Call Blending

Y.H. Hoekstra

May 2, 2002

Preface

This paper is a part of the study "Business Mathematics and Informatics" at the Free University in Amsterdam. The paper is a compulsory subject in the fourth year of the study. The idea is that the student makes a little survey of a particular subject on his own. The subject is chosen by the student himself.

The subject of this paper came to my mind because a girlfriend of mine works in a call center. She told me some stories about her call center. Once she told me that she has time to make her homework at her work. I became interested in how that was possible. Soon it became clear to me that she has to wait until a call arrives. When no calls arrive at all or very little calls arrive she could spent her time on her homework. At that particular moment it came to my mind that, from a managerial perspective, it could be better to let her also make phone calls to people. In other words she could combine so-called inbound and outbound calls, which is called *call blending*.

Figure 1: The working environment of a call center

Another reason to choose for this subject was the course "Modeling of business processes". One of the subjects of this course is call centers. During this course I recognized my problem in the subject of so-called call blending. It appeared that a lot work can be done on this subject. An idea for this paper was born.

Some people were of great help. I would like to thank them here. In the first place I would like to thank Sandjai Bhulai for the discussions we had and the help he offered in writing this paper. I would like to thank Rob van der Mei and Bart Gijsen, my expert supervisors of the graduation project I did at KPN research during my study. They gave me some time to work on this paper.

Yanko Hoekstra.
Amsterdam, April 2002

Contents

Notation

- λ = number of arrivals per time unit
- μ_1 = inverse of the mean service time for inbound calls
- μ_2 = inverse of the mean service time for outbound calls
- s = number of available agents
- x = number of inbound calls in the system
- y = number of outbound calls in service
- c = threshold, i.e., $x + y \geq c$ should be satisfied
- $\rho = \frac{\lambda}{\mu_1}$ = load without type 2 jobs
- δ = probability to do nothing when finished serving a job at level $x + y = c + 1$

Chapter 1

Introduction

This paper describes a certain aspect of call centers called call blending. There are two kind of calls in a call center. The first type of calls are phone calls that arrive at the call center, referred to as **inbound calls**. The call center itself also contacts other people by, e.g., e-mail, faxes, and phone calls. These calls are referred to as **outbound calls**.

In most call centers the **agents**, the employees who make inbound and outbound calls in call centers, exclusively handle either inbound or outbound calls. In this system it could be the case that agents working on inbound calls have no work to do because no phone calls arrive. This approach seems quite expensive. Another problem with the inbound calls is the fact that the arrivals per time unit fluctuates over the day. This makes it hard to schedule your agents because you need a different number of agents for, i.e., each half hour. A simple solution to both problems is to mix the inbound and outbound calls. This is what **call blending** is about. An agent working on inbound calls and outbound calls seems more effective than an agent only handling inbound calls. Another advantage is that it makes the scheduling of the agents easier; it is possible to use the same amount of agents the whole day while you increase the average productivity of the agents. Thus call blending offers a way to increase resource utilization in a call center.

There is also a disadvantage. An inbound call has to wait longer if an inbound call arrives and the agent is handling an outbound call it cannot interrupt. In other words, say we have s agents available and we apply call blending, in this case queueing times will be larger than the case that we also have s agents en do not apply call blending. But if you combine your agents available for inbound and your agents available for outbound it may become possible to get an advantage in productivity and lower queueing times due to call blending.

In this paper we assume a certain policy to be optimal for call blending and we will derive expressions for the mean queueing time of inbound calls and derive the mean throughput for the outbound calls. We also derive the productivity of the agents. These expressions will be implemented in a web-based tool which can be used by call center managers. With a simulation tool we will show that the expressions indeed lead to the expected waiting time. We also show that advantages can be accomplished by call blending.

- **Chapter 2** gives an overview of call blending in business life;
- **Chapter 3** describes the call blending model in more detail;
- **Chapter 4** derives the waiting time and throughput under call blending;
- **Chapter 5** discusses the results of Chapter 4 compared with simulation and shows some examples of performance gains that can be achieved;
- **Chapter 6** shows the web based CALLculator. This chapter can be read without reading the other chapters. This is a convenient tool for call center managers;
- **Chapter 7** presents conclusions and topics for further research.

Chapter 2 represents the business part of this paper. Chapters 3 and 4 represent the mathematical part, and Chapters 5 and 6 represent the informatics part of the paper.

Chapter 2

Call blending in practice

Call centers in general have two distinct groups to handle their calls. One group for the inbound calls and one group for the outbound calls. During my visit to the call center “tesselaar”, in Haarlem, I mentioned that it seemed strange to me that they did not apply call blending. They said to me that the techniques developed for call blending were not good. The only call blending they apply, and according to them all the call centers in the Netherlands, is the blending of various inbound calls.

