Proceedings of the 2003 Winter Simulation Conference
S. Chick, P. J. Sánchez, D. Ferrin, and D. J. Morrice, eds.

CALL CENTER SIMULATION MODELING: METHODS, CHALLENGES, AND OPPORTUNITIES

Vijay Mehrotra

Department of Decision Sciences
College of Business
San Francisco State University
1600 Holloway Avenue
San Francisco, CA 94123, U.S.A.

Jason Fama

Engineering Group
Blue Pumpkin Software Inc.
884 Hermosa Court
Sunnyvale, CA 94085, U.S.A.

ABSTRACT

Using stochastic models to plan call center operations, schedule call center staff efficiently, and analyze projected performance is not a new phenomenon, dating back to Erlang's work in the early twentieth century. However, several factors have recently conspired to increase demand for call center simulation analysis.

- Increasing complexity in call traffic, coupled with the almost ubiquitous use of Skill-Based Routing.
- Rapid change in operations due to increased merger and acquisition activity, business volatility, outsourcing options, and multiple customer channels (inbound phone, outbound phone, email, web, chat) to support.
- Cheaper, faster desktop computing, combined with specialized call center simulation applications that are now commercially available.

In this tutorial, we will provide an overview of call center simulation models, highlighting typical inputs and data sources, modeling challenges, and key model outputs. In the process, we will also present an interesting "real-world" example of effective use of call center simulation.

1 INTRODUCTION: "WHY CALL CENTERS?"

The trend in our economy from manufacturing towards services is well documented. One notable facet of this transition towards services has been the explosion of the call center industry.

Mehrotra (1997) defines call centers as "Any group whose principal business is talking on the telephone to customers or prospects." In this paper, we will refer to the individuals who talk on the phone with customers as "agents."

While the size of the industry is difficult to accurately determine, a plethora of statistics from diverse sources reflect that fact that this is a huge and growing global industry. Most stunning: Mandelbaum (2001) cites a study that an estimated 3% of the United States population works in this industry. Most recent: an explosion of outsourced call centers springing up in India, the Philippines, the Caribbean, and Latin America, serving overseas customers in the United States and Western Europe as well as growing domestic market needs.

From a mathematical perspective, call centers are interesting for a variety of reasons:

- Call centers typically handle more than one type of call, with each distinct call type referred to as a "queue" (as discussed below, this usage is not consistent with our normal definition of a queue).
- Inbound calls within each queue arrive at random over the course of time.
- In many centers, agents make outbound calls to customers, either proactively (typically for telemarketing or collections activities) or as a follow-up to previous inbound calls.
- Each call is of a random duration, as is the work (data entry, documentation, research, etc.) that agents must do after completing the phone call.
- Through Automatic Call Distribution ("ACD") and Computer Telephony Interaction ("CTI") devices, inbound calls can be routed to agents, groups, and/or locations, with advancements in these routing technology supporting more and more sophisticated logic over time.
- Individual agents can be skilled to handle one type of call, several types of calls, or all types of calls, with different priorities and preferences specified in the routing logic.

Thus, call centers can be thought of as stochastic systems with multiple queues and multiple customer types. As we discuss below, there are great challenges associated with managing these systems effectively.

To summarize, call centers are of interest both because of the sheer size of the industry, both in the United States and overseas, and because of the operational and mathematical complexity associated with these operations, which makes it difficult for decision makers to understand system dynamics without effective modeling.

The remainder of this tutorial is organized as follows. In Section 2, we motivate the need for and value of simulation in the context of effective call center management. In Section 3, we discuss how call centers make use of simulation models, focusing on the key output statistics that are used for system performance evaluation. In Section 4, we provide a modeling framework for call center simulation, and discuss the key inputs associated with simulation models, introducing the concepts through the formulation of a simulation model. In Section 5, we discuss business decisions associated with this model and explore some of the results of our analysis. Finally, in Section 6, we propose likely future directions for call center simulation.

Note: Throughout this paper, we will use the term "call center" and focus our discussion on centers that are processing only phone calls (either inbound, outbound, or both). Another common term in this industry is "contact center," which refers to centers handling not only phone calls but other types of customer contacts such as email, fax, paper, and/or chat sessions. We have chosen to focus on call centers here for clarity of exposition. However, leveraging the ideas presented here from phone-only call centers to multi-channel contact centers is a straightforward extension that we have also engaged in extensively.

