities by improvements in efficiency. Starting point of cutting costs and improving service level is understanding the costs. Hishinuma (2007) stated that 60 to 70% of the costs of call centers are labor related. Clearly, in order to improve efficiency of call centers, it is a must to determine the staffing levels carefully. The more agents in call centers mean the quicker a service representative contacts the customer; but unfortunately the companies would not be able to afford any agent inefficiently just waiting for the customers to call. Therefore, it is important to analyze how to plan a call center in the best way possible. Naturally, the decision has to be given based on performance criteria. A well known metric for measuring the performance of the call center is the fraction of callers who abandon to total incoming calls, which is an obvious and key performance indicator, that we also will use (Hishinuma et al., 2008).

This thesis is dedicated to analyzing the service level and improving agent staffing in an inbound call center. Analysis of the service level includes determination of necessary number of agents for the planning horizon for a given set of arriving calls. The sample incoming calls' data is provided by a private bank's call center in Turkey. Based on the sample data, we will use queueing and analytical models together with simulation and a heuristic to determine the necessary number of agents. For each time period in the planning horizon, the minimum number agents that can process target percentage of incoming calls will be clarified.

Improving agent staffing is the employee allocation part of this study. Necessary number of agents should be working at the relative periods but since employees have boundaries for working hours, meal times, coming and departing times, their shifts are to be determined in an optimal way. Determination of the shift starting times and the number of employees working on these shifts is a key part of call center model.

There has been a lot of publications about call centers. This thesis uses some of these methods and combinations of them. Firstly, we perform an application of a private Turkish banks call center data to a queueing model which determines agent requirements. Then we use the result of the analtyical queueing model as input for a simulation combined with a heuristic algorithm. Eventually, we determine the shift starting times and the number of employees to work at each shift. At this part, we have a contribution to the shift optimization algorithm. We provide probabilistic meal breaks for the agents so that loss of workforce could be mitigated to a degree. The solution of the optimization model is the answer to the decisions of the main problems, the number of employees and their shifts' starting times.

We will introduce remarkable literature in the following section. In part 3, our model's framework will be explained in detail including determination of number of agents, simulation, adjustment heuristic and shift optimization model. Next part is the

Computational Analysis in which we show our computations and the final optimum shifts for the data of the call center. In Scenario Analysis and Results, we will show the different optimization scenarios and their affects to the final shifts. We will conclude in the latter section.

2. LITERATURE SURVEY

Call center employee staffing problem has attracted researchers since the industry started growing. The problem includes aspects from queueing, shift determination, employee allocation and employee scheduling models. Studies covering these models are briefly introduced in this section after general description of the call center process.

2.1. Call Center Process

Call centers have two different characteristics; inbound and outbound. Inbound call centers are the ones in which there are incoming calls and responding agents. In these call centers, the arriving call enters the queue if there are no available agents, waits for its turn and then faces the agent. The customer may abandon the call if the transaction is not completed before the customers patience limit. The time, which passes until the agent response is called patience time.

The other type; outbound call centers are designed for calling the customer, for sales or Customer Relationship Management (CRM) purposes. Agents call the customers from a prepared list. Some call centers are also managed to work both ways so that the agents make outbound calls when there is no necessity for more agents to handle incoming calls (Gans et al., 2003). We will be disregarding outbound operations in this study.

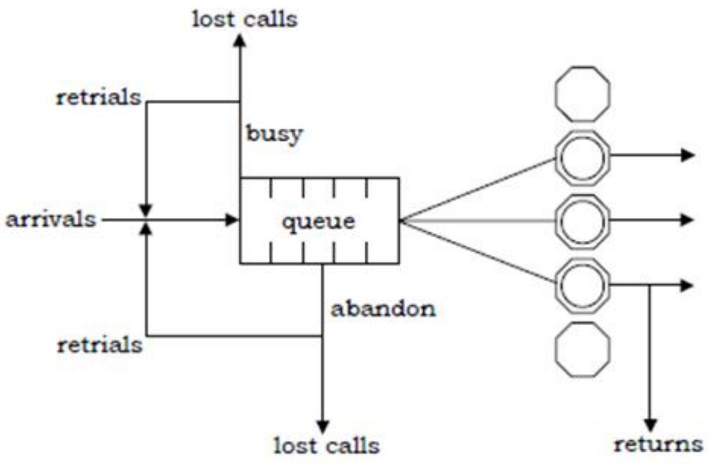

Figure 2.1. Scheme of an inbound call center (Koole and Mandelbaum, 2002).

In Figure 2.1, a simple inbound call center is demonstrated. There are a limited number of work stations and a limited number of serving agents. A customer calling the line is a new arrival and new arrivals enter the queue if there are no available agents. During any time in the queue, the customer can stop waiting, hang up the phone and abandon the call which makes it a lost call. If there are customers in the queue, the agent serves the longest waiting customer in the queue whenever the previous customer departs from the system. Thus the system is a First In First Out system.

In practical life, of course call center operations is not as simple as illustrated in Figure 2.1. The customers need one or a couple of a numerous complex possible transactions. Since agents are needed to be trained well for these operations in order not to cause dissatisfaction, it is not feasible to train all the agents for all the skills. Instead, the capabilities and trainings necessary for transactions are grouped and so are the agents. The agents are trained well but only for some subset of the tasks. Which eventually bring the necessity to distribute the calls to the appropriately trained agents. This process is generally handled by Automatic Call Distributors (ACD). The determination of the number of skills to the agents and determination of the number of agents are studied in multiskill call center skill based routing problems. Pisacane (2009) focuses on the agent scheduling in a multiskill call center. Our model is a single route model in this study.

2.2. Call Center Staffing Model

Management of call centers require answers to the following questions: “How many agents should work at each time of day?”, “What time should their shifts start?”, “How many shifts are there going to be?”, “How many employees should work at each shift?” and lastly “How could the employees be allocated to the shifts?”. All these questions need thorough investigation and analysis. Call center models and studies are supposed to cover and analyze these decisions. The milestone Kolesar et al. (1975) is the model referred as SIPP (Stationary, Independent, Period by Period) and is similar to Buffa et al. (1976)’s stepwise approach. This approach for workforce allocation is widely accepted and followed in literature (Mason et al., 1998) (Ingolfsson et al., 2002).

Buffa’s model and SIPP consist of four steps:

i. Forecast period-by-period demand rates for the planning horizon. The period is commonly 15 minutes, 30 minutes, or one hour, but could be any other time unit. The planning horizon is one day to one month in general.

ii. Estimate minimum necessary number of agents for each period. For each period, a determined number of agents are required to handle calls. The number of working agents should be estimated by the goal of service level.

iii. Determine the set of permissible shifts. Shifts are constructed based on the pre-defined rules such as allowable working hours, part-time/ full-time options, breaks and even over-time. According to the rules and needs, the starting times of shifts are to be determined.

iv. Select a set of shifts that minimizes labor cost (labor cost is number of agents unless cost per agent is differentiated), while providing at least the minimum necessary number of employees in each time period.

We will investigate these steps in detail now.

2.2.1. Demand Forecast

The first step; forecasting is generally left out of scope in the papers since the input load levels are generated through databases' ACD reports (Gans et al., 2003). These reports provide incoming call data for even years, so that any average or statistical forecasting model can be used as input for the second step of the approach.

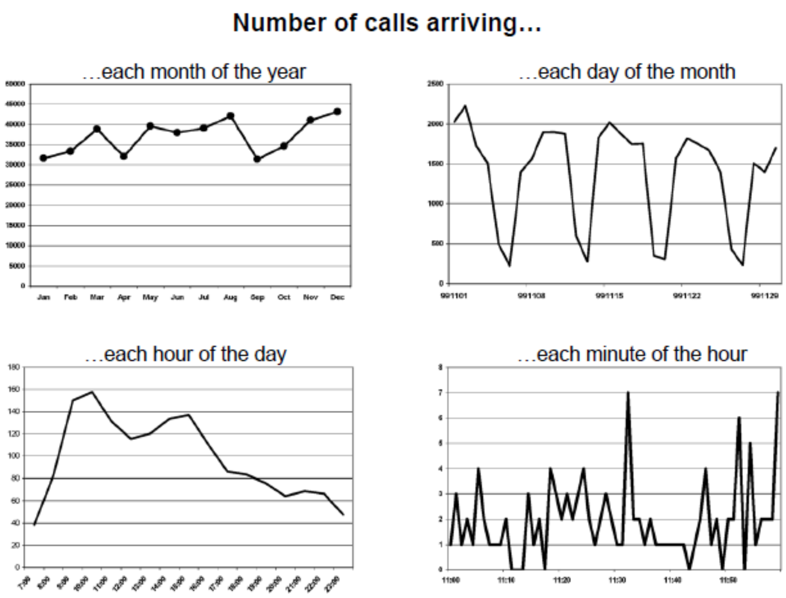

Figure 2.2. Fluctuation of demand (Gans et al., 2003).

Unfortunately, demand fluctuation is quite high in call centers because of seasonal effects, economic changes, sales opportunities, advertisements and promotions. As a result, this uncertainty has a huge effect in call center management despite being disregarded by the models. Figure 2.2 illustrates fluctuation in demand.

2.2.2. Determination of Necessary Number of Agents

The studies related to call centers may be assumed to have started when Agner Krarup Erlang -defined as "a mathematician, statician and an engineer who invented the fields of traffic engineering and queueing theory" (Wikipedia, 2010)- published his famous Erlang loss formula in 1917 (Takacs, 1969).

Erlang's formula and any analytical method estimates the resulting service level for a given number of agents. Here, service level estimations need to be clarified; there are a couple of estimations for service level measurements most of which are the fraction of the number of answered calls to the number of total incoming calls and the fraction of the number of calls handled within a determined number of seconds to the number of total incoming calls (Avramidis et al., 2010). Nevertheless, these and all smilar methods are parallel to each other since the core aim is to make contact with the maximum number of customers in the most efficient way. In this thesis, we define service level as the first one mentioned above; the fraction of the number of answered calls to the number of total incoming calls.

The estimations generate the resulting service level, but of course the service level may not be satisfactory enough. So, these analytical methods should be used in a heuristical manner by increasing or decreasing the number of agents given to the model until reaching the minimum number of agents for the target service level. After reaching these work requirements, some researchers used simulations to test the service levels generated by the analytical queueing models (Green et al., 2003).