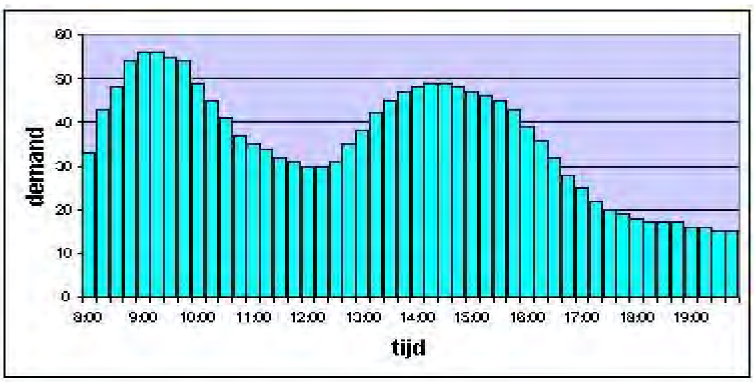

Figure 2.1: Typical arrival structure over the day

The research on the internet showed some packages that include packages for call blending. For example CISCO [2] is one of the manufacturers that delivers a call blending option. The call blending technique they use is as follows. Everybody works on outbound calls and the inbound calls are assigned to the first available agent. This technique assures you that the productivity of the agents is maximal. But the probability that an inbound call has to wait is 100% assuming that you always have outbound calls available. Another disadvantage is that the packages do not include estimates of the expected waiting times. This is very important because call centers agree on a certain service level agreement, for example 80% of the calls has to be answered within 20 seconds.

In call centers the arrival of inbound calls varies over the day, e.g. see Figure 2.1. One of the problems in call centers is the scheduling of the agents for the varying arrival rates. Call blending can give a solution in this case. The scheduling of the agents becomes easier because you can assign the agents dynamicly. So if you notice that certain agents are not needed on the inbound calls, you can assign them to the outbound calls. This will increase the productivity while you still meet the service demands if you apply a good strategy. The maximization of the productivity is essential in call centers since the salary costs of agents account for 60-70% [3] of the total operating costs. So the main objective in call blending is to maximize the productivity while you still meet your demands.

This paper will give a strategy that can be used in call blending. The simulation chapter will show some results for this strategy and shows that it becomes possible to increase productivity and throughput while you use the same number of agents if you do not apply call blending.

Chapter 3

Model description

The exact model formulation is as follows. There are two types of calls, inbound calls and outbound calls. The service times of these calls are independently exponentially distributed with parameter

$$\mu_1 = \frac{1}{\text{mean inbound service time}} \tag{3.1}$$

for the inbound calls and parameter

$$\mu_2 = \frac{1}{\text{mean outbound service time}} \tag{3.2}$$

for the outbound calls. The inbound calls are assumed to arrive according to a Poisson process with rate

$$\lambda = \frac{1}{\text{mean time between arrivals}}. \tag{3.3}$$

The call center is assumed to have an infinite number of phone lines available for inbound calls that cannot be served yet. We also assume that there is an infinite number of outbound calls available. The number of agents available is represented by s. The long-term average waiting time of inbound calls should be below a constant α. Waiting excludes the service time. The objective for the outbound calls is to serve as many jobs as possible while the time constraint for the inbound calls is still met.

The following control actions are possible. The moment an agent finishes service, or, more generally at any moment that an agent is available, he can take one of the following three actions:

- Start serving an inbound call, if one or more phone calls are waiting;
- Start serving an outbound call;
- Remain idle.

We describe the threshold policy as follows. Let y be the number of outbound calls in service and let x be the number of inbound calls in the system. There is a level c, called the threshold, such that if $x + y < c$, then the proposed action is to schedule $c - x$ calls, if inbound calls are available they have a higher priority since we want to minimize the long term average waiting time for the inbound calls. If you served an outbound call first then the throughput for that moment is higher but we are interested in the long term throughput which stays the same. The average waiting time for the inbound calls however increases so if an inbound call is available you should schedule this first. If $x + y \geq c$, then the proposed action is to schedule no outbound calls but only inbound calls. This policy is optimal if μ_1 is equal to μ_2 and nearly optimal if this is not the case (see Bhulai and Koole [1]).

Assume that a threshold policy with threshold c gives an expected waiting time of p and that the threshold policy with threshold $c + 1$ gives an expected waiting time of r. However you need a policy with an expected waiting time of q with $p < q < r$. To achieve this we introduce randomization between level c and level $c + 1$. The randomization works as follows. When finished serving an inbound or outbound call at level $c + 1$ and no inbound calls available there are two possibilities. The first possibility is to schedule an outbound call with probability $1 - \delta$, in that case $c + 1$ agents are serving a call. The second possibility is to do nothing with probability δ, in that case only c agents are serving a call. Note that when $\delta = 1$ we have a normal threshold policy without randomization. The parameter δ must meet the following condition:

$$0 < \delta \leq 1. \tag{3.4}$$

Figure 3.1 shows the model as a two dimensional Markov Chain. The x axis represents the number of inbound calls in the system. The y axis represents the number of outbound calls served at the moment. Now we

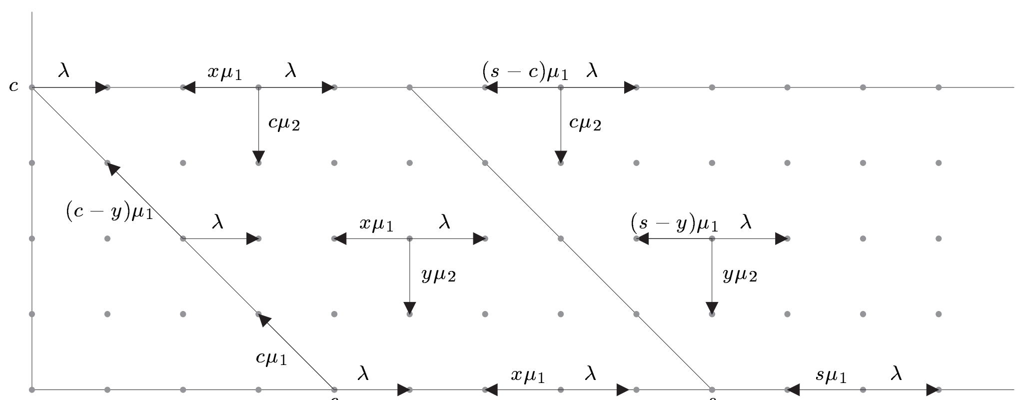

Figure 3.1: System dynamics.

define q_x^y as the stationary probability that the system is in state (x, y).