2 CALL CENTER MANAGEMENT CHALLENGES AND THE NEED FOR MODELS

Those responsible for managing call centers face a very difficult set of challenges. At a high level, they must strike a balance between three powerful competing interests, as shown in Figure 1 below.

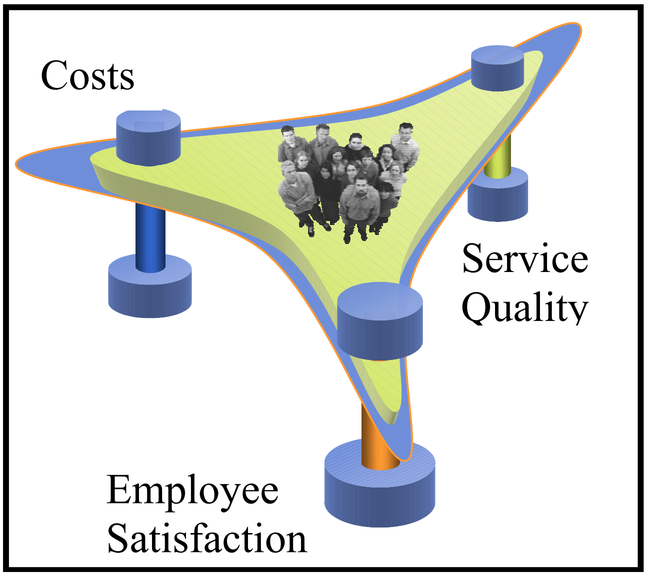

Figure 1: The Call Center Management Balancing Act

On a day-to-day basis, while simultaneously keeping costs, service quality, and employee satisfaction, these executives and managers must (implicitly or explicitly) answer a number of important questions for which decision support models are valuable:

- How many agents should we have on staff with which particular skills? How should we schedule these agents' shifts, breaks, lunches, training, meetings and other activities?
- How many calls of which type do we expect at which times?
- How quickly do we want to respond to each type of inbound call?
- How should we cross-train our agents? How should we route our calls to make the best use of these resources?
- Given a forecast, a routing design, and an agent schedule, how well will our system perform?
- What is our overall capacity? How will a spike in call volumes impact our overall performance?
- How is our center doing right now? What has changed since we did our last forecast and published our schedules? If the changes are significant, what can I do to respond to minimize the impact on the rest of the day or week?

There are a variety of mathematical methods (see Grossman et al. 2001 and Mandelbaum 2001 for more discussion of this) and associated software to help call center personnel as they try to address these types of questions, most notably workload forecasting models based on time series and agent scheduling optimization solutions.

However, over the past several years, simulation has emerged to play an important role in the call center design and management arena.

3 HOW CALL CENTERS USE SIMULATION

There are three major ways that simulation is utilized within the call center industry:

1. Traditional Simulation Analysis: A simulation model is built to analyze a specific operation, with inputs obtained from a variety of data sources, as discussed in Section 4 below.
2. Embedded Application – ACD/CTI Routing: Many of the leading ACD and CTI applications include a routing simulation to provide insights to routing design engineers about the impact of different decisions.
3. Embedded Application – Agent Scheduling: Already a complex scheduling problem (see Andrews and Parsons 1989 for a more detailed discussion), optimal call center agent scheduling is even more complex when both calls and agents are non-homogeneous. Many commercial scheduling software applications, including the one developed by

the authors' company, make use of simulation as part of their overall optimization engine.

In each of these cases, the key output statistics typically include the some or all of the following metrics:

- Queue Statistics: The two dominant queue statistics for inbound queues and call centers are Average Speed of Answer ("ASA") and Percent of Calls Answered with a queue time of less than some defined value ("PCA" or, more commonly, "Service Level)." Note that for each queue this statistic is interesting at the interval level (typically 15 minutes, 30 minutes, or one hour) and also at the aggregate daily and weekly levels; additionally, management is interested in the overall performance across a collection of queues that draw upon a common pool of agent resources.
- Abandonment Statistics: For most inbound call centers, particularly those focused on customer service and/or sales, a great deal of attention is paid to the overall number of customers who abandon (that is, hang up and thus leave the queue before being served). This is known to be a significant indicator of customer satisfaction (see Feinberg et al. 2000 for a recent published study on this). Many centers will look at the more restrictive metric of number of customers abandoning beyond the target Service Level parameter, based on the rationale that a certain waiting time in queue (as defined by the Service Level parameter, which ranges from 5 seconds to several minutes across companies and industries) is inevitable.
- Volume Statistics: For outbound queues and call centers, the real statistic of interest is Right Party Connects ("RPCs"). That is, for all of the attempted calls that were made, what percentage of these calls reached the targeted individual (as opposed to no answer, answering machine, or some other human being). Outbound contact center managers are typically interested in RPC on both an absolute and a percentage basis. For inbound queues, the Calls Handled statistic is of interest, and is easily derived by subtracting Abandoned Calls from the total number of incoming calls (referred to as "Offered Calls").