The analytical models' assumptions are well defined by Kolesar et al. (1975) with SIPP. The periods are assumed to be stationary and independent from each other. In fact, this approach also has handicaps since in business life, periods may neither be stationary nor independent. Studies also have been made to cover these handicaps (Whitt, 1999), (Green et al., 2001).

2.2.2.1. Erlang-B

The Erlang's loss formula describes a telephone exchange in which calls arrive with a Poisson process of density λ. There are m homogenous lines (resources) in the system. If the entering call finds a free line, then the transaction starts and is completed after service time which is independent and identically distributed (IID). Otherwise the call is lost. The holding times for each call are mutually independent (Takacs, 1969). The Erlang's formula demonstrates the probability P_b (Blocking Probability) that all lines will be busy at time t and the incoming call will be lost is;

$$P_b = B(E, m) = \frac{\frac{E^m}{m!}}{\sum_{i=0}^{m} \frac{E^i}{i!}} \tag{2.1}$$

$E = \lambda h$ is the total amount of traffic offered in Erlang unit. (2.2)

This model is referred to as *Erlang-B* (Angus, 2001). Erlang-B is remarkable for network traffic but although being fundamental for explanation, it is not the accurate model for call centers. Call center queueing distributions are better explained by the Erlang-A model (Mandelbaum and Zeltyn, 2005). We will be explaining Erlang-A later.

2.2.2.2. Erlang-C

The second Erlang model is Erlang-C. The main difference between Erlang-C and Erlang-B is waiting spaces for the customer. In Erlang-B the arriving call is dismissed if there are no available servers at the time of arrival. Meanwhile, in Erlang-C, the calls wait for some IID time of exponential distribution, to be served. The deficiency of Erlang-C is the calls may not abandon the system. Although being frequently used for call centers (Mandelbaum and Zeltyn, 2005), it is not the able to project a real-life call center.

The model is an *M/M/N+M* queue which has a Poisson arrival process with rate λ, IID service times with an exponential distribution having mean $1/\mu$, N homogenous parallel servers, infinite waiting room and the First-in-First-out service discipline (Whitt, 2005).

$$P_W = \frac{\frac{A^N}{N!}\frac{N}{N-A}}{\sum_{i=0}^{N-1}\frac{A^i}{i!} + \frac{A^N}{N!}\frac{N}{N-A}} \tag{2.3}$$

$A = \lambda h$ is the total amount of traffic offered in Erlang unit (Same as 2.2).

P_W is the probability that an incoming customer has to wait for service.

Since holding times are also taken into account, Erlang-C is a more complex model than Erlang-B but Erlang-C might lead to instability in larger traffic load.

2.2.2.3. Erlang-A

Erlang-A model is also an *M/M/s+M* model but is the abandonment version of Erlang-C (Mandelbaum and Feldman, 2008). Taking abandonment after some patience time into account, Erlang-A model would result in lower necessity of agents than the Erlang-C model. Moreover, this enhancement brings stability to the system.

Halfin and Whitt (1981) call this approach Quality and Efficiency Driven (QED) regime for the *M/M/N* queue. Zeltyn and Mandelbaum (2005) include also Quality Driven and Efficiency Driven operational regimes whose difference is priority being given to efficieny or service level.

Erlang-A model is very commonly used because it is a good explanator by taking abandonment and patience time into account and it is better than Erlang-C for call center estimations, though, it is very complicated and some researchers choose to make approximations to this model or use Erlang-C instead (Whitt, 2005). We will be explaining similar Markovian models in the following part.

2.2.2.4. Queueing Models

There are a lot of published studies related to queueing models. More and more researchers work on common distributions, which is a necessity that cannot be satisfied by Erlang models.

Bacceli and Hebuterne (1981) focus on a *GI/GI/1+GI* model concerning impatient customers who leave after waiting for some time when all servers are busy. Mandelbaum and Momcilovic (2009) generalize QED regime with *G/GI/N+GI* queue, Brandt and Brandt (2002) work on asymptotic results of the *M/M/s+GI* system and publish a Markovian approximation. Zeltyn and Mandelbaum (2005) point out that the classic Erlang-C *M/M/n* queueing model is insufficient in explaining abandonment and defend *M/M/n+G* by introducing patience time for the waiting calls having a common distribution *G*. Mandelbaum and Feldman (2008) research *M/M/s* (time varying Erlang-C) with non exponential service times and an asymptotic approach.

Borst et al. (2004) use an *M/M/N* queue with *N*, the number of agents, being large and develop an asymptotic framework. Besides, Jennings et al. (1994) suggest an approximate procedure and compare with the Markovian *M/M/s* model.

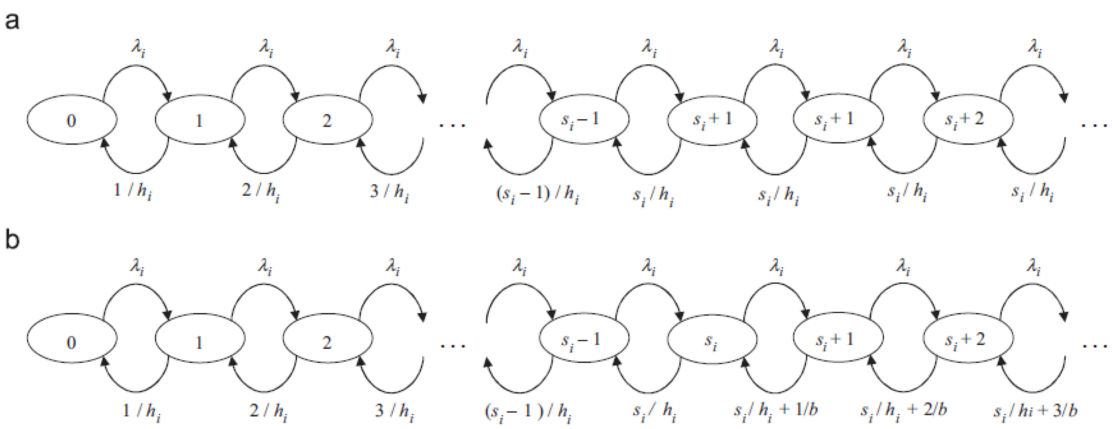

Figure 2.3. Transition diagrams for queueing models. (a) basic model and (b) improved model including abandonment (Diets, 2011).

As a result, queueing models like Erlang-A have been a popular topic for the researchers. Modifications of the model to cover common distributions and asymptotic approaches were published.

2.2.2.5. Square-Root Staffing Rule

Square-Root Staffing is a robust and general rule generated from an *M/M/s+M* Erlang-A model, providing fast and easy solution for the estimation of necessary servers when offered load (rate times time unit in Erlang unit) is given. It is also derived from the Markovian model and it is known as a rule of thumb.

The rule recommends *N** servers with the formula below;

$$N^* = R + y^*(\frac{a}{c})\sqrt{R} \tag{2.4}$$

where c is the staffing cost per agent per hour and a is the waiting cost per customer per hour. R is then the offered load estimated the same way as 2.2. $y^*(r), r \geq 0$ is an easily computable function described in Borst et al. (2004). Moreover, Square-Root Safety Staffing rule is explained by Jennings et al. (1996), Kolesar and Green (1998), and in the Cal Center Tutorial of Gans et al. (2002).

2.2.2.6. A Markovian Approximation

Whitt (2005) proposes an algorithm for the basic call center queueing model *M/GI/s/r+GI* in order to compute rapid approximations for the standard performance measures. The published model assumes a Poisson arrival process, IID service and customer abandonment times with general distributions, *s* servers and *s+r* waiting slots. The paper explains that empirical studies of call centers show non exponential distributions, that is why he uses general distributions. This decision make the model very difficult to analyze directly. Instead, a Markovian *M/M/s/r+ M(n)* queueing model where *M(n)* is the state-dependent abandonment rate is proposed. The model characterizes the steady-state waiting time distributions via their Laplace transforms and numerically invert the transforms to compute approximate distributions. Finally the results are compared and it is shown that the approximation performs well.

We will adapt Whitt (2005)'s model in our study to estimate the necessary number of agents and use a heuristically iterated simulation afterwards.

2.2.3. Shift Determination and Employee Allocation

The task of determining shift starting times in the planning horizon and the number of employees for each shift is referred to as *workforce allocation* problems in the literature (Mason et al., 1998). Since determination of shifts, determination of the number of employees to work on these shifts and scheduling are combined in each other, we will survey the literature together in this part.

According to Ingolfsson et al. (2002), Dantzig was the first to formulate this problem as a mathematical program. Dantzig (1954) was a comment on Edie's "Traffic Delays at Toll Booths" research and was published three months afterwards, in August 1954. The article proposed a method to find an optimal schedule for the workers so that necessary amount of employees could be provided with minimum resource usage. The method is explained here in detail:

Employees will be working seven hours a day, excluding one half hour meal break after working for three and a half hours.

Let x_i be the model variable, the number of employees who start their shift at time i, where $i= 0, 1, 2, \ldots, n$ corresponding to the half-hour time periods.

Let $a_{ti}= 1$ *if* $t= i, \ldots, i+ 6, i+ 8, .. , i+14$ (where $i+7$ corresponds to the meal period) and $a_{ti}= 0$ otherwise. This parameter determines whether a worker, starting his shift at period i, will work at period t or not.

Then, optimization is a minimization problem so that

$$\text{Minimize } \sum_{i=0}^{n}(x_i) \quad \text{where} \quad (x_i \geq 0) \tag{2.5}$$

Under this goal, assignments are to be chosen which satisfy:

$$\sum_{i=0}^{n}(a_{ti}x_i) \geq \mathrm{b_t}, \qquad (t= 1, 2, \ldots, n+ 14) \qquad (2.6)$$

where b_t is the required number of servers for period t. The optimum assignment problem then can be stated as finding a set of non-negative x_i that satisfies 2.5 under the constraints 2.6.

It was stated (Dantzig, 1954) that the method was solvable by hand in a couple of hours. It is obvious that this integer programming problem can easily be solved with an optimization tool.

This model tries to minimize the total cost while covering the necessary staffing for each period. Because of this approach, this model is called “set covering” (Gans et al., 2003). There has been modifications and improvements on this core model. Brusco and Jacobs (1984) suggest a similar pattern tour-scheduling model which determined part time and full time shifts. Taylor and Huxley (1989) aim to minimize shortage of employees and surplus of employees by applying penalties. Easton and Rossin (1991) interestingly take employee satisfaction into consideration. Beaumont (1997) works on a mixed integer scheduling problem. The model tried to minimize the costs while satisfying the needs. Diets (2011) formulates a quadratic scheduling algorithm which repeats forming schedules until reaching the target service level.