To end this chapter we give an overview of the assumptions presented to you in this chapter. **Model assumptions summarized**

- Inbound calls arrive according to a Poisson process with parameter λ;
- The duration of inbound calls is exponentially distributed with parameter μ_1;
- The duration of outbound calls is exponentially distributed with parameter μ_2;
- Customers do not abandon;
- An infinite number of outbound calls is present;
- An infinite number of phone lines is available;

Chapter 4

Solving the model

4.1 Stationary probabilities for $\delta = 1$ and $c = s$

We start with $y = c = s$ and $x + y \geq s$.
Note that there are no inbound calls in service at this level. The following equilibrium equation holds

$$(\lambda + s\mu_2)q_{x,s} = \lambda q_{x-1,s}. \tag{4.1}$$

The solution for this level will be:

$$q_{x,s} = \left(\frac{\lambda}{\lambda + s\mu_2}\right) q_{x-1,s} = \left(\frac{\lambda}{\lambda + s\mu_2}\right)^x q_{0,s} \qquad \text{for } x \geq s - y, y = s. \tag{4.2}$$

For convenience we define in the rest of this section

$$\gamma_s := \frac{\lambda}{\lambda + s\mu_2}, \tag{4.3}$$

and

$$K_{s,0} := q_{0,s}. \tag{4.4}$$

So the solution can be written as follows:

$$q_{x,s} = K_{s,0}\gamma_s^x \tag{4.5}$$

Now examine level $y = s - 1$ and $x + y \geq s$
The equilibrium equations are as follows.

$$\Big(\mu_1 + \lambda + (s-1)\mu_2\Big) q_{x,s-1} = \lambda q_{x-1,s-1} + \mu_1 q_{x+1,s-1} + s\mu_2 q_{x,s}. \tag{4.6}$$

We start with solving the homogeneous equation with respect to level $s-1$. Suppose that $q_{x,s-1} = \gamma_{s-1}^{x-1} q_{1,s-1}$ with γ_{s-1} to be determined. Then the homogeneous equation looks as follows.

$$(\mu_1 + \lambda + (s-1)\mu_2)\gamma_{s-1}^{x-1} q_{1,s-1} = \lambda\gamma_{s-1}^{x-2} q_{1,s-1} + \mu_1\gamma_{s-1}^{x} q_{1,s-1}, \tag{4.7}$$

thus

$$q_{1,s-1}\gamma_{s-1}^{x-2}[\lambda-(\mu_1+\lambda+(s-1)\mu_2)\gamma_{s-1}+\mu_1\gamma_{s-1}^2]=0. \tag{4.8}$$

Since $\gamma_{s-1}>0$, due to the fact that the states we investigate in the previous equation can and will be reached in finite time, we have to make sure that the equation between brackets equals zero. This can be achieved by using the abc-formula. One of the solutions is greater than 1 and therefore not useful. The result for γ_{s-1} is:

$$\gamma_{s-1}=\frac{(\mu_1+\lambda+(s-1)\mu_2)-\sqrt{(\mu_1+\lambda+(s-1)\mu_2)^2-4\mu_1\lambda}}{2\mu_1}. \tag{4.9}$$

Now we have the homogeneous solution:

$$q_{x,s-1}=\gamma_{s-1}^{x-1}q_{1,s-1}, \qquad \text{for } x\geq 1, y=s-1. \tag{4.10}$$

For the general solution we need a particular solution to the inhomogeneous equation. As particular solution we try

$$q_{x,y}=K_{s-1,1}\gamma_s^x. \tag{4.11}$$

With $K_{s-1,1}$ the second constant for level $y=s-1$ ($K_{s-1,0}$ is the first constant for this level used in the general solution for the homogeneous part). After substitution in the inhomogeneous equation we get the next equation:

$$K_{s-1,1}\gamma_s^{x-1}[\lambda-(\mu_1+\lambda+(s-1)\mu_2)\gamma_s+\mu_1\gamma_s^2]+s\mu_2K_{s,0}\gamma_s^x. \tag{4.12}$$