4 CALL CENTER SIMULATION MODELLING

4.1 Framework

The biggest challenge of call center simulation modeling is the definition and organization of model inputs. Figure 2 below illustrates our framework for call center simulation model definition and key inputs.

As reflected in Figure 2, call center simulation models feature a diverse range of inputs from multiple data sources, and as with all simulation designs, there are decisions to be made about the level of detail to include in the model.

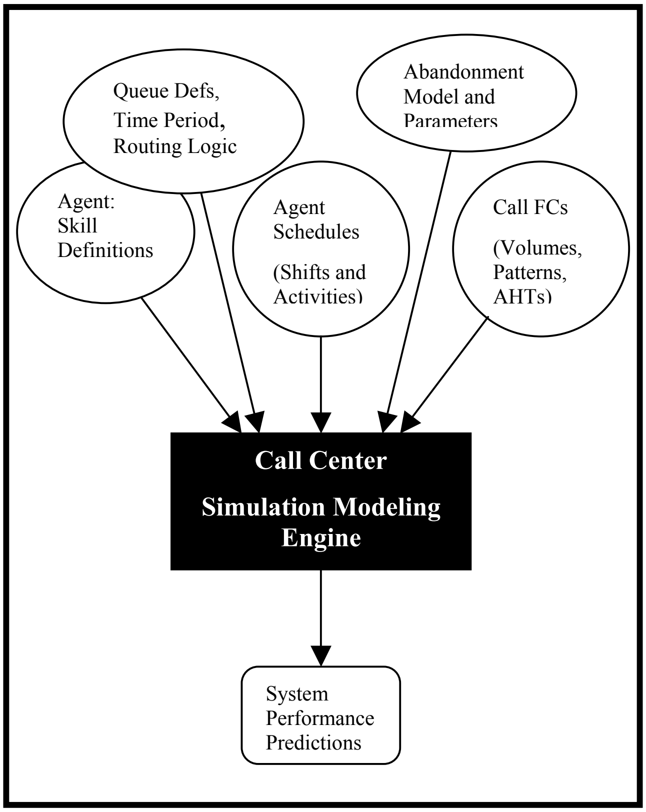

Figure 2: Call Center Simulation Modeling Framework

In the sections below, we discuss these key input areas in more detail. In the process, we will use our example model to illustrate these modeling concepts.

4.2 Key Inputs: Queue Definitions, Time Period, and Routing Logic

The basic building blocks of a call center simulation model are the calls, the agents, and the time period during which the call center is open. In turn, the basic routing logic connects the way that the calls interact with the people during that time period.

Typical call center simulation models contain more than one queue (as single queue models are ordinarily dealt with analytically) and run for a period of one day, one week, or multiple weeks.

Our example model is for a Collections call center. As is typical for call centers, this operation is part of a larger business context, in which creditors' records are being monitored on a regular basis for potential delinquency. Once a customer falls into delinquency, several things happen: (a) the information on the account is added to a list of prospects for an Outbound collections call; (b) they are notified about the state of their credit by mail; and (c) additional limitations may be placed on the account.

The call center itself features two queues: an Inbound queue and an Outbound queue. The time period that we are using for the analysis is one week.

The two agent groups and the basic routing logic are illustrated in Figure 3. Calls from the Inbound queue will arrive and be served by an agent from Group #1, the Inbound-Only group. If no agent from this group is available, the calls will wait in queue. If, after some predefined period of time, the Inbound call has not yet been served, it will then also queue for an agent from Group #2, the Cross-Trained Outbound group.