Moreover, different approaches have also been used for workforce allocation. One example is Hishinume et al. (2007) trying a Tabu search heuristic. Helber and Henken (2008) publish an approximation of a dynamic stochastic system covering multiple scenarios for a multiskill contact center. Mason et al (1998) suggest an integrated simulation, heuristic and optimization method.

More information on shift determination and scheduling algorithms may be found in Gans et al. (2003) and more information on simulations can be found in Mandelbaum, (2003).

Taking all into consideration, employee scheduling is a very important problem that would affect profitability and organizational satisfaction of employees by preventing unnecessary employees to work and better management of off days. Since minimal staffing, target service level and satisfaction of employees' scheduling preferences are conflicting objectives, optimization methods are necessary (Easton and Rossin, 1991). Since determination of shifts on the 7/24 basis and allocation of employees change the staffing level critically, all required effort should be spent on this decision by call centers.

There are a lot of publications in literature spanning a wide area of aspects about call center. Still, some points are in need for deeper investigation. These issues include multi-skill routing models in which calls and agents are grouped according to specialties and they should be matched. Shift scheduling algorithms also can be broadened. Different shift options, part-time and full-time choices, costs of shifts, meal-breaks, coffe-times and over-time decisions are improvements for shift scheduling algorithm. In the future, it would be reasonable to expect studies in these areas, one of which is this thesis with a touch by integrating probabilistic lunch breaks to scheduling algorithm.

3. FRAMEWORK OF THE CALL CENTER MODEL

The aim of the study is to determine shift starting times and the number of agents to be planned for each shift for a call center whose demand is assumed to be known. For this reason, we constructed a model consisting of four levels; estimation of initial number of agents, simulation of the process, adjustment of the number of agents by a heuristic model and determination of shifts through an integer programming optimization model.

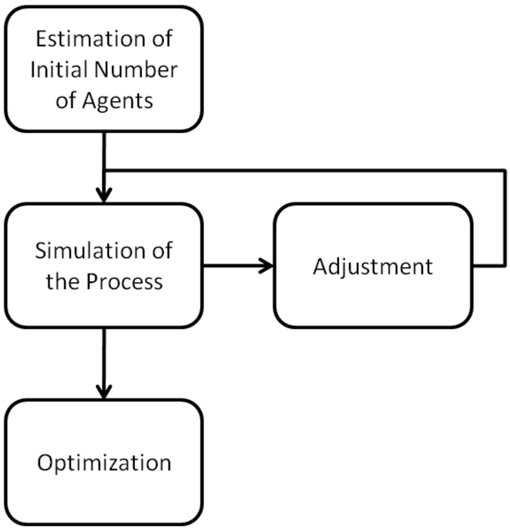

Figure 3.1. Framework of the model.

Details of the steps of the model will be explained in the following sections.

3.1. Determination of Initial Number of Agents

The model needs an input of required number of agents for each period that can serve more than a constant percentage of the customers. To reach this number, our model uses an analytical approximation and then simulations in a heuristical manner. The analytical approximation provide the initial number agents for the simulation. In this section we explain the model we used to reach an initial agent estimation to start the simulation.

For this estimation, we reviewed Erlang-B, Erlang-C and Erlang-A and decided to use Ward Whitt's Markovian approximation (Whitt, 2005) because it performs quite as good as an exact model and it covers general distributions.

Whitt (2005) explains a Markovian approximation having state-dependent abandonment rates, for the *M/GI/s/r+GI* queueing model. It is claimed that his model does not need more computational effort, performs equally well with Erlang-A model *M/M/s/r+M* (which covers exponential distributions) and it is easy to apply to the generalized *M/GI/s/r+GI* model.

The model is a Markovian *M/M/s/r+M(n)* queueing model where *s* is the number of servers and *r* is the number of waiting spaces. Then, let *N(t)* be the number of customers in the system at time *t*. In the stated Markovian model the stochastic process of $\{N(t): t \geq 0\}$ is a birth and death process whose arrival rate is λ and death rate is μ_k. Simply, death rate is service rate if there are no waiting customers and it is service rate plus total abandonment rate if there are waiting customers.

Let *k* be the number of customers in the system, death rate is estimated via 3.1.

$$\mu_k = \begin{cases} 1 \leq k \leq s & k * \mu \\ s+1 \leq k \leq s+r & s * \mu + \delta_{k-s} \end{cases} \quad (3.1)$$

Total abandonment rate of the system is shown in 3.2. Here, α is the random abandonment rate of a single customer waiting to be served.

$$\delta_{k-s} = \alpha * (k - s) \quad (3.2)$$

The steady state probabilities are estimated through local balance equations. A recursive method would be reasonable to estimate the probabilities.

$$x_k = \frac{x_{k+1} * \mu_{k+1}}{\lambda} \; \forall \text{ all } k \geq 0,\; x_k = 1 \text{ for } k = s \quad (3.3)$$

In order to normalize the probabilities, let the sum be,

$$y = \sum_{k=0}^{s+r} x_k \tag{3.4}$$

and normalized probabilities:

$$p_k = \frac{x_k}{y} \qquad 0 \le k \le s + r \tag{3.5}$$

After finding steady state probabilities of being in state k, we aim for the probability of being served or abandonment. Next, we estimate the probability that an admitted customer finds k customers in the system;

$$p_k^a = \frac{p_k}{1 - p_k} \qquad 0 \le k \le s + r \tag{3.6}$$

Thus, an admitted customer would wait in the system only if it finds more customers than the servers. If there are less customers, the new customer would not have to wait.

$$P(NoWait) = \sum_{k=0}^{s-1} p_k^a \qquad 0 \le k \le s + r \tag{3.7}$$

The case of arrivals, joining the queue makes the model a little more complicated.

$$\gamma_{k,j} = \frac{\alpha}{s * \mu + (k - j + 1) * \alpha} \tag{3.8}$$

$$m_{k,j} = \frac{1}{s * \mu + (k - j + 1) * \alpha} \tag{3.9}$$

The model defines *departure event* as an abandonment from the queue or start of service to a customer in the queue by completion of service of another customer being processed. As a result of both, the number of customers in the queue decreases by one. The first departure event for a customer is the departure event that happens first during the time the customer is in the system. Then 3.8 is the probability that the k^{th} customer will abandon at the j^{th} departure event which is an abandonment or start of service for the customer. 3.9 is the mean time between $(j\text{-}1)^{st}$ and j^{th} departure event. Whitt's original study could only make approximations for these values since they are only exact in the exponential case.

Finally we have reached our service level estimators; service level is accepted as the probability of service in this thesis. 3.10 is the probability that customer k eventually receives service.

$$\Gamma_k=(1-\gamma_{k,1})*(1-\gamma_{k,2})*..*(1-\gamma_{k,k}) \text{ for } k=0,1,2,\ldots,r \quad (3.10)$$

After all, the probability that a customer eventually completes its service is *P(S)*;

$$P(S)=\sum_{k=0}^{s-1}p_k^a+\sum_{k=0}^{r-1}p_{s+k}^a*\Gamma_k \quad (3.11)$$

$$P(A)=1-P(S) \quad (3.12)$$

This model will provide *P(S)* for a given input. But it should be run recursively while modifying the number of agents so that target service level will be reached at the best efficiency level.

Pseudecode for the algorithm is demonstrated in Figure 3.2.

```
Step 1:     Set initial values
Set initial servicetime, μ, α , r, targetservicelevel
Read λ and s from input file
Step 2:     Make computations
For each period
        Do until servicelevel > targetservicelevel
                For k= 1 to s+r
                        Compute δ_k-s  using Eq. 3.2
                        Compute μ_k using Eq. 3.1
                Next k
                For k= 1 to s+r
                        Compute x_k using Eq. 3.3
                Next k
                Compute y using Eq. 3.4
                For k= 1 to s+r
                        Compute p_k using Eq. 3.5
                Next k
                For k= 1 to s+r
                        Compute p_k^a using Eq. 3.6
                Next k
                Calculate P(NoWait) using Eq. 3.7
                For k= 1 to s+r
                        For j= 1 to k
                        Compute γ_k,j using Eq. 3.8
                        Compute m_k,j using Eq. 3.9
                        Next j
                Next k
                Compute Γ_k using Eq. 3.10
                Compute P(A) using Eq. 3.11
                Compute P(S) using Eq. 3.12
                If servicelevel < targetservicelevel
                        s = s +1
                End if
        Loop
        Record s as final s for this period
Next period
```

Figure 3.2. Pseudecode of the Markovian approximation.

An example period for the approximation is the first period of our computations (see section 4.2). An average of 64.42 calls are expected for this period, arrival rate is 2.15. First simulation starts with 1 agents and generate a service level of 7.8%. Next estimations provide inadequate service levels; 21.2% for 2, 35.5% for 3, 49.2% for 4, 61.9% for 5, 73.1% for 6 until finally 82.3% for 7 agents. We estimate all periods in this manner in our model framework.

3.2. Simulation of the Process

This part of the model aims to test the actual performance the number of agents in each peridod. In order to test the performance, we simulate an inbound call center using the tool Arena. The simulation generates calls, directs them into a queue, controls if there is a free agent. If there is a free agent, it processes the transaction, otherwise, the call waits in the queue for patience time and abandons the system.

The results of the simulation is written into a text file including the times of processed and abandoned calls. Thus, for each period, the simulated service level is estimated. Based on the service levels, the number of agents are changed according to the adjustment heuristic algorithm explained in section 3.3.

The simulation starts by generating calls based on the input in the schedule object. Declaration is hourly means of the exponential distribution of the number of incoming calls for each period. After the call is generated, it is taken into a queue. Here, the hold object controls whether there is an available agent or not. If there is a free agent, the call is released immediately and taken into process object which would delay it for processing time generated by an exponential distribution. Otherwise, the call is kept for some time generated by another exponential distribution. This time deflects patience time in the call center. When patience time is over, the call is again released as an abandonment. The abandoning call leaves the system.

Our expectation with the simulation is basically consistence with the estimation of Markovian approximation. Results of the simulation may be over or under previously estimated levels, each way, we will be able to reach optimum number of agent by the help of the heuristic approach.