Now subtract the following equation:

$$K_{s-1,1}\gamma_s^{x-1}[\lambda-(\lambda+s\mu_2)\gamma_s]. \tag{4.13}$$

Note that this term equals zero, because the term between brackets is zero due to the homogeneous equation at level s. After subtraction we derive the following equation:

$$K_{s-1,1}\gamma_s^{x-1}[-((\mu_1-\mu_2)\gamma_s+\mu_1\gamma_s^2]+s\mu_2K_{s,0}\gamma_s^x, \tag{4.14}$$

or

$$\gamma_s^x[K_{s-1,1}\cdot\Big(-(\mu_1-\mu_2)+\mu_1\gamma_s\Big)+s\mu_2K_{s,0}], \tag{4.15}$$

which is equal to zero when

$$K_{s-1,1}=\frac{s\mu_2}{(\mu_1-\mu_2)-\mu_1\gamma_s}K_{s,0}. \tag{4.16}$$

For the general solution we have to multiply the homogeneous solution with a constant C. We define

$$K_{s-1,0}:=C\cdot q_{1,s-1}. \tag{4.17}$$

The general solution for level $y = s - 1$ is now as follows

$$q_{x,s-1} = K_{s-1,0}\gamma_{s-1}^{x-1} + K_{s-1,1}\gamma_s^x. \tag{4.18}$$

Now examine level $0 < y < s$ and $x + y \geq s$.
The equilibrium equations are as follows.

$$\Big((s-y)\mu_1+\lambda+y\mu_2\Big)q_{x,y} = \lambda q_{x-1,y}+(s-y)\mu_1 q_{x+1,y}+(y+1)\mu_2 q_{x,y+1} \tag{4.19}$$

We start with solving the homogeneous equation with respect to level y. Suppose that $q_{x,y} = \gamma_y^{x-(s-y)}q_{s-y,y}$ with γ_y to be determined. Then the homogeneous equation looks as follows.

$$q_{s-y,y}\Big((s-y)\mu_1+\lambda+y\mu_2\Big)\gamma_y^{x-(s-y)} = q_{s-y,y}[\lambda\gamma_y^{x-1-(s-y)}+(s-y)\mu_1\gamma_y^{x+1-(s-y)}], \tag{4.20}$$

thus

$$q_{s-y,y}\gamma_y^{x-1-(s-y)}[\lambda - \Big((s-y)\mu_1 + \lambda + y\mu_2\Big)\gamma_y + (s-y)\mu_1\gamma_y^2] = 0. \tag{4.21}$$

Since $\gamma_y > 0$ due to the fact that each state we investigate in the previous equation can and will be reached in finite time, we have to make sure that the equation between brackets equals zero. This can be achieved by using the abc-formula. One of the solutions is greater than 1 and therefore not useful. The result for γ_y is:

$$\gamma_y = \frac{\Big((s-y)\mu_1 + \lambda + y\mu_2\Big) - \sqrt{\Big((s-y)\mu_1 + \lambda + y\mu_2\Big)^2 - 4(s-y)\mu_1\lambda}}{2(s-y)\mu_1}. \tag{4.22}$$

Now we have the homogeneous solution:

$$q_{x,y} = \gamma_y^{x-(s-y)}q_{s-y,y}, \qquad \text{for } x \geq s-y, 0 < y < s. \tag{4.23}$$

For the general solution we need a particular solution to the inhomogeneous equation. As particular solution we try

$$q_{x,y} = \sum_{i=y+1}^{s} K_{y,i-y}\cdot\gamma_i^{x-(s-i)}. \tag{4.24}$$

After substitution of this particular solution we see that we get a combination of linear equations in γ_i which can be solved separately. We solve for $i = y+1, \ldots, s$ separately. So we solve $s - y$ equations. Now let's introduce $j = i - y$. So values for j are $j = 1, \ldots, s - y$.

After substitution and focusing on γ_{y+j} in the inhomogeneous equation we get the next equation.

$$K_{y,j}\gamma_{y+j}^{x-(s-y-j)-1}[\lambda-\Big((s-y)\mu_1+\lambda+y\mu_2\Big)\gamma_{y+j}+(s-y)\mu_1\gamma_{y+j}^2]+(y+1)\mu_2K_{y+1,j-1}\gamma_{y+j}^{x-(s-y-j)}. \quad (4.25)$$

Now subtract the following equation:

$$K_{y,j}\gamma_{y+j}^{x-(s-y-j)-1}[\lambda-\Big((s-y-j)\mu_1+\lambda+(y+j)\mu_2\Big)\gamma_{y+j}+(s-y-j)\mu_1\gamma_{y+j}^2]. \quad (4.26)$$

Note that this term equals zero because the term between brackets is zero due to the homogeneous equation at level y. Note that even at level s this equation will equal zero since $s-y=0$. After subtraction we derive the following equation:

$$K_{y,j}\gamma_{y+j}^{x-(s-y-j)-1}[-j(\mu_1-\mu_2)\gamma_{y+j}+j\mu_2\gamma_{y+j}^2]+(y+1)\mu_2K_{y+1,j-1}\gamma_{y+j}^{x-(s-y-j)} \quad (4.27)$$

which equals

$$\gamma_{y+j}^{x-(s-y-j)}[-\Big(j(\mu_1-\mu_2)-j\mu_1\gamma_{y+j}\Big)K_{y,j}+(y+1)\mu_2K_{y+1,j-1}], \quad (4.28)$$

which is equal to zero when

$$K_{y,j}=\frac{(y+1)\mu_2}{j(\mu_1-\mu_2)-j\mu_1\gamma_{y+j}}K_{y+1,j-1}. \quad (4.29)$$