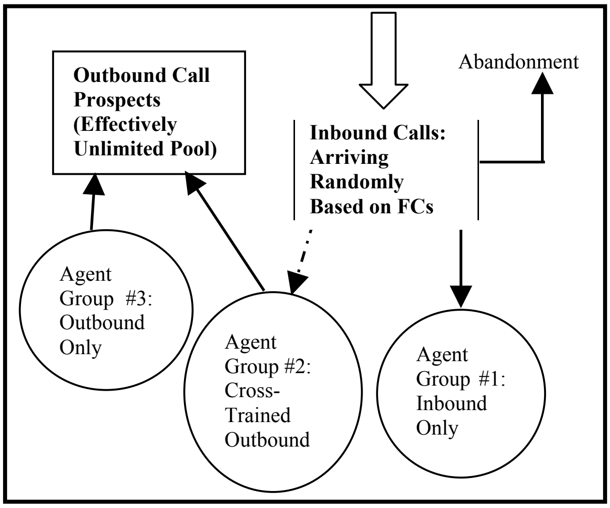

Figure 3: Queues, Agent Groups, and Routing Logic for Example Model

Meanwhile, Group #2 and Group #3 agents will be logged into a Predictive Dialer, which places Outbound calls to prospects from a list. When an answer is detected by the Predictive Dialer, the call is automatically routed to one of the Outbound Specialists or one of the available Cross-Trained Outbound agents. The agent then engages in a collections discussion (if they have reached the Right Party) or leave a message (if they have reached an answering machine or another party on the same phone number). Along with mailings to delinquent customers, these messages will generate some calls to the Inbound queue.

4.3 Key Inputs: Call Forecasts

Call forecasts are typically driven by a combination of historical data, time series model, and expert judgment. There are two major types of call forecasts: call volumes and Average Handling Time. Both are required for any basic call center simulation.

Due to the telecommunication and call center industries' history of using steady state M/M/n queue formulas to derive the number of agents needed for each time interval, it has been customary to translate call volume forecasts into λ values for Poisson arrivals and AHT forecasts into μ values for Exponential service times.

A great deal of research has been conducted on call volume forecasting models, and the interested reader is referred to Mabert (1985) and Andrews and Cunningham (1995) for valuable discussions on this topic.

Forecasts must be created for each queue for each time interval in the simulation period.

The most common call center forecasting approach is to create weighted averages of historical data for specific time intervals over the course of a week. For example, the initial call volume forecast for 8:15 a.m. - 8:30 a.m. next Tuesday might be computed as the average of call volume for the 8:15 a.m. - 8:30 a.m. period on the past several Tuesdays. From here, alterations may be (or more commonly, should be!) made based on additional information (e.g. specific marketing activities for a sales center or emerging product issues for a technical support center) that may cause volume to differ substantially from previous patterns.

4.3.1 Average Handle Time Forecasts

As mentioned earlier, most call center models assume that call handling times are exponentially distributed. We would recommend using more accurate distributional information about call handling times whenever possible. For example, it is common to find technical support call center for which call handling times are bi-modal (easy cases with a shorter mean, harder cases with a longer one).

However, the primary reason that the call center industry accepts the assumption of exponential handling times is because the ACD and CTI devices (the primary source for historical call volume data) store only *average* handling times at the interval level. With a dearth of consistent second moment information available, we have thus accepted this assumption far more often than we would like; in particular, we have modeled exponential handling times in the numerical example presented in Section 5.

<u>Note</u>: in this paper, we refer to Average Handling Time, or AHT. However, when obtaining data from ACD reports, it is not uncommon to find two fields that are then summed together to compute AHT: Average Talk Time ("ATT") and After Call Work ("ACW").

4.4 Key Inputs: Agent Schedules

Agent schedules can be thought of as a series of activities taking place over the course of a day. For example, an agent who comes to work at 8:00 am for a nine-hour shift may have a 15-minute break at 9:45 am, lunch at 11:30 am, an on-line training course from 1:00 – 2:00 pm, and a break at 3:15 pm before leaving work at 5:00 pm.

From a simulation perspective, each agent is viewed as a resource to perform certain types of work. Note that in the call center context, agents are actually productive only during the interval in which the agents are scheduled to be actually handling phone calls.

In addition, it is conventional to model agents as completing the task that they are engaged, even if it extends past the time at which they are to switch activities. That is, an agent within our simulation will be modeled as completing the phone call that he is working on before leaving for a break or a lunch.

A common step in call center simulation is to translate a set of individual agent schedules into a matrix of resources, where the dimensions of the matrix are defined by the number of Agent Groups and the number of Time Intervals.