Table 3.1. The objects used in Arena.

Object	Explanation
Create	Generates incoming calls
Assign	Assigns time information to the generated call
Hold	Controls the duty agents. If there is a free agent, releases the call, otherwise the call waits for some time generated by an exponential distirubiton
Decide	Decides whether the released call will be disposed or processed
Process	The call is processed for a time generated from an exponential distriubution
Dispose1	Exit of processed cals
Dispose2	Exit of abandoned calls
Readwrite1	Keeping the record of answered calls in a text file
Readwrite2	Keeping the record of abanadoned calls in a text file
Schedule1	Hourly means of each periods exponential distribution is declared
Schedule2	Number of working agents for each period is declared

The figure below shows the simulation model including the objects.

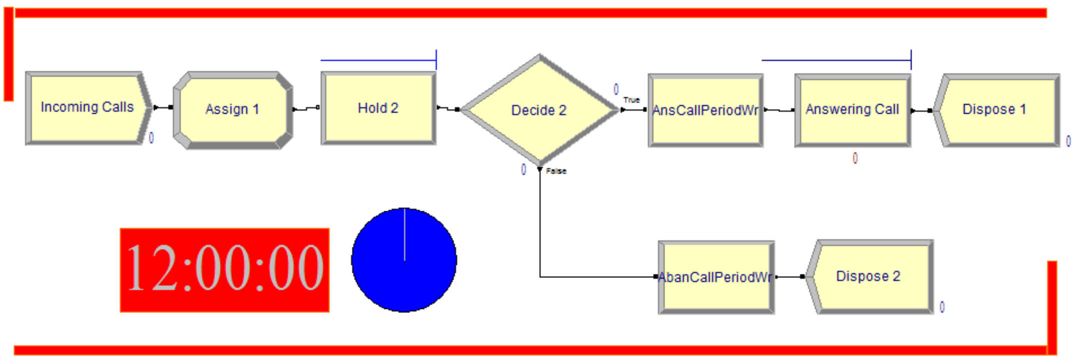

Figure 3.3. Arena simulation model.

Arena simulation provides testing opportunity to our analtyical Markovian approximation model. The simulation is run long enough in order to increase accuracy of the results.

3.3. Adjustment Heuristic

The results of the simulation may be over or under the determined constant level of service. In these cases, adjustments are necessary to find the ideal minimum number of agents for the shift determination algorithm.

A weight based heuristic is applied at this step. The heuristic finds the gap between the target service level and simulated service level. The number of agents are modified according to the gap. The larger the gap is, the larger the number of agents to be added or subtracted.

The advantage brought by this heuristic instead of just adding or subtracting agents is the speed gained through large jumps if agent numbers are high. Consider for the target service level of 80%. The heuristic subtracts 2 employees from the simulation result of 100% if the simulated number is 10, meanwhile it subtracts of 20 agents from the same simulation result if the simulated number is 100.

The heuristic stops for one period if the simulated number of agents that generates satisfactory service level, has occurred twice in the last three simulations. After the first time a number above target service level occurs, the model will decrease the number of agents related with the over-achievement range, thus a lower service level will be generated with a smaller number of agents. Then, if still there are too many agents, another decrease will occur, otherwise an increase will happen. To get closer and guarantee that we reach our focus, we add the constraint that the adjusted numbers will be smaller than the previous ones unless the adjusted number is one. If the adjusted number is one, then our algorithm is just fluctuating above an below the target service level. Hereby we understand that the number of agents above the service level is our target and we do not make any more adjustments.

Step 1: Set initial number of agents for each period

Step 2: Make simulation of 24 hours

Step 3: Estimate simulated periods' service levels

Step 4: Check whether target service levels are reached

For each period

If *service level> target service level* and *simulation count of the number of agents for that period >=2* then

End

Else go on with step 5

Step 5: Make adjustments

For each period

If *service level< target service level* then

Compute adj as: (*target service level-service level)/(target service level)×agent number*

Provide adj<previous adj unless adj=1

Increase agent number by nearest greater integer of adj

Record adj as previous for this period

Else

Compute adj as: (*target service level-service level)/(target service level)×agent number*

Provide adj<previous adj unless adj=1

Decrease agent number by nearest larger greater adj

Record adj as previous for this period

Loop

Go to step 2

Figure 3.4. Pseudecode for the heuristic algorithm.

As an example, consider our computation of period 30 (See section 4.3). Simulation begins with 52 agents to handle calls. Results generate a service level of 94% which is 14% over our target service level. Then we calculate 14% of 52, that equals to 7.28. We thn decrease the number of agents by nearest larger integer 8 -note that we round the calculation to the nearest larger integer since we would not like to tolerate a jump of magnitude zero, thus, we always change the number of agents. Next simulation starts with 44 agents and the result is 90%. We decrease the number of agents by 5. 39 agents result in a service level below our target: 75%. We then divide 39 by 75 and multiply it by 80, in order to estimate how many agents can provide 80% of service level if 39 agents are able to handle 75% of calls. The result is 41.6 leading us to our next simulation with 42 agents.

Then, we are again over target and decrease by 1. Afterwards, 41 agents do not satisfy our target. Lastly, we increase by 1 and simulate with 42 agents, since we had another simulation run with 42 agents before, this time we stabilize our estimation and go on with 42 agents in the following runs.

3.4. Optimization Model

The steps prior to the optimization of the shifts will provide minimum number of agents for each period. The shift determination algorithm uses these numbers and aims to find the minimum total number of agents that provide necessary level of service.

The method we use is a derivation of Dantzig's old and milestone optimization algorithm (2.3 and 2.4) with a few changes.

3.4.1. Probabilistic Lunch Breaks

Our call center is 7/24 on-line. The calls are desired to be handled all day. The agents work nine hours a day, including one hour lunch break. The first difficulty is to handle calls in 24 hours. Minimum three shifts would cover the working hours. The second issue is the lunch breaks since calls still are handled during the breaks. An improvement we introduce with our model is probabilistic lunch breaks in order to distribute the loss of workforce due to lunch. The agents start their shifts, work for three hours, then they may do their lunch break for four periods with a probability of 0.5. Thus, each agent will be missing for an hour, but the lunch periods will be distributed for the total workforce. After the lunch break, the agents work for another four hours.

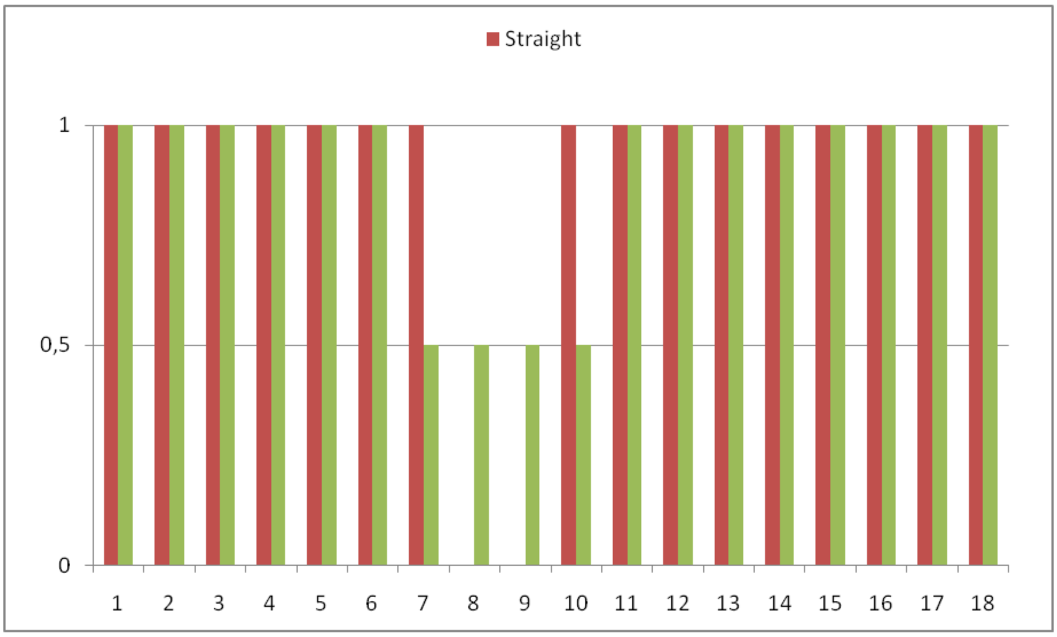

Figure 3.5. a_{ik} values in probabilistic and strict lunch breaks.

Vertical axis of Figure 5 is a_{ik} (explained in 3.4.2), and horizontal axis is the periods of one single shift. Probabilistic distribution of agents to the lunch break provide half of workforce during the meal times.

Table 3.2. Meal options for the agents

	Meal Period 1	Meal Period 2
Option 1	7	8
Option 2	8	9
Option 3	9	10
Option 4	7	10
Option 5	8	10
Option 6	7	9

Table above demonstrates the lunch period options of the employees. In order to keep workforce in adequate levels, the number of employees choosing one option should not be different than the others. From the agent perspective, probabilistic lunch breaks would be acceptable. The consecutive lunch options provide a full hour to the employees and the other options may offer an interesting choice to especially smoking ones. In fact, options 5 and 6 are probable but not necessary. Equal distribution of agents to the first four options would provide 50% of the employees working during all lunch periods.

3.4.2. Problem Formulation

Let x_i be the variable (integer) denoting the number of employees who start their shift at time i, and,

y_i be the variable which determines whether a shift starts at period i or not. So, y_i is a binary variable. $i = 0, 1, 2, \ldots, k-1$ corresponding to the half-hour time periods. (note that we have $k=48$ periods in our model)

Let a_{ik} be the parameter determining whether an agent, starting the shift at period i will be working at period k or not. a_{ik} is a table having values 1, 0.5 and 0 to carry the information of all possible shifts.

Then, optimization is a minimization problem so that

$$\text{Minimize } \sum_{i=1}^{k}(x_i) \qquad \text{where } (x_i \geq 0) \qquad (3.41)$$

Under this goal, assignments are to be chosen which satisfies:

$$\sum_{i=1}^{k}(a_{ik}x_i) \geq \mathrm{r_k}, \qquad (k = 1, 2, \ldots, 48) \qquad (3.42)$$

$$x_i \leq g * y_i \qquad (i = 1, 2, \ldots, 48) \qquad (3.43)$$

$$\sum_{i=1}^{k} y_i \leq t \qquad (k = 1, 2, \ldots, 48) \qquad (3.44)$$

where r_k is the required number of servers for period k, g is a large number and t is the number of allowable shifts.