If we add the solutions of the linear equations for each γ_i we get a particular solution. For the general solution we need to add the homogeneous solution multiplied with a constant C_y and the particular solution. We define:

$$K_{y,0}:=C_y\cdot q_{s-y,y}. \quad (4.30)$$

The general solution for level $0<y\le s$ is now as follows

$$q_{x,y}=\sum_{i=y}^{s}K_{y,i-y}\cdot\gamma_i^{x-(s-i)} \quad (4.31)$$

Now examine level $y=0$ and $x+y\ge s$.
If you take a closer look to the equations that hold for $0<y<s$ and $x+y\ge s$, you should notice that the solution also holds for $y=0$. Up till now we have only $s+1$ unknown constants, which can be solved with the remaining equations for $x+y=s$.
Now examine level $y=s$ and $x+y=s$.
The equilibrium equation for this case is as follows:

$$\lambda q_{0,s}=\mu_1q_{1,s-1}, \quad (4.32)$$

so

$$q_{1,s-1} = \frac{\lambda q_{0,s}}{\mu_1}. \tag{4.33}$$

Now examine level $0 < y < s$ and $x + y = s$.
The equilibrium equations are as follows:

$$(\lambda+(s-y)\mu_1)q_{s-y,y} = (s-y)\mu_1 q_{s-y+1,y}+(s-y+1)\mu_1 q_{s-y+1,y-1}+(y+1)\mu_2 q_{s-y,y+1}, \tag{4.34}$$

so

$$q_{s-y+1,y-1} = \frac{(\lambda + (s-y)\mu_1)q_{s-y,y} - (y+1)\mu_2 q_{s-y,y+1} - (s-y)\mu_1 q_{s-y+1,y}}{(s-y+1)\mu_1}. \tag{4.35}$$

Now it is possible for each stationary probability to express it in terms of one constant. We also know that

$$\sum_{x=0}^{\infty}\sum_{y=0}^{s} q_{x,y} = 1. \tag{4.36}$$

This determines the last constant. Now all $q_{x,y}$ are determined for the particular case that $c = s$ and $\delta = 1$. In the CALLculator described in a later chapter we determine the $q_{x,y}$ up to $x + y = n$. Where n is a figure for the accuracy of the calculation.

4.2 Stationary probabilities for $\delta = 1$ and $c < s$

Start with $y = c$ and $x + y \geq s$
For $y = c$ and $x + y \geq s$ the following equilibrium equation holds

$$((s-c)\mu_1 + \lambda + c\mu_2)q_{x,c} = \lambda q_{x-1,c} + (s-c)\mu_1 q_{x+1,c}. \tag{4.37}$$

Suppose $q_{x,c} = \gamma_c^x q_{s-c,c}$, then

$$((s-c)\mu_1 + \lambda + c\mu_2)\gamma_c^x q_{s-c,c} = \lambda\gamma_c^{x-1} q_{s-c,c} + (s-c)\mu_1\gamma_c^{x+1} q_{s-c,c}, \tag{4.38}$$

or

$$q_{s-c,c}\gamma_c^{x-1}[\lambda - ((s-c)\mu_1 + \lambda + c\mu_2)\gamma_c + (s-c)\mu_1\gamma_c^2] = 0. \tag{4.39}$$

Since $\gamma_c > 0$ due to the fact that each state we investigate in the previous equation can and will be reached in finite time, we have to make sure that the equation between brackets equals zero. This can be achieved by using the abc-formula. One of the terms is not useful since it is greater than 1. The result for γ_c is:

$$\gamma_c = \frac{((s-c)\mu_1 + \lambda + c\mu_2) - \sqrt{((s-c)\mu_1 + \lambda + c\mu_2)^2 - 4(s-c)\mu_1\lambda}}{2(s-c)\mu_1}. \tag{4.40}$$

We define

$$K_{c,0} := q_{s-c,c}. \tag{4.41}$$

Since the equation only holds for $x + y \geq s$ the solution will be as follows:

$$q_{x,c} = K_{c,0}\gamma_c^{x-(s-c)} \qquad \text{for } x \geq s-c, y = c. \tag{4.42}$$

Continue with $0 \leq y < c$ **and** $x + y \geq s$.
A closer look at these levels tell us that the equibrilium equations are equal to the case that $c = s$. Since the solution for level $y = c$ for $c < s$ equals the solution for level $y = c$ and $c = s$, we have the same solution as in the previous section. Only γ_c differs from the previous section. So the solution for $c < s$ and $x + y \geq s$ for the stationary probabilities is as follows.

$$q_{x,y} = \sum_{i=y}^{c} K_{y,i-y} \cdot \gamma_i^{x-(s-i)} \tag{4.43}$$

With $K_{y,i-y}$ recursively defined as

$$K_{y,j} = \frac{(y+1)\mu_2}{j(\mu_1 - \mu_2) - j\mu_1\gamma_{y+j}} K_{y+1,j-1}, \tag{4.44}$$

and

$$\gamma_y = \frac{\Big((s-y)\mu_1 + \lambda + y\mu_2\Big) - \sqrt{\Big((s-y)\mu_1 + \lambda + y\mu_2\Big)^2 - 4(s-y)\mu_1\lambda}}{2(s-y)\mu_1}. \tag{4.45}$$