In our example, we have leveraged the fact that our schedules are at a 15-minute level of granularity, and therefore prior to running the simulation we have converted these schedules into a number of on-phone agents for each group for each 15-minute interval.

4.5 Key Inputs: Abandonment Model and Parameters

Abandonment is one of the most hotly debated topics in call center management and research. There are two basic questions that must be answered in order to effectively model customer abandonment behavior:

1. What is the customer's tolerance for waiting, and at what point will this customer hang up and thereby leave the queue?
2. How likely is the customer to call back, and after how long?

Many researchers (e.g. Hoffman and Harris 1986, Andrews and Parsons 1993) have examined the challenge of modeling these problems from both an empirical point of view and from an analytic perspective.

From our experience, these questions are difficult to answer not only because of the mathematical complexity of the queue dynamics but also because of a lack of observable data about customer abandonment and retrial. While many surveys have been done, we have observed great differences in customer behavior across different industries and different companies' operations. In addition, information provided to callers about expected waiting time and/or position in queue can have a marked impact on abandonment behavior.

In our example model, simulated customers arrive at the call center and are served by an agent if one is available. If not, they join the queue, at which point they are also assigned a "life span" which is drawn from an exponential distribution. If a customer's life span expires while they are still waiting in queue, they then abandon the queue.

That is, we represent customers' tolerance for waiting in queue as an exponential random variable (as suggested by Garnett et al. 2002). We refer to the mean of this distribution as "the patience factor."

Given this modeling choice, we must still with the challenge of selecting the patience factor, which we estimate from historical data about callers' time in queue.

We do not include caller retrial in the example model.

4.6 Key Inputs: Agent Skills

Our definition of "Agent Skills" is comprised of three major types of inputs for each agent or group of agents:

1. What calls is the agent capable of handling?
2. Given a choice of multiple calls waiting, which will the agent handle ("call priority")
3. How fast will the agent be able to handle each type of call, and how often will the agent resolve the issue successfully ("call proficiency")

When combined with routing logic and call forecasts, these attributes fully specify the queueing model to be simulated.

In our example, we have three distinct groups of agents, each with different skills:

- Agent Group #1 (Inbound Only) handle only Inbound calls on a First-Come-First-Served basis. These agents have a call proficiency of 1.0 for Inbound Calls, so that their AHT is equal to the forecasted AHT for Inbound Calls.
- Agent Group #3 (Outbound Specialists) handle only Outbound calls. These agents have a call proficiency of 1.0 for Outbound Calls, meaning that their AHT is equal to the forecasted AHT for Outbound Calls.
- Agent Group #2 (Cross-Trained Outbound) handle both Inbound and Outbound calls. These agents have a call proficiency of 1.0 for Outbound Calls, meaning that their AHT is equal to the forecasted AHT for Outbound Calls. However, these agents will give priority to Inbound Callers if there are any waiting in queue, and have a call proficiency of 2.0 for Inbound calls, reflecting the relative inefficiency of cross-training (see Pinker and Shumsky 2000 for more discussion of this phenomenon both in and out of the contact center).

4.7 Other Modeling Considerations

4.7.1 Shrinkage

It is well known that a certain amount of agent time will be lost, either in large blocks (unanticipated shift cancellations, partial day absences for personal reasons) or in small blocks (late arrivals to the call center, extra-long breaks, trips to the bathroom).

There is an important distinction between two different kinds of lost agent time. On one hand, agent time that is known to be lost prior to the creation and publication of a schedule has essentially no additional impact on the simulation model beyond the fact that this particular agent is not included in the schedule.

On the other hand, scheduled time that is not worked, either because of unexpected absences or because of lack of rigorous adherence with agent schedules, is time that should be accounted for in the simulation if this represents a known phenomenon (e.g. higher absenteeism on Mondays). In the call center industry, this is known as "shrinkage" and it is a major management problem as well as significant modeling challenge.

Most call centers have significant levels of shrinkage – we have seen many sites with over 30% overall. We have included a shrinkage level of 10% in our example model.

4.7.2 Additional Detail for Outbound Queues

As we have discussed earlier, the workflow associated with Outbound calls is very different than the logic for Inbound queues. At heart, this modeling difference stems from the fact that inbound calls are characterized by a random arrival pattern; in contrast, the outbound dialing pattern can be scheduled but each call features a random outcome (right party connect, wrong party connect, no answer).