Constraint 3.42 guarantee that the number of working employees at period k are more than the number of required employees r_k , estimated in previous sections.

Constraint 3.43 removes the risk of assigning employees to a shift which is not active.

The last constraint, 3.44 sets the number of allowable shifts to keep the call center operations manageable. Otherwise, the optimization algorithm could have generated up to 48 that would bring about administrational problems. We will use five to eight shifts in our scenarios.

Optimum assignment problem then can be stated as finding a set of non-negative x_i that satisfies (3.42), (3.43) and (3.44) under the minimization goal (3.41). The model is an integer programming lineer optimization. It is easily solvable. The solution of this problem is the main workforce plan for our call center; shift starting times and the number of employees that will work at each shift.

Our framework has starts with the given input data, uses analytical methods to estimate number of agents, improves results of the agent estimations with a heuristic combined simulation and then finds the optimum shifts with an integer programming model. The next chapter explains how we ran all these steps in our study.

4. COMPUTATIONAL ANALYSIS

This section includes computations conducted in the model. The model was descriped in detail in section three. The framework requires solution of an analytic model, simulation of the call center, adjustment heuristic and optimization model.

4.1. Input Data

Call center models start with the demand forecast step. Applied to our work, this step is simply determining the number of arriving calls for each period in the planning horizon. We have received ACD reports from a private Turkish bank's inbound call center. One month full data consisted of the number of incoming calls, the average waiting time and average call transaction time for every period. The calls were grouped according to the transactions, so data was appropriate for a multiskill analysis but due to our focus, we assumed all incoming data as a single type of transaction.

Incoming calls seem to increase around regular working hours, from 09.00 to 18.00 (periods 18 to 36 in the table). So, more employees should be assigned for these hours. During night time, rate is lower than one.

Average transaction time is 150 seconds including preparation time for the agent and average patience time for the customers is 120 seconds. Using these values in our model, we also set our target service level as 80% and make our computations to reach this target. Now we will use these figures to determine work requirements for each period.

Table 4.1. Average number of incoming calls and λ

Period	Avg. # of Calls	λ (#/min)
1	64,42	2,15
2	44,23	1,47
3	26,42	0,88
4	15,32	0,51
5	10,03	0,33
6	7,71	0,26
7	5,10	0,17
8	4,61	0,15
9	3,65	0,12
10	3,52	0,12
11	3,10	0,10
12	3,42	0,11
13	5,84	0,19
14	10,58	0,35
15	18,55	0,62
16	35,00	1,17
17	74,32	2,48
18	169,81	5,66
19	338,26	11,28
20	406,42	13,55
21	491,77	16,39
22	529,74	17,66
23	572,32	19,08
24	599,74	19,99
25	536,16	17,87
26	532,65	17,75
27	541,39	18,05
28	618,94	20,63
29	631,68	21,06
30	618,10	20,60
31	622,68	20,76
32	607,23	20,24
33	605,13	20,17
34	573,13	19,10
35	545,68	18,19
36	478,45	15,95
37	378,74	12,62
38	358,65	11,95
39	322,13	10,74
40	284,87	9,50
41	275,74	9,19
42	249,16	8,31
43	222,94	7,43
44	193,48	6,45
45	171,74	5,72
46	152,06	5,07
47	126,48	4,22
48	95,39	3,18

4.2. Determination of Initial Number of Agents

According to the bank's 24 hour incoming call levels, we need to determine work requirements for each period so that we can find an optimum solution that will handle these calls. The method we use is described in section 3. 1.

We use M.S. Excel for determination of initial number of agents. M.S. Excel formulas and macros using Visual Basic Applications coding language helped us run the algorithm in seconds.

We delivered these parameters; $1/\mu = 2.5$ (min), $\mu = 0.4$ (number/min), $\alpha = 0.5$ (number/min), $r = 1000$ and target service level $P(S) = 0.8$. We define service level as the fraction of answered calls to the number of total incoming calls.

The method also required initial number of servers so that it would estimate a service level for that period and increase or decrease the number of agents in case of necessity. Our initial setting was 2 agents for each period since it is not allowed in the call center to let 1 agent stay alone. We increased the number of agents one by one until we reached our target service level.

The number of necessary servers and the final service level reached by the method is demonstrated in Table 4.

Table 4.2. Final number of servers and probability levels.

Period	Final s	Final P(S)
1	7,00	0,82
2	6,00	0,89
3	4,00	0,85
4	3,00	0,85
5	3,00	0,92
6	3,00	0,95
7	2,00	0,82
8	2,00	0,83
9	2,00	0,86
10	2,00	0,87
11	2,00	0,88
12	2,00	0,87
13	3,00	0,97
14	3,00	0,91
15	4,00	0,93
16	5,00	0,87
17	8,00	0,84
18	16,00	0,83
19	29,00	0,80
20	35,00	0,82
21	42,00	0,82
22	45,00	0,81
23	48,00	0,80
24	51,00	0,82
25	46,00	0,83
26	45,00	0,81
27	46,00	0,82
28	52,00	0,81
29	53,00	0,81
30	52,00	0,81
31	53,00	0,82
32	51,00	0,80
33	51,00	0,81
34	49,00	0,83
35	46,00	0,81
36	41,00	0,82
37	33,00	0,83
38	31,00	0,81
39	28,00	0,81
40	25,00	0,82
41	24,00	0,81
42	22,00	0,81
43	20,00	0,82
44	18,00	0,84
45	16,00	0,83
46	14,00	0,81
47	12,00	0,81
48	10,00	0,85

The figures in the table are also demonstrated in the chart below.

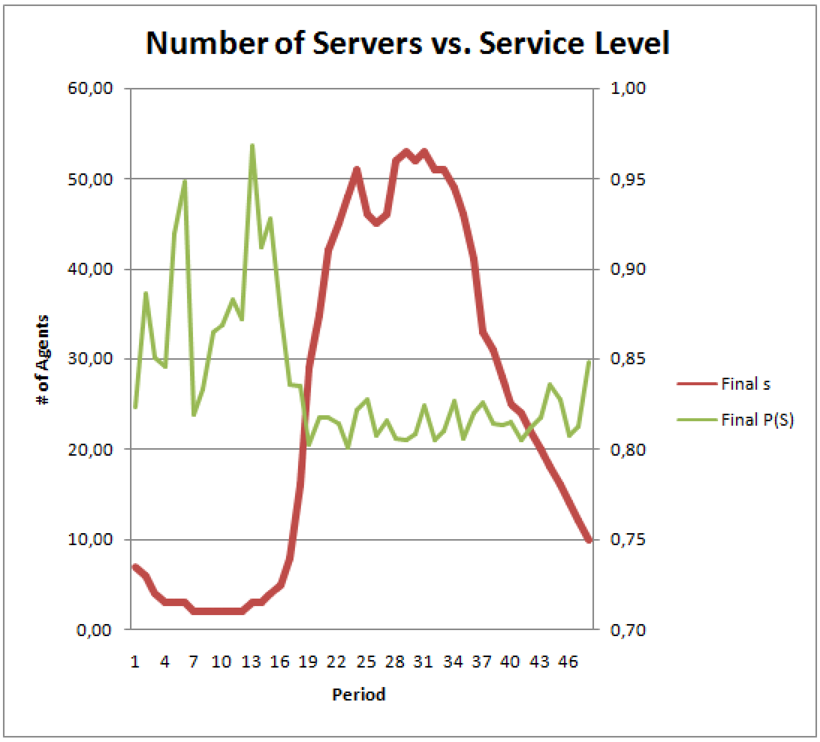

Figure 4.1. Number of servers vs. service level.

Whitt's algorithm shows us the need of 53 agents at maximum. There is a sharp increase in necessity between 8.00 to 11.00 in the morning and a decrease starting at 16.00.

It is also reasonable that at higher number of agents, achieved service level decreases to levels of 80%, but when there are only a couple of working agents, the algorithm cannot be able to keep service level under 90% since a decrease in the number of agents would decrease service level to lower than target.

The result of the method can now be tested by the simulation.

4.3. Simulation and Adjustment Heuristic

We have simulated the results of agent requirement estimations in Arena. The model we built in the tool could generate inter-arrival time, patience time and process time from exponential distributions. We choose to run the simulation for 2400 hours equivalent to 1000 replications for each run.

The simulation wrote the simulation time and period of processed and abandoned calls into text files, then the text files are imported into M.S. Excel and using pivot tables, simulated service levels are found for each period.

The table below contains the number of agents that are simulated. Shadowed cells are stabilized number of agents who can handle more than target percentage of calls. The algorithm increases and decreases the number of agents based on the simulation results listed below. For example, the first period generated a service level of 92% in the first run with 7 agents. The heuristic tried to decrease seven agents by 12% -which is excess-. Although 12% of seven is lower than one, the heuristic decreases the nearest larger integer, so six agents are simulated in the second run. In the second run, again the service level was exceeded, thus, five agents are simulated in the third run. But now, service level decreased to 77% which is lower than our target. The algorithm this time increases the number of agents by one again, which is six the second time. Since we just simulated five, we know that it was insufficient, six is our appropriate number of employees for this period and it stays so until the end of heuristic.

Table 4.3. Simulated number of agents.