Continue with $x + y < s$.
Up till now only $c+1$ constants are left to be determined. This can be done with the remaining finite number of equations. The idea is as follows. Take the equilibrium equations for level $x+y = s$. In each of these equations there is just one $q_{x,y}$, the $q_{x,y}$ for which $x + y = s - 1$, which is not determined in the $c+1$ unknown constants. So now it becomes possible to determine these stationary probabilities in terms of the unknown constants. Now we take a closer look at the equilibrium equations for level $x + y = s - 1$ and solve the stationary probabilities for $x + y = s - 2$. Continue in this way until all stationary probabilities are determined in terms of the unknown constants. For the determination of the stationary probabilities with $x + y = s - 1$ you can use the following equation:

$$q_{x,y} = \frac{\Big(\lambda + y\mu_2 + (x+1)\mu_1\Big)q_{x+1,y} - (y+1)\mu_2 q_{x+1,y+1} - (x+1)\mu_1 q_{x+2,y}}{\lambda}. \tag{4.46}$$

For the determination of the stationary probabilities with $x+y < s-1$ you can use the following equation:

$$q_{x,y} = \frac{\Big(\lambda + y\mu_2 + (x+1)\mu_1\Big) q_{x+1,y} - (y+1)\mu_2 q_{x+1,y+1} - (x+2)\mu_1 q_{x+2,y}}{\lambda}, \quad (4.47)$$

with,

$$q_{x,y} = 0 \qquad \text{for } y > c, \quad (4.48)$$

as expected since you never exceed c inbound calls.
Now all stationary probabilities are determined in the $c+1$ constants. From here on we can use the equilibrium equations for $x+y=c$ to leave only one constant to determine. First we determine $q_{1,c-1}$ in terms of $q_{0,c}$ with

$$q_{1,c-1} = \frac{\lambda q_{0,c} - \mu_1 q_{1,c}}{\mu_1}. \quad (4.49)$$

Now the other $q_{x,c-x}$ are determined in terms of $q_{0,c}$ with

$$q_{x+1,y-1} = \frac{(\lambda + x\mu_1) q_{x,y} - (y+1)\mu_2 q_{x,y+1} - (x+1)\mu_1 q_{x+1,y}}{(x+1)\mu_1}. \quad (4.50)$$

Now all probabilities are expressed in one constant. We also know that

$$\sum_{x=0}^{\infty} \sum_{y=0}^{s} q_{x,y} = 1. \quad (4.51)$$

Now all $q_{x,y}$ are determined for the particular case that $c < s$ and $\delta = 1$. In the CALLculator described in a later chapter we determine the $q_{x,y}$ up to $x+y=n$. Where n is a figure for the accuracy of the calculation.

4.3 Stationary probabilities for $\delta < 1$

This case is not much distinct from the previous problems. Now there exists an extra level $c+1$. If we recalculate the previous sections with this new level, it becomes clear that all equilibrium equations are the same except for the levels $x+y=c$ and $x+y=c+1$.

It is possible for all stationary probability to express it in $c+2$ constants. Now it becomes possible to solve these constants in a similar manner as in the case of $\delta = 1$. The only difference is the fact that we get $c+2$ equations in $c+2$ constants. The equations can be solved by a solver program.

4.4 Determination of throughput, expected waiting time, and productivity

Since all stationary probabilities are known it becomes possible to determine e.g., the throughput of outbound calls, the expected waiting time, the probability to wait for an agent, and the variance of the waiting time. In our case we focus on the throughput of the outbound calls, the productivity of the agents, and the expected waiting time.

The throughput ξ of the outbound calls is given by:

$$\xi = \mu_2 \sum_{x,y} y q_{x,y}. \tag{4.52}$$

The productivity of the agents is defined between 0 and 1 as follows:

$$\omega = 1 - \frac{1}{s} \sum_{y=0}^{s} \sum_{x=0}^{s-y} \Big(s - (x+y) \Big) q_{x,y} \tag{4.53}$$

Let $[x]^+ = \max\{x, 0\}$ and define $W(x,y)$ recursively by:

$$W(x,y) = \frac{1}{(s-y)\mu_1 + y\mu_2} [1 + (s-y)\mu_1 W(x-1,y) + y\mu_2 W(x,[y-1]^+)] \tag{4.54}$$

for $x + y \geq s$, and 0 otherwise. The expected waiting time for the inbound calls is given by

$$\mathbb{E}W_q = \sum_{x,y} W(x,y) q_{x,y}. \tag{4.55}$$

Chapter 5

Simulation

We developed a simulation program for two purposes. The first purpose is to show that the equations are correct, the second purpose is to show some advantages of call blending. We could have used the CALLculator (discussed in the next chapter) for the second purpose, but the CALLculator does not implement randomized policies yet. The simulation program on the other hand is able to use the randomized policies between a level c and a level $c+1$.

Assume that we have the following parameters:

- $\lambda = 0.5$;
- $\mu_1 = 0.4$;
- $\mu_2 = 0.2$;
- $s = 5$.

Note that the service time for the outbound calls is twice as large as the service time for the inbound calls.

Assume that we do not apply call blending. In that case we have 6 possibilities to divide the group. We can use the Erlang delay formula to calculate the expected waiting time in each of these cases. The throughput can be determined by multiplying the number of agents serving outbound calls with μ_2. Then for each of the cases the following results hold.