In addition, as discussed in Section 3 above, the performance metrics associated with Outbound queues are quite different (overall RPCs achieved, rather than the queue and abandonment statistics that are typically used to evaluate Inbound queues). In order to effectively estimate the number and pattern of RPCs, simulation models require information about the probabilities that a given dial achieves an RPC, which typically varies by time of day, as well as the AHT associated with an RPC.

To model one level deeper, one might consider actually representing the detailed logic of the predictive dialer (see Samuelson 1999 for more on predictive dialer logic). However, this level of detail was not necessary for the types of business decisions being addressed by our example model, and so we have not included detailed dialer logic in it.

5 EXAMPLE: ROUTING STRATEGIES FOR A COLLECTIONS CALL CENTER

5.1 Operational Problem and Business Decisions

Throughout Section 4, we have described parts the simulation model associated with this example. The call center of interest is illustrated in Figure 3, and the formulation was motivated by discussions with several blended inbound-outbound centers about optimal system design.

In our example, the call center is open Monday - Friday from 7:00 am to 6:00 pm. There are 50 Inbound agents (Group #1) and a total of 150 Outbound agents (Group #2 and Group #3). We treat agent schedules for each of the three agent groups, as well as call forecasts (a total of about 20,000 calls for the week) for the Inbound calls, as fixed inputs for this simulation model. In addition, we assume that there is an effectively unlimited pool of customers to contact with Outbound calls.

The operational problem facing the management of this call center is focused on call routing and agent skilling. Underlying this problem is the classical tension between *specialization* and *cost*.

In terms of specialization, Inbound agents are far more effective in handling Inbound calls than Outbound agents, while Inbound calls disrupt the rhythm and effectiveness of Outbound agents; for both of these reasons, it would be far better to have specialized agents for Inbound and Outbound calls respectively.

In terms of cost, there is a management objective of handling 80% of Inbound calls within 60 seconds for each interval of the day. With dedicated agents, this translates into a larger amount of Inbound agents *required* than are actually *available*. Current staffing levels, therefore, will result in longer than desirable waiting times, which in turn is correlated with higher abandonment rates.

Specifically, the business decisions to be addressed are as follows:

- Of the 150 Outbound-skilled agents, how many of them should be enabled to handle Inbound calls and included in the Cross-Trained Outbound group?
- If no Inbound agents are available, how quickly should an Inbound call be offered to the Cross-Trained Outbound group?

In an ideal world, there would be an empirical "right answer" to these questions, a mathematically optimal solution that could be determined through sequential simulation runs.

In practice, however, such decisions typically involve substantial trade-offs that are difficult to value in relation to one another, and simulation's role is to quantify the impact of different possible decisions.

The key output metrics for these simulations are:

1. Phone Service Levels (% of Inbound calls handled within 60 seconds).
2. Abandonment Levels (% of Inbound callers hanging up prior to receiving service).
3. Right Party Connects (total number of Outbound calls completed to the correct individuals).
4. Number of Overflows (of Inbound calls to Cross-Trained Outbound group).

5.2 Numerical Results

5.2.1 Determining the Number of Replications

For each of the individual scenarios that are discussed below, we ran multiple replications of the simulation model and computed estimates for performance measures based on the average of the run length.

For purposes of determining the number of runs for each scenario, we focused on average weekly Service Level for the Inbound queue as the statistic of interest. After each run, we would examine overall standard deviation of this statistic across all runs to date. We continued to run additional iterations until this overall standard deviation was under 2.5%, which we had set arbitrarily as our confidence threshold.

5.2.2 Base Case

Our baseline scenario is one with no Outbound Cross-Trained agents. This base case is listed as Scenario 1 in Table 1 below.

Table 1: Simulation Results For Base Case and Initial Cross-Training Scenarios

Scenario #	Number Cross-Trained	SL %	Abandon %	Number Inter-rupts	Total RPCs for Week
1.	0	56.3	16	0	2621.8
2.	10	70.1	10	1288	2494.4
3.	20	80	6.3	2211	2400.5
4.	30	86.7	4	2816	2339
5.	40	92.1	2.3	3242	2301.7
6.	50	93.5	1.7	3374	2284
7.	75	97.1	0.8	3715	2255.1
8.	100	98	0.6	3726	2250.4

From this base case, it was clear that the Inbound Agent Group alone cannot deliver the desired Service Level (80% within 60 seconds), and that the Abandonment Rate is also much higher than desired.