Period	Sim1	Sim2	Sim3	Sim4	Sim5	Sim6	Sim7	Sim8
1	7	6	5	6	6	6	6	6
2	6	5	4	3	4	4	4	4
3	4	3	4	4	4	4	4	4
4	3	2	3	3	3	3	3	3
5	3	2	3	3	3	3	3	3
6	3	2	3	3	3	3	3	3
7	2	2	2	2	2	2	2	2
8	2	2	2	2	2	2	2	2
9	2	2	2	2	2	2	2	2
10	2	2	2	2	2	2	2	2
11	2	2	2	2	2	2	2	2
12	2	2	2	2	2	2	2	2
13	3	2	3	3	3	3	3	3
14	3	2	3	3	3	3	3	3
15	4	3	4	4	4	4	4	4
16	5	4	3	2	3	3	3	3
17	8	6	5	6	6	6	6	6
18	16	13	11	12	11	12	12	12
19	29	24	21	22	23	22	23	23
20	35	29	26	27	28	27	28	28
21	42	35	31	33	32	33	33	33
22	45	38	33	35	36	35	36	36
23	48	41	36	38	39	38	39	39
24	51	43	38	40	41	40	41	41
25	46	39	34	36	37	36	37	37
26	45	38	34	36	37	36	37	37
27	46	39	34	36	37	36	37	37
28	52	44	39	41	42	41	42	42
29	53	45	40	42	43	42	43	43
30	52	44	39	42	41	42	42	42
31	53	45	39	42	43	42	43	43
32	51	43	38	41	42	41	42	42
33	51	43	38	41	40	41	41	41
34	49	41	36	39	38	39	39	39
35	46	39	35	37	38	37	38	38
36	41	35	31	33	32	33	33	33
37	33	28	25	26	27	26	27	27
38	31	26	23	25	24	25	25	25
39	28	24	21	22	23	22	23	23
40	25	21	19	20	19	20	21	21
41	24	20	18	19	20	19	20	20
42	22	19	17	18	17	18	17	18
43	20	17	15	16	15	16	16	16
44	18	15	13	14	13	14	14	14
45	16	13	11	12	13	12	13	13
46	14	12	10	11	12	11	12	12
47	12	10	9	10	9	10	10	10
48	10	8	7	8	8	8	8	8

Table 4.4. Simulation results (Service levels).

Period	Sim 1	Sim 2	Sim 3	Sim 4	Sim 5	Sim 6	Sim 7	Sim 8
1	0,92	0,86	0,77	0,86	0,86	0,86	0,86	0,86
2	0,96	0,86	0,82	0,68	0,82	0,82	0,82	0,82
3	0,95	0,79	0,95	0,95	0,95	0,95	0,95	0,95
4	0,96	0,71	0,96	0,97	0,97	0,97	0,97	0,97
5	0,99	0,67	0,99	0,99	0,99	0,99	0,99	0,99
6	1,00	0,62	1,00	1,00	1,00	1,00	1,00	1,00
7	0,99	0,99	0,99	0,99	0,99	0,99	0,99	0,99
8	0,99	0,99	0,99	0,99	0,99	0,99	0,99	0,99
9	0,99	1,00	1,00	1,00	1,00	1,00	1,00	0,99
10	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
11	1,00	0,99	0,99	0,99	1,00	1,00	1,00	1,00
12	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
13	1,00	0,56	1,00	1,00	1,00	1,00	1,00	1,00
14	0,99	0,68	0,99	0,99	0,99	0,99	0,99	0,99
15	0,99	0,79	0,99	0,99	0,99	0,99	0,99	0,99
16	0,97	0,85	0,81	0,62	0,81	0,82	0,81	0,81
17	0,95	0,86	0,75	0,84	0,84	0,84	0,84	0,84
18	0,95	0,90	0,76	0,82	0,76	0,81	0,81	0,82
19	0,94	0,90	0,76	0,78	0,82	0,78	0,81	0,82
20	0,95	0,90	0,76	0,79	0,81	0,79	0,82	0,82
21	0,95	0,91	0,76	0,80	0,78	0,80	0,80	0,80
22	0,94	0,91	0,74	0,79	0,81	0,79	0,81	0,81
23	0,94	0,91	0,75	0,79	0,81	0,79	0,81	0,81
24	0,95	0,91	0,76	0,79	0,81	0,80	0,81	0,81
25	0,94	0,91	0,74	0,79	0,80	0,78	0,81	0,81
26	0,94	0,90	0,75	0,80	0,82	0,80	0,81	0,82
27	0,94	0,91	0,75	0,79	0,81	0,79	0,81	0,80
28	0,94	0,91	0,76	0,79	0,81	0,79	0,81	0,81
29	0,94	0,91	0,75	0,79	0,81	0,79	0,81	0,81
30	0,94	0,90	0,75	0,80	0,79	0,81	0,80	0,80
31	0,95	0,91	0,74	0,80	0,82	0,80	0,82	0,82
32	0,94	0,90	0,74	0,80	0,82	0,80	0,82	0,81
33	0,94	0,90	0,74	0,80	0,78	0,80	0,80	0,80
34	0,94	0,90	0,74	0,80	0,78	0,80	0,80	0,80
35	0,93	0,90	0,76	0,80	0,82	0,80	0,82	0,82
36	0,94	0,91	0,76	0,80	0,78	0,80	0,80	0,80
37	0,93	0,90	0,76	0,79	0,81	0,79	0,81	0,81
38	0,93	0,90	0,75	0,81	0,78	0,81	0,81	0,81
39	0,93	0,90	0,75	0,79	0,82	0,79	0,82	0,82
40	0,93	0,89	0,77	0,80	0,76	0,80	0,83	0,83
41	0,93	0,89	0,75	0,79	0,83	0,79	0,83	0,83
42	0,93	0,90	0,78	0,81	0,78	0,81	0,78	0,81
43	0,93	0,89	0,77	0,81	0,77	0,81	0,81	0,81
44	0,93	0,90	0,75	0,80	0,76	0,81	0,80	0,80
45	0,93	0,88	0,72	0,78	0,83	0,78	0,83	0,83
46	0,92	0,89	0,73	0,80	0,85	0,79	0,84	0,85
47	0,92	0,87	0,78	0,83	0,78	0,83	0,82	0,84
48	0,94	0,87	0,76	0,84	0,84	0,85	0,84	0,84
Total	0,94	0,90	0,75	0,80	0,80	0,80	0,81	0,81

The simulation results of all periods have reached above 80% at our last run, thus, heuristic has ended in eight runs. Besides independent periods' service levels, 81% of overall service level is also achieved.

Final number of agents is determined by this heuristic combined simulation. At least these numbers are required for the related periods in order to answer more than 80% of the incoming calls.

4.4. Optimization Model

We used GAMS tool to solve the integer programming problem of our call center model. The used model is described in section 3.4. We have introduced probabilistic lunch breaks and ran the simulation with shift constraint (3.44) changing from three to 48.

Non-Probabilistic vs. Probabilistic Lunch Breaks are illustrated in Appendix B Table B.1-2.

This tables show shift starting times and the number of employees allocated to that shift. For instance, at the strict meal break four shift model, 6 agents will start their shifts at period 1: 00.00 a.m.. On the other hand, at the probabilistic meal break for shift model, 12 employees will start their first shifts at period 7: 03.30 a.m.. Total workforce usage is shown at the end of the table.

Our probabilistic break approach save us 16 agents in four shift case. Savings are two, one and two agents for maximum five, siz and seven shifts respectively. Beside this analytical explanation, letting all the agents go to meal break at the same time is not a sensible management approach, realistic planning would distribute loss of work according to incoming calls in a call center.

Now we will analyze computational results for the probabilistic break approach.

Computational results of the final optimization model is provided in Appendix B Table B.3.

Naturally, the problem was infeasible with two shifts since the model needs coverage for 48 periods. The most sensible solutions seem to be five, six or seven shifts since less shifts require too many agents and more shifts are hard for administrative perspectives.

To manage our call center in four shifts, 85 agents are necessary although we know that our peak requirement was 43 at periods 29 and 31. The shortage at shifts cause the algorithm to use twice as our peak. Five shifts decrease the necessity to 80 agents and six shifts decrease the need to 76. If our model had no constraints for shifts, we could lower the need to 66 employees with at 20 shifts. Thus, 66 is our lower bound.

The first reasonable shift plan mean two additional shifts except obligatory three shifts. Starting times of these shfits are; 4 agents at 00.00, 5 at 06.30, 34 at 09.00, 19 at 12.00 and 18 at 16.00. Main shifts are 09.00, 16.00 and the night shift 00.00. This is an appropriate and manageable planning for a call center.

Six shifts require three additional shifts. Starting times of all six shifts are; 2 agents at 00.30, 8 at 06.30, 32 at 09.30, 13 at 12.00, 5 at 13.00 and 14 at 16.00. Distribution of these shifts are similar to five shift plan. Now, the algorithm was able to start five agents at 13.00 to handle another peak and lower total requirement to 76 agents.

Seven shifts is harder for administration. Now, there are four additional shifts which make transportation, meal and even shift control extremely hard to be provided by the company. Agents have to be well informed about the rules of the shifts and obey these rules. Starting times of the shifts are; 12 agents at 05.30, 12 at 09.00, 15 at 09.30, 13 at 12.00, 10 at 14.30, 9 at 18.30 and 2 at 20.30. Total number of workforce is 73 agents, three agents lower than six shifts. The difference could be considered worthless by management since shift starting times are more problematic now.

The managers of the call center should analyze how to provide agent rigths for especially additional shifts and see how much cost does every single additional shift cause. Then, one of our plans can be implemented to achieve successful operations.

4.5. Managerial Insights

Operational management of call centers require decisions of the number of shifts, shift starting times, the number of employees that will work at each shift and weekly allocation of employees. Our call center model is an end-to-end model; starting with the incoming calls, covering all necessary decisions and resulting in the shift starting times and the number of employees for each shift, which can be used as a support tool for these decisions.

4.5.1. How to Use This Framework

Our framework is able to calculate the necessary number of agents and their shift starting times when a forecast of incoming calls and the constraint of the number of shifts are provided. A call center can forecast incoming demand based on its ACD system, and directly adapt this model to see the number of necessary agents at a specific level of shifts. The number of shifts could be changed in the model with the same forecast so that a good decision for the number of shifts can be reached.

4.5.2. On Number of Shifts

The number of shifts is a key decision in call centers. In our model, number of shifts is a parameter, but we have investigated the effects of the number of shifts to efficiency and optimization based on our input data. Firstly, a 24-hour operating call center needs at least three shifts, otherwise it is not possible to handle calls in all day. Thus, our model is infeasible with two shifts. With three shifts, there are no reserve agents for meal breaks which make the number of necessary agents very large. Increasing the number of the shifts enables algorithm to assign shifts to urgently requiring periods so that agents can be used more optimally. Theoretically, a maximum number of 48 shifts are possible to minimize the number of necessary employees.

On the other hand, increasing the number of shifts causes problems for operational management. Initially, transportation and meals should be handled for each shift. More importantly, controlling working hours of agents, planning and communicating the plans to the agents is harder to manage when there are too many shifts. Of course, a balance has to be kept between these constraints. A way to keep this balance is to see the changes in necessary number of agents. The marginal change of the number of agents between two shift levels should be considered. For instance, the increase of number of shifts from four to five could generate an efficiency gain of two or fifteen agents. The final decision would be different in these cases.