Inbound agents	Outbound agents	Expected Waiting time	productivity	throughput
0	5	INF	1	1
1	4	INF	1	0.8
2	3	1.602	0.85	0.6
3	2	0.222	0.65	0.4
4	1	0.038	0.45	0.2
5	0	0.006	0.25	0

Now assume that we do apply call blending for this case. The results for the throughput, the waiting time and the productivity can be seen in Figure 5.1. This figures show the result of the simulation and the result of the calculations done in the previous chapter (the CALLculator, described in the next chapter, is used for the calculations). The productivity for the inbound calls equals the productivity minus the throughput for the outbound calls. Now note that this productivity remains the same regardless of the policy. This was to be expected because the inbound works that arrives has to be done in each case.

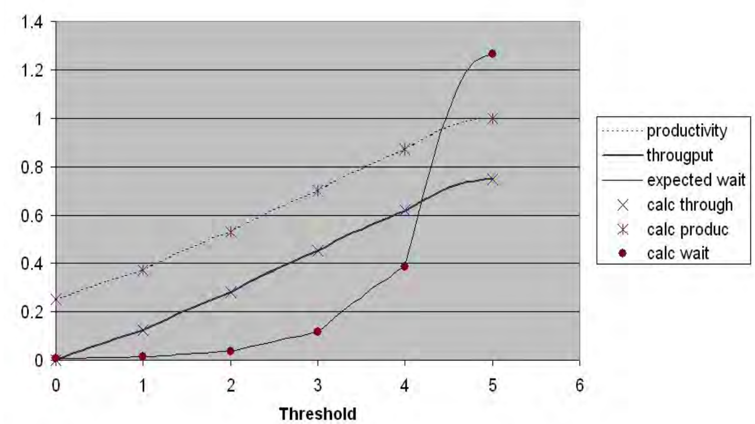

Figure 5.1: Simulation and calculated results

We can now start comparing the case with no call blending and with call blending. Note that the calculated values lie on the curves of the simulation, this confirms the correctness of the formulas derived for call blending. Assume that we did the calculation taking minutes as time unit. If we do not apply call blending and assign only zero or one agent to the inbound calls, then the waiting time becomes very large (infinite). If we assign two agents to the inbound calls we still have to wait more than 1 minute and thirty seconds. If we assign three agents to the inbound calls we get an acceptable waiting time of about 13 seconds. In this case the throughput equals 0.4. Now we consider the call blending case. If we use the threshold of 3 we get an average waiting time of 7 seconds and a throughput of 0.45. So in the call blending case the waiting time is smaller and the throughput is bigger. By applying the randomization technique you could set the throughput to 0.40 as in the case with no call blending but with lower waiting times. It is also possible to maximize the throughput while the expected waiting time equals the waiting time in the case with no call blending. This will result in a higher throughput.

We showed only one case in our paper. But the result holds in much more cases. Only if the outbound calls last much longer then you do not have advantage in both waiting time and throughput but just in one of them. Note however that the outbound calls in this case last twice as long as the inbound calls. So better results can be obtained.

Chapter 6

The web-based CALLculator

We developed a calculator to get a quick answer on the throughput of outbound calls and the waiting time for inbound calls with user specified values for the various parameters. This chapter gives an overview of the CALLculator and explains how to use the calculator. In Figure 6.1 you see a screen-shot of the CALLculator.

On fore-hand you have to determine which time unit you will use, e.g., seconds, minutes, or hours. Once you have chosen this time unit stick to it with every variable else the garbage in garbage out principle holds. The first step is to estimate the number of inbound calls that will arrive in your time unit. Then determine λ (lambda) as follows:

$$\lambda = \text{mean number of incoming calls in your time unit.} \tag{6.1}$$

The next thing to do is to estimate the mean service time (remember your time unit) for an inbound call. You then can determine μ_1 by the following equation:

$$\mu_1 = \frac{1}{\text{mean service time for inbound calls}}. \tag{6.2}$$

The last estimation to make is the mean service time for an outbound call. With this time you can calculate the μ_2 for the CALLculator as follows:

$$\mu_2 = \frac{1}{\text{mean service time for outbound calls}}. \tag{6.3}$$

As a call center manager you should have a feeling for how many agents to schedule. If you have really no clue you can use the following idea. First determine how many agents you should use for the inbound calls. This can be done with the help of the Erlang delay formula on the web located at http://www.cs.vu.nl/~koole/erlang.html. Only three parameters are required. The first is the amount of arrivals per minute, the second is the mean service time for the inbound calls, and the third is the number of agents. Increase and decrease the number of agents until you satisfy the