5.2.3 Varying Cross-Training Levels

We then began to vary the number of Outbound-Skilled Agents who were included in the Cross-Trained Outbound group, assuming for these initial experiments that Inbound calls would immediately overflow to Cross-Trained Outbound agents whenever all Inbound Only agents were busy. The impact of this cross training on the population of Inbound callers is dramatic, as even limited cross training has a big impact on Service Levels and Abandonment Rates. In addition, there is an equally obvious negative impact of this cross training on the Outbound call statistics. These trade-offs are evident in Table 2 below.

Based on these preliminary simulations, we chose to focus on cross-training a total of 30-40 Outbound agents. From here, we turned our attention to defining parameter for how long Inbound calls should wait before being made available to Cross-Trained Outbound agents.

5.2.4 Varying the Wait Time Parameter for Overflowing Inbound Calls

Results for different scenarios associated with 30 and 40 Cross-Trained Outbound agents are shown in Tables 2 and 3.

Table 2: Simulation Results - 30 Cross-Trained Agents

Scenario #	Wait Until Overflow	SL %	Abandon %	Number Interrupts	RPCs for Week
4	0 seconds	86.7	4	2816	2339
9.	15 seconds	85.9	4.2	2592	2360.3
10.	30 seconds	82.6	4.7	2403	2380.6
11.	45 seconds	78.3	5.7	2173	2414.3
12.	60 seconds	67.9	7.6	1769	2453.8

Table 3: Simulation Results - 40 Cross-Trained Agents

Scenario #	Wait Until Overflow	SL %	Abandon %	Number Interrupts	RPCs for Week
5	0 seconds	92.1	2.3	3242	2301.7
13.	15 seconds	90.2	2.7	3057	2324.7
14.	30 seconds	87.7	3.3	2727	2347.2
15.	45 seconds	81.9	4.3	2428	2377.9
16.	60 seconds	70.1	6.3	1998	2427.2

5.2.5 Summary

The different scenarios that we have simulated have enabled us to (a) hone in on the right levels of cross training to meet the Service Level goals with the current staffing levels and (b) examine trade-offs between different scenarios in terms of the key model outputs.

For example, consider Scenarios 3, 10, 14, and 15, all of which deliver SLs at or above the 80% target. The answer to which of these is the “best” choice will of course depend on the relative value of RPCs, Service Levels, and Abandoned customers. However, it is interesting to note that Scenario 3 produces essentially the same SL and RPC values as Scenarios 10 and 15 – but with a substantially higher abandonment rate. In turn, the tangible difference between Scenario 14 and 15 enables managers to explicitly quantify the level of increased Service Level and decreased abandonment rates against the decreased number of RPCs.

Finally, it is worth mentioning that while we have shown summary statistics for sixteen scenarios here, it is relatively easy for us to produce more detailed statistics and also to vary different parameters to examine any number of other cases. This flexibility, in turn, enables managers and analysts to develop a sense for system dynamics and also to proactively answer common senior management questions such as "what would a 10% increase in call volume next week do to us?" or "what is the value of adding an outsourcer to help us during our peak months?"

6 WHAT THE FUTURE HOLDS FOR CALL CENTER SIMULATION

Looking out into the future, we see two major trends impacting call center simulation. First of all, operational complexity will continue to grow: more queues, more different agent schedules, more diverse skilling combinations and routing rules. This will put pressure on analysts to not only build richer models, but also to define output metrics that enable them – and their management – to understand the bigger picture as well as the more minute statistics. Even in the very simple numerical example above, it is easy to see how one can become overwhelmed with the sheer volume of numbers that simulation can produce.

In addition, as executives begin to understand that the call center is a key component in their customer value delivery chain, we foresee an increased desire to understand the risks inherent in any particular operational configuration. In particular, we see interesting and important opportunities in randomizing not only call arrival patterns and handling times but also overall call volumes, and using techniques from risk analysis and experimental design along with simulation models to quantify system capacity and delivery risks.

Finally, we hope for and expect improvements in the quality of data provided for quantitative analysis. In particular, increased accuracy and detail associated with handle time distributions, waiting time distributions, and abandonment time distributions will lead to better model inputs and more robust results.

ACKNOWLEDGMENTS

The authors would like to thank the call center directors, managers, and executives that we have had the chance to work with over the past several years. Through our professional and personal interactions with these overworked and underappreciated individuals, we have learned a great deal about call center operations, management, and data sources, all of which has contributed greatly to our ability to model and analyze these types of systems. We have also gotten a first-hand sense of the pressures that these individuals work under, and hope that our experience and our models can help provide insight and support to them.