According to our model, five or six shifts seem to handle calls in an optimal way. Three shifts are obligatory and two or three additional shifts decrease the number of agents with an efficient magnitude. Lower number of shifts cause too many employees and higher number of shifts could require more managerial staff. Naturally, in case a company can bear the problems of more shifts, it would be better to plan more number of shifts.

Eventually, the best way to make this decision is considering the operational constraints of the company and the efficiency gain that each shift generates.

4.5.3. Suggestions on Break Policy

An important aspect for the operational management of call centers is breaks. Researches on this topic generally neglected this area but from the workforce management perspective, breaks cause large gaps in daily shifts. A standard shift is nine hours including one full hour break, except coffee breaks of the employees. The full hour meal break could be handled without a problem in case call center ceases its operations. Unfortunately, operation continues 24 hours all year. Then, it becomes a must to manage the meal breaks so that some employees will stay for the operations during others' lunch time. This leads the call center to distribute meal breaks to two or three hours and let the agents give their break sequentially. This is the main idea in our probabilistic meal break approach.

Since distribution of meal breaks is the obvious solution for the non-stop working call centers, researches should be incorporated with more flexible and applicable models. From this perspective, our model provides acceptable results which enables operational managers use this framework for workforce planning and allocation.

4.5.4. Application of the Model

Our model suggests the number shifts, shift starting times and the number of employees for the given demand levels. So, in order to use our framework an estimation of incoming calls should be ready. Then, the model should be run with changing the shift constraints from three to the highest manageble number of shifts.

Final results are the total number of employees at each shift constraint. Efficiency levels would increase together with the number of allowable shifts. The remaining decision of the operational manager is then, the number of shifts.

5. CONCLUSION AND FURTHER STUDIES

Call center industry, although being quite new with a a history of only five decades, is an important business area exceeding $340 billion in the world. The market is growing especially in developing economies such as Turkey. Since growth in call center means growth in budgets whose major share of the costs are employee costs, optimization of resources and efficiency gain vital importance. In order to increase efficiency and generate savings, analysis and optimization models for call centers are necessary.

Reasonably, there has been a lot of publications in this area. Researchers have paid much attention, starting with Erlang's formula in 1917 (Takacs, 1969) and increasing in the last two decades (Callaghan and Thompson, 2002). Most of this research, like this study, take inbound call centers into consideration. Generally, inbound call center problems are handled in four steps; demand forecast, workforce requirement determination, shift determination and employee allocation.

Demand forecast is the part in which number of incoming calls are estimated. Advances in technology and usage of ACD's made it possible to analyze exact historical data for accurate forecasts. Second part is workforce requirement determination. It is the most popular step in literature for call centers. Queueing and Markovian and even simulation models are used for workforce requirement determination. Next parts are optimization and scheduling models which determine the number of shifts and allocate employees to these shifts.

The goal of this study is analysis of a call center and application of an end-to-end model based on real data from a private Turkish bank. Our framework consists of work requirement determination based on a Markovian approximation, simulation, adjustment heuristic for the simulation, shift determination and employee allocation. From this perspective, this thesis proposes an integrated multidimensional industrial engineering model related to stochastic processes, simulation and optimization models.

Our contribution to literature is investigating the research points in determination of possible shifts. We propose probabilistic lunch breaks instead of exact lunch plans for the agents in order to decrease loss of workforce. In business life, daily plan of the agents is both a problem and advantage for the managers. It is a problem since it is hard to manage the operations during the agents' meal times but it is an advantage because not like academic papers, agents' schedules can be used very flexibly in daily operations.

The model we propose in this thesis has unique aspects such as integrating the Markovian Approximation with the simulation model, adjustment heuristic and a probabilistic lunch break optimization model. Moreover, we make an application of a real company's data and observe the results. In literature, most papers are just related to one part of the call center problem. This paper is a full model of call centers and it can serve as a decision support tool for call center managers.

Call center problem has attracted a lot of attention from researchers, but still there is a lot of room for research starting with the first step of call center approaches: demand forecast. Unfortunately it has been disregarded by many models despite having a strong effect. We also have neglected demand forecast in our study. Certainly, forecasting models are needed to be integrated to call center studies.

Another main research area is multiskill routing. Only minority of the papers take routing into consideration since it increases complexity of the problems. Whereas in business life, it is not possible to handle the calls and employee training without distributing the skills to the agents.

Besides being a common application in business, one important and very rarely studied area is customer prioritization. In call centers, incoming calls are prioritized based on the importance of the customer in the company's CRM levels. Some customers are more important in financial means, than the others for the company. Thus, in case they make a call to the call center, they are not kept in the same queue, they wait less than any other customer. Remarkably, it is also possible to model agent prioritization to the calls. In this case, agents do not handle the calls randomly, there is a choice and options for the ACD to assign a call to the agent that can handle the call in the best way possible.

Although literature works include scheduling algorithms, weekly scheduling still has space for development. Weekly scheduling decisions have an impact on satisfaction of the agent requirements, since scheduling algorithms will provide adequate resting spaces for the employees. In fact, dynamic and presonalized scheduling and agent determination algorithms are necessary since instead of stable shifts, weekly determination of agent shifts and determination of each agent seperately will change optimization levels.

Based on experience, part-time, full-time, over-time, coffee and meal break determination is also a key research area while it has not been taken into consideration by many studies. In real business life, meal breaks, coffee breaks and over-time decisions are very flexible. In case there are too many incoming calls, agents may use their meal breaks in 20 minutes and skip their coffee breaks in the day. Moreover, some agents can stay for working over-time. Although integrated probabilistic lunch breaks for the agents in order to decrease loss of workforce, most of these options are not yet explored.

One final direction for research is end-to-end models. Call center model has different sub-problems, but from the business perspective, only the models covering the whole process will be adaptable.

All in all, call center literature has been published for more than 40 years with a sharp increase in the last two decades. But still, there are major areas which has room for research in the future.

APPENDIX A: GAMS OPTIMIZATION CODE

```
set
 i 'periods for starting times' /1*48/
 k 'periods for the demand'    /1*48/
;
PARAMETERS
a(i,k) determines if the agent starting his shift at period is working at period k /
$include C:\Users\Asus\Documents\gamsdir\projdir\callcenterAik.inc
/
r(k) required number of agents for period k /
$include C:\Users\Asus\Documents\gamsdir\projdir\callcenterRk.inc
/
;
VARIABLES
X(i) the number of agents who will start their shift at period i
y(i) 1 if there is an active shift starting at period i, 0 otherwise
zobj objective value
integer variable x(i)
binary variable y(i)
;
EQUATIONS
objectivefunc objective
constraint1(k)
constraint2(i)
constraint3
;
objectivefunc.. zobj =e= sum(i,x(i));

constraint1(k).. sum(i,x(i)*a(i,k)) =g= r(k);
constraint2(i).. x(i) =l= 10000*y(i)   ;
constraint3.. sum(i,y(i)) =l= 5        ;

model ccopt /all/
option MIP=CPLEX;
solve ccopt using MIP minimizing zobj;
```

APPENDIX B: LUNCH BREAK COMPARISON

Table B.1. Non-probabilistic versus probabilistic lunch breaks periods 1-24.

	Strict	Prob.	Strict	Prob.	Strict	Prob.	Strict	Prob.
Period\ Number of Shifts	4	**4**	**5**	**5**	**6**	**6**	**7**	**7**
1	6			4				
2						4		
3								
4								
5								
6								
7		12						
8								
9							4	
10								
11					12			12
12			12					
13				5		8		
14							8	
15								
16								
17								
18	42			34	24			12
19		30	29			32	19	15
20								
21							13	
22								
23					6			
24				19		13		13

APPENDIX B: EMPLOYEE ALLOCATION RESULTS

Table B.2. Non-probabilistic versus probabilistic lunch breaks periods 25-48.

	Strict	Prob.	Strict	Prob.	Strict	Prob.	Strict	Prob.
Period\ Number of Shifts	4	**4**	**5**	**5**	**6**	**6**	**7**	**7**
25	37	27			19			
26			25			5	16	
27								
28								
29								10
30								
31								
32			12	18	12	14		
33							10	
34								
35								
36								
37								9
38								
39								
40							5	
41								2
42		16	4		4			
43	16							
44								
45								
46								
47								
48								
Total	**101**	**85**	**82**	**80**	**77**	**76**	**75**	**73**

Table B.3. Final optimization results.

Period\ Number of Shifts	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**20**	**48**
1			4							
2				4						
3	6									
4										
5										
6										
7		12								
8										
9										
10										
11					12	5		6	6	4
12							6			1
13			5	8					2	1
14									3	2
15								3		
16										
17						4				
18	82		34		12	22	20	6	8	9
19		30		32	15		6	21	13	11
20							4		2	4
21						10		2	2	3
22									3	3
23							3	5	2	2
24			19	13	13	13	13	8	5	6
25		27							3	2
26				5					2	2
27								6		2
28							7		2	1
29					10	6			2	2
30									1	2
31									2	2
32			18	14		8	4	9	2	3
33									1	
34										1
35										
36	36								2	
37					9					1
38										
39										
40										
41					2	4		4		2
42		16					8		3	
43										
44										
45										
46										
47										
48										
Total	**124**	**85**	**80**	**76**	**73**	**72**	**71**	**70**	**66**	**66**

REFERENCES

Adams, J., S., 1963, “Toward an Understanding of Inequity”, *Journal of Abnormal and Social Psychology*, Vol. 67, No. 5, pp. 422-436.

Angus, I., 2001, “An Introduction to Erlang B and C”, *Telemanagement 187*.

Atlason, J., Epelman, M. A., Henderson, S G., 2008, “Optimizing Call Center Staffing Using Simulation and Analytic Center Cutting-Plane Methods”, *Management Science,* Vol. 54, No. 2, February 2008, pp. 295-309.

Avramidis, A. N., Chan, W., Gendrau, M., L’Ecuyer, P., Pisacane, O., 2010, “Optimizing Daily Agent Scheduling in a Multiskill Call Center”, *European Journal of Operational Research,* 200 (2010), pp. 822–832.

Bacceli, F., Hebuterne, G., 1981, “On Queues With Impatient Customers”, *Performance ’81*.

Beaumont, N., 1997, “Scheduling Staff Using Mixed Integer Programming”, *European Journal of Operational Research,* 98 (1997), pp. 473-484.