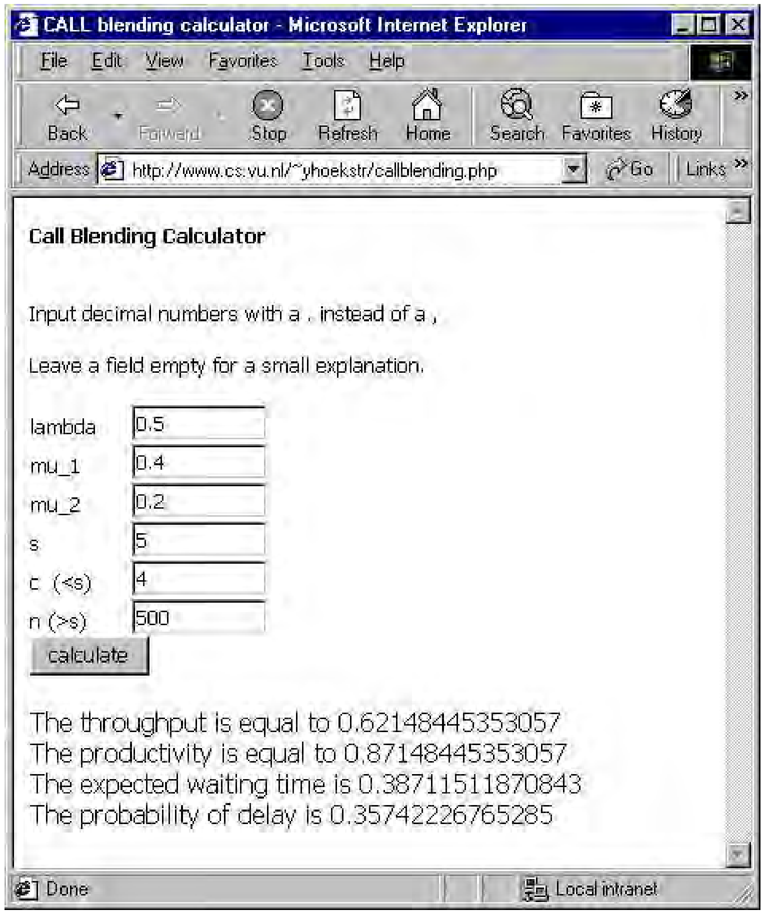

Figure 6.1: The CALLculator.

waiting time criterion. Call the number of agents found $s_{inbound}$. The second step is to determine how many agents you need to realize the throughput of the outbound calls. This can be done with the following equation:

$$s_{outbound} = \text{desired outbound number served in time unit} \cdot \frac{1}{\mu_2}. \tag{6.4}$$

Then a start for the number of agents (s) is as follows:

$$\text{s} = s_{inbound} + s_{outbound}. \tag{6.5}$$

You do not have to stick to this number. You can try to take another number of agents. You could try a smaller number of agents and see if it is still possible to meet your demands as discussed in the previous section where you saw that you needed less agents.

The parameter c is the so-called threshold. We describe the threshold policy as follows. Let y be the number of outbound calls in service and let

x be the number of inbound calls in the system. The level c, called the threshold is such that if $x + y < c$, then the proposed action is to schedule $c - x$ calls while inbound calls have a higher priority. If $x + y \geq c$, then the proposed action is to schedule no outbound calls but only inbound calls. In this paper there is no general answer for the best threshold to use. You should experiment by yourself. Take the threshold as high as possible while you meet your demands like waiting time and throughput.

By setting δ (not shown in CALLculator) you can randomize the threshold between c and $c+1$. This work as follows in practice. Add the number of outbound calls in service and the number of inbound calls in the system (so the ones in service and the ones in the queue). If this number equals $c + 1$ and an agent just finished his job, then with probability δ you do nothing and with probability $1 - \delta$ you schedule an outbound call. Up till now the CALLculator does not provide the calculations for a threshold policy with randomization between two levels i.e., $\delta = 1$ in the CALLculator. If you would like to know the outcome of a randomized policy you should estimate it by interpolation with a higher degree polynomial function. You could use Excel for this interpolation.

The last parameter n in the CALLculator is only for the accuracy of the calculations the bigger this number the better the accuracy but the longer the calculation time. You could use the following equation to get an idea of what number you can take.

$$n = s + 4 \cdot \text{expected queue length.} \tag{6.6}$$

Again try to change this parameter to see how the throughput behaves. You should take it at least as large as s, notice that in the particular case that you should take n equal to s that you would not account for queueing times. The next thing to do is click calculate and the CALLculator shows you the expected queueing time for an inbound customer and the throughput realized. Note that these numbers represent long term behavior and do not say anything about a percentage of customers that has to wait for a certain time. You cannot simply say that 50% of the customers waits less than the average queueing time.

An example of how to use the CALLculator is in Figure 6.1. For this example the following holds. The mean time between arrivals of inbound calls is 2 minutes, the average duration of an inbound call is 2.5 minutes, the average duration of an outbound call is 5 minutes, the number of agents available is 5, the threshold used is 4 and the accuracy is set to 500. When you click the calculate button you get the result as shown in Figure 6.1

Chapter 7

Conclusion

Conclusion
If call centers apply call blending it is possible to increase productivity and throughput significantly. While on the other hand at the same time it is possible to decrease the waiting times. The CALLculator developed for the calculations works good and can be used by call center mangers to get an idea of what results can be obtained by applying call blending.

Topics for further research
At first the CALLculator could be extended to randomized policies ($\delta < 1$). The CALLculator could also be extended with a comparison with the case that no call blending is applied. The CALLculator could also give the probability that a certain inbound call has to wait longer than say 10 seconds. This can be done because all stationary probabilities are already determined.

We could also look at service times that are not exponentially distributed. Another idea is to investigate the arrival rate, do the customers really arrive according to a Poisson process. We assumed that customers did not abandon and that our queues are infinite. What if queues are finite and customers are able to abandon. The last point we like to mention is that you could investigate agents with various skills.

Bibliography

[1] S. Bhulai and G.M. Koole*A Queueing model for Call Blending in Call Centers* To appear in IEEE Transactions on Automatic Control.

[2] http://www.cisco.com Manufacturer for call blending solutions.

[3] G. M. Koole and A. Mandelbaum *Queueing Models of Call Centers An Introduction* October 2001. This article can be obtained at http://www.cs.vu.nl/obp/callcenters.