REFERENCES

Andrews, B. H. and S. M. Cunningham. 1995. L.L. Bean Improves Call Center Forecasting. *Interfaces* 25:1-13.

Andrews, B. H. and H. L. Parsons. 1989. L.L. Bean Chooses an Agent Scheduling System. *Interfaces* 19:1 – 9.

Andrews, B. H. and H. L. Parsons. 1993. Establishing Telephone-Agent Staffing Levels Through Economic Optimization. *Interfaces* 23:14-20.

Feinberg, R. A., I. Kim, B. Hokama, K. Ruyter, and C. Keen. Operational Determinants of Caller Satisfaction in the Call Center. *International Journal of Service Industry Management* 11:131-141.

Garnett, O., A. Mandelbaum, and M. L. Reimann. 2002. Designing a Call Center With Impatient Customers. *Manufacturing and Service Operations Management* 4:208-227.

Grossman, T. A., D. A. Samuelson, S. L. Oh, and T. R. Rohleder. 2001. Call Centers. In *Encyclopedia of Operations Research*, ed. S. L. Gass and T. M. Harris, 73-76. Norwell: Kluwer Academic Publishers.

Hoffman, K. L. and C. M. Harris. 1986. Estimation of a Caller Retrial Rate for a Telephone Information System. *European Journal of Operational Research* 27:207-214.

Mabert, V. A. 1985. Short-Interval Forecasting of Emergency (911) Workloads. *Journal of Operations Management* 5:259-271.

Mandelbaum, A. 2001. *Call Center Research Bibliography with Abstracts*, Technical Report, Technion, Israel Institute of Technology.

Mandelbaum, A. and N. Shimkin. 2000. A Model for Rational Abandonments from Invisible Queues. *Queueing Systems, Theory, and Application* 36:141-173.

Mehrotra, V. 1997. Ringing Up Big Business. *OR/MS Today* 24:18-24.

Pinker, E. and R. Shumsky. 2000. The Efficiency-Quality Tradeoff of Crosstrained Workers. *Manufacturing and Service Operations Management* 2:32-48.

Saltzman, R. and V. Mehrotra. 2001. A call center uses simulation to drive strategic change. *Interfaces* 31:87-101.

Samuelson, D. A. 1999. Predictive Dialing For Outbound Telephone Call Centers. *Interfaces* 29:66-81.

AUTHOR BIOGRAPHIES

VIJAY MEHROTRA is an Assistant Professor in the College of Business at San Francisco State University.

Vijay joined the SFSU faculty in the fall of 2003 after over ten years in the operations management consulting field. Most recently, he was a Vice President with the Solutions Group at Blue Pumpkin Software. Prior to joining Blue Pumpkin, he was co-founder and CEO of Onward Inc., an operations management consulting firm based in Mountain View, CA that focuses on the successful application of Operations Research techniques to business applications.

Vijay's research interests include applications of stochastic processes and optimization, queueing networks, and the adoption of models and information technology by individuals and organizations. Within the simulation area, he has been actively involved with modeling semiconductor manufacturing facilities, electric power production systems, container ship traffic, as well as call center operations.

Over the course of his consulting career, Vijay has worked with a wide variety of clients in many industries, including EDS, Intuit, Hewlett Packard, Charles Schwab, AOL, Hewlett Packard, IBM, General Electric, Sykes, CIGNA, Tyecin Systems, National Semiconductor, and Remedy.

Vijay holds a B.A. degree in Mathematics, Economics, and Statistics from St. Olaf College and an M.S. and Ph.D. in Operations Research from Stanford University. He is the past President of the Northern California INFORMS chapter, and writes a regular column in *OR/MS Today* entitled "Was It Something I Said?" His email address is <Vmehrotra@sfsu.edu>.

JASON FAMA is an analyst and developer in the algorithms group at Blue Pumpkin Software. While at Blue Pumpkin, Jason has been actively involved in the development of forecasting, queueing, and scheduling algorithms, and with effectively embedding efficient simulation models within Blue Pumpkin's overall workforce optimization framework.

Prior to joining Blue Pumpkin, had worked as a researcher Rockwell's Palo Alto Laboratory. Jason holds a B.S. degree in Economics and Computer Science from the University of California at Berkeley. His email address is <jfama@bluepumpkin.com>.