Borst, S., Mandelbaum, A., Reiman, M. I., 2004, “Dimensioning Large Call Centers”, *Operations Research,* Vol. 52, No.1.

Brandt, A., Brandt, M., 2002, “Asymptotic Results and a Markovian Approximation for the M(n)/M(n)/s + GI System”, *Queueing Systems,* 41, 73–94.

Brusco, M. J., Jacobs, L. W., 1998, “Personnel Tour Scheduling When Starting - Time Restrictions are Present”, *Management Science,* Vol. 44, No.4, April 1998.

Brusco, M. J., Jacobs, L. W., 2001, “Starting-Time Decisions in Labor Tour Scheduling: An Experimental Analysis and Case Study”, *European Journal of Operational Research,* 131 (2001) pp. 459-475.

Brusco, M. J., Jacobs, L. W., Bongiorno, R. J., Lyons, D. V., Tang, B., 1995, “Improving Personnel Scheduling At Airline Stations”, *Operations Research,* Vol. 43, No.5, September-October 1995.

Buffa, E.S., Cosgrove, M.J., Luce, B.J., 1976. “An integrated work shift scheduling system”, *Decision Sciences*, 7 (4), pp. 620– 630.

Cagrimerkezidernegi.org, 2008, *Dünya ve Türkiye Çağrı Merkezi Sektörü*, http://cagrimerkezleridernegi.org/uploads/Dunya-ve-Turkiye-Cagri-Merkezi-Sektoru.pdf, 6.02.2011.

Callaghan, G. and Thompson, P., 2002, “We Recruit ‘Attitude’: The Selection and Shaping of Routine Call Center Labour,” *Journal of Management Studies*, 39 (2), pp. 233-254.

Dantzig, B., 1954, “A Comment on Edie's Traffic Delays At Toll Booths”, *Journal of the Operations Research Society of America*, Vol. 2, No. 3 (Aug., 1954), pp. 339-341.

Dean, A. M., 2007, “The Impact of the Customer Orientation of Call Center Employees on Customer’ Affective Commitment and Loyalty”. *Journal of Service Research*, Vol. 10, No. 2, November 2007, pp. 161-173.

Deery, S., Iverson R. and Walsh. J., 2002, “Work Relationships in Telephone Call Centres: Understanding Emotional Exhaustion and Employee Withdrawal,” *Journal of Management Studies*, 39 (4), pp. 471-496.

Deslauriers, A., Pierre L'Ecuyer, P., Pichitlamken, J., Ingolfsson, A., Avramidis, A., N., 2005, *Markov chain models of a telephone call center with call blending*. http://apps.business.ualberta.ca/aingolfsson/documents/PDF/ctmc1.pdfi, March 2011.

Dickson, D. R., Markova, G, Bohn, T., 2010, "Does Customer Service Make a Difference? An Empirical Study of a Professional Association". *Journal of Applied Management and Entrepreneurship*; Jul 2010, p. 62.

Diets, D. C., 2011, *Practical Scheduling for Call Center Operations*, www.elsevier.com/locate/omega, March 2011.

Easton, F. F., Rossin, D. F., 1991, "Equivalent Alternate Solutions for the Tour Scheduling Problem", *Decision Sciences*, Nov/Dec 1991, 22, 5, p. 985.

Edie, L. C., 1954, "Traffic Delays at Toll Booths", *Journal of the Operations Research Society of America*, Vol. 2, No. 2(May., 1954), pp. 107-138.

Feldman, Z., Mandelbaum, A., 2008, "Staffing of Time-Varying Queues to Achieve Time-Stable Performance", *Management Science* doi 10.1287/mnsc.1070.0821ec pp. ec1–ec20.

Fleischer, J., 2006, "It's Not the Cost It's the Productivity", *Call Center Magazine*, Apr 2006, 19, 4, p. 12.

Gans, N., Koole, G., Mandelbaum, A., 2003, "Telephone Call Centers: Tutorial, Review and Research Prospects", *Manufacturing & Service Operations Management* 5:79-141.

Green, L., V., Kolesar, P. J., Soares, J., 2003, "An Improved Heuristic For Staffing Telephone Call Centers With Limited Operating Hours", *Production and Operations Management*, Vol. 12, No. 1, Spring 2003.

Green, L., V., Kolesar, P. J., Soares, J., 2001, "Improving the SIPP Approach for Staffing Service Systems That Have Cyclic Demands", *Operations Research*, Vol. 49, No.4, July-August 2001.

Helber, S., Henken, K., 2010, "Profit-Oriented Shift Scheduling of Inbound Contact Centers With Skills-Based Routing, Impatient Customers and Retrials", *OR Spectrum*, 32.109-134, DOI 10.1007/s00291-008-0141-8.

Hishinuma, C., Kanakubo, M., Goto, T., 2007, "An Agent Scheduling Optimization for Call Centers", In: IEEE, 2007 IEEE Asia-Pacific Services Computing Conference.

Hurriyet, 2011a, *Çağrı Merkezinde Kiralama Dönemi*, http://www.hurriyet.com.tr/ekonomi/17518557.asp?gid=384, 19.04.2011.

Hurriyet, 2011b, *Çağrı Merkezleri Teşvik Kapsamına Alındı*, http://www.hurriyet.com.tr/ekonomi/17585502.asp, 19.04.2011.

Ingolfsson, A., Campello, F., Wu, X., Cabral, E., 2010, "Combining Integer Programming and the Randomization Method to Schedule Employees", *European Journal of Operational Research*, 202 (2010), pp. 153–163.

Ingolfsson, A., Haque, M. A., Umnikov, A., 2002, "Accounting for Time-Varying Queueing Effects in Workforce Scheduling", *European Journal of Operational Research*, 139 (2002), pp. 585-597.

Jennings, O. B., Mandelbaum, A., Massey, W. A., Whitt, W., 1996, "Server Staffing to Meet Time-Varying Demand", *Management Science*, Vol 42, No 10, October.

Kolesar, P. J., L. V. Green, 1998, “Insights on Service System Design From a Normal Approximation to Erlang’s Delay Formula”, *Production Oper. Management,* 7, pp. 282–293.

Kolesar, P. J., Rider, K. L., Crabill, T. B., Walker, W. E., 1975, “A Queueing-Linear Programming Approach to Scheduling Police Patrol Cars”, *Operations Research*, Vol. 23, No. 6, November-December.

Koole, G., Mandelbaum, A., 2002, “Queueing Models of Call Centers: An Introduction”, *Annals of Operations Research*, 113, pp. 41-59.

Lywood, J., Stone, M., Ekinci, Y., 2009, “Customer experience and profitability: An application of the empathy rating index (ERIC) in UK call centres”, *Journal of Database Marketing & Customer Strategy Management*, 16, pp. 207 – 214. doi: 10.1057/dbm.2009.24.

Mandelbaum, A., 2004, *Call Centers Research Bibliography With Abstracts*, http://ie.technion.ac.il/serveng, 26.04.2011.

Mandelbaum, A., Momcilovic, P., 2009, *Queues With Many Servers and Impatient Customers*, http://iew3.technion.ac.il/serveng/References/MM0309.pdf , 11.03.2011.

Mandelbaum, A., Zeltyn, S., 2005, *Service Engineering in Action: The Palm/Erlang-A Queue, with Applications to Call Centers*, http://iteseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.100.pdf, 13.02.2011.

Mason, A., J., Ryan, D. M., Panton, D. M., 1998, “Integrated Simulation, Heuristic and Optimization Approaches to Staff Scheduling”, *Operations Research*, Mar/Apr 1998; 46, 2, p. 161.

Pisacane, O., 2009, “Agent Scheduling in a Multiskill Call Center”, *4OR-A Quarterly Journal Operational Research*, 7:199-2002 doi 10.1007/s10288-008-0078-4.

Read, B. B., 2001, “Winning Outsourcing Plays”, *Call Center Magazine*, Mar 2001, 14, 3, p. 58.

Riggs, B., Thyfault, M. E., 1999, “The Modern Call Center”, *InformationWeek*, Oct 4, 1999, 755, p. 53.

S. Halfin and W. Whitt, 1981, “Heavy-Traffic Limits for Queues with Many Exponential Servers”. *Operations Research*, 29, pp. 567–587.

Takacs, L., 1969, “On Erlang’s Formula”, *The Annals of Mathematical Statistics*, Vol. 40, No. 1, pp. 71-78.

Taylor, P. E., Huxley, S. J., 1989, “A Break from Tradition for the San Francisco Police: Patrol Officer Scheduling Using an Optimization-Based Decision Support System”, *The Institute of Management Sciences*, doi 0091-2102/89/1901/0004501.25.

Whitt, W., 2005, *Sensitivity of Performance in the Erlang-A Queueing Model to Changes in the Model Parameters*, http://docs.google.com/viewer?a=v&q=cache:7xy7ucUchk0J:citeseerx.ist.psu.edu/viewdoc/download?doi%3D10.1.1.120.4624%26rep%3Drep1%26type%3Dpdf+Sensitivity+of+Performance+in+the+Erlang-A+Queueing+Model+to+Changes+in+the+Model+Parameter&hl=tr&gl=tr&pid=bl&srcid=ADGEESgSOknY4W4YxVaxquiUFKt7CahBFHeFBK6mwFz_xTD0D1mvQUHB9VZDafAXOC4Y_I7eHhX9J6Hsb2iSzUaVf9cRsBeqbwF0zOUX4IBIsNxMhxcAPbnh76OPMW6OG5AifJ-fkIDx&sig=AHIEtbQza0NG0q0ZXKln8Ky209jIQgAoIQ&pli=1, 21.10.2010.

Whitt, W., 1999, "Dynamic Staffing in a Telephone Call Center Aiming to Immediately Answer All Calls", *Operations Research Letters*, 24, pp. 205-212.

Whitt, W., 2005, "Engineering Solution of a Basic Call-Center Model", *Management Science*, Feb 2005, 51, 2, p. 221.

Wikipedia, 2010, *Agner Krarup Erlang*, http://en.wikipedia.org/wiki/Agner_Krarup_Erlang, 16.02.2011.

Wikipedia, 2011, *Call Centre*, http://en.wikipedia.org/wiki/Call_center, February 16.02.2011.

Zeltyn, S., Mandelbaum, A., 2005, *Call Centers With Impatient Customers: Many-Server Asymptotics of the M/M/n+G Queue*, http://iew3.technion.ac.il/serveng/References/references.html, 06.01.2011.