

A METHOD FOR ESTIMATION OF REDIAL AND RECONNECT PROBABILITIES IN CALL CENTERS

Sihan Ding

Department of Stochastics
Center for Mathematics and Computer Science
Amsterdam, The Netherlands

Ger Koole

Department of Mathematics
VU University Amsterdam
Amsterdam, The Netherlands

Rob van der Mei

Department of Stochastics
Center for Mathematics and Computer Science
Amsterdam, The Netherlands

ABSTRACT

In practice, many call center forecasters use the total inbound volume to make forecasts. In reality, besides the *fresh calls* (initial call attempts), there are many *redials* (re-attempts after abandonments) and *reconnects* (re-attempts after answered calls) in call centers. Neglecting redials and reconnects will inevitably lead to inaccurate forecasts, which eventually leads to inaccurate staffing decisions. However, most of the call center data sets do not have customer-identity information, which makes it difficult to identify how many calls are fresh. Motivated by this, the goal of this paper is to estimate the number of fresh calls, and the redial and reconnect probabilities. To this end, we propose a model to estimate these three variables. We formulate our estimation model as a minimization problem, where the actual redial and reconnect probabilities lead to the minimum objective value. We validate our estimation results via real call center data and simulated data.

1 INTRODUCTION

In an inbound call center, a manager typically uses historical call data sets to forecast the future call volumes. Based on the call volume forecast, one can make staffing decisions. An inaccurate forecast inevitably leads to inaccurate staffing decisions (see Steckley, Henderson, and Mehrotra (2010)). There is extensive literature on different forecasting methods applied to call centers. Andrews and Cunningham (1995) used the Autoregressive Integrated Moving Average (ARIMA) method to forecast the inbound call volume of L.L. Bean's call center. Taylor (2012) adjusted the traditional Holt-Winters exponential smoothing method to the Poisson count model with gamma-distributed arrival rate, and took both intraweek and intraday patterns into account in his model. Taylor (2008) compared the accuracy of a few forecasting models for a British retail bank call center. He concluded that for forecasting horizons up to two or three days ahead, seasonal ARIMA and Holt-Winters model are more accurate, while for longer lead times, simple taking the historical average is better. Shen and Huang (2008) used the Singular Value Decomposition (SVD) method to reduce the dimension of square-root-transformed call center data. Then they applied time series and regression analysis techniques to make distributional forecasts. Besides the forecasts, they also developed a method to dynamically update the forecasts when early realizations of the day are given. The doubly stochastic model built by Jongbloed and Koole (2001) addresses the issue of high variability in call arrival

volume. This model was then further developed in Avramidis, Deslauriers, and L'Ecuyer (2004), where three variants of doubly stochastic model were analyzed and compared. Ibrahim and L'Ecuyer (2013) added the correlation between different call types into a model with additive seasonality, interday correlation and intraday correlation. A multiplicative way to model the intraweek and intraday pattern was used by Gans et al. (2009).

Call center forecasting models aim to achieve the minimum error in the forecasts, where total inbound volumes are used. In this paper, we show that the true inbound volume (we refer to it as the fresh volume from now on) is more appropriate to be used when one makes forecasts, since it is independent of the service levels in the call center. In contrast, the total inbound volumes are influenced by the service levels and staffing decisions of the call centers, due to the redial and reconnect customer behaviors. We define redials as re-attempts of abandoned calls, and reconnects as re-attempts of answered calls. Data analysis of a real call center reveals that a significant fraction of the inbound call volume involves redials and reconnects. The reason for customers to redial is clear, since abandoned customers did not get their questions answered in their initial attempts. There are several reasons for customers to reconnect. For example, a customer may check what is the status of his previous request. Also, solutions offered by agents may not be effective for customers, hence, they may reconnect. Koole (2013) gives more insights into redials and reconnects.

In order to know the fresh volume, one would need customer-identity information in the data set, such that redials and reconnects can be filtered out. However, in most of the call center data sets, customer-identity information is either not recorded or not accessible, i.e., we do not know who is the caller of each call. In other words, we do not know whether a call is a fresh call, a redial, or a reconnect. Furthermore, the fresh volume is not stable due to the existence of seasonality and trend. On the other hand, the redial and reconnect probabilities are expected to be more stable over time, since they represent customer behavior. In this paper, we will show how to estimate the number of fresh calls with the assistance of the redial and reconnect probabilities.

Hoffman and Harris (1986) are the first ones who address the issue that the total volume does not represent the true demand in call centers. Aiming to have a more accurate forecast for the call volume, they estimate the redial probability for the U.S tax-payer service telephone center. However, due to the assumptions they made, the model in Hoffman and Harris (1986) is difficult to apply to other call center data sets. First of all, they only consider redial behavior, but neglect reconnect behavior. Secondly, in their model, the fresh call arrival rate is assumed to be a constant among certain hours of the day, whereas in most call centers the arrival rate is far from constant over the day, exhibiting certain intraday pattern, see Shen and Huang (2008); Gans et al. (2009); Ibrahim and L'Ecuyer (2013). In this paper, we formulate this estimation problem as a minimization problem, where the lowest objective value is attained when the actual redial and reconnect probabilities are chosen. We analytically show that the redial probability, the reconnect probability and the fresh volumes cannot be accurately estimated simultaneously. Nevertheless, if one variable is given, the other two variables can be estimated accurately. To allow intraweek seasonality, we adjust our model to a linear programming problem, which is easy to solve. We show both via simulated data and real call center data that our estimations are close to the real values.

The remainder of the paper is organized as follows. In section 2, we describe the queueing model. We also show a simulation example of such a model to understand the influence of redials and reconnects on the total volume. In section 3, we present our estimation model for constant unknown arrival rate. We validate this estimation model in section 4 via simulated data sets. In section 5, we analyze a real center data to understand the redial and reconnect behaviors, and we adjust the estimation model such that the model is suitable for data with intraweek seasonality. The results from estimating real call center data via the adjusted model are shown in section 6.

2 MODEL DESCRIPTION

Consider the queueing model illustrated in Figure 1. We assume that calls arrive according to a Poisson process. We refer to these calls as *fresh calls*. There are s agents who handle inbound calls. An arriving

call will be handled by an available agent, if there is any. Otherwise, he will wait in a queue with infinite buffer size. The calls are handled in the order of arrival. After a generally distributed amount of time H , a waiting customer who did not get connected to an agent will lose his patience and abandon. We assume $\mathbb{E}H = \theta < \infty$. With probability p , an abandoned customer will enter the redial orbit, and he will redial after some generally distributed amount of time Γ_{RD} , with $\mathbb{E}\Gamma_{RD} = \delta_{RD} < \infty$. We refer to these calls as *redials*. With probability $1 - p$, this customer will not call back, and this call is considered as a ‘lost’ call. We assume that the service time B of a customer has a general distribution with mean $\mathbb{E}B = 1/\mu < \infty$. After the call has been finished, this customer will enter the reconnect orbit with probability q , and he will reconnect after some generally distributed time Γ_{RC} , with $\mathbb{E}\Gamma_{RC} = \delta_{RC} < \infty$. We refer to these calls as *reconnects*. We assume that p and q do not depend on customers’ experiences in the system. These experiences include holding time, waiting time and the number of times that customers have already called. We use this queueing model to represent the situation of a single-skill call center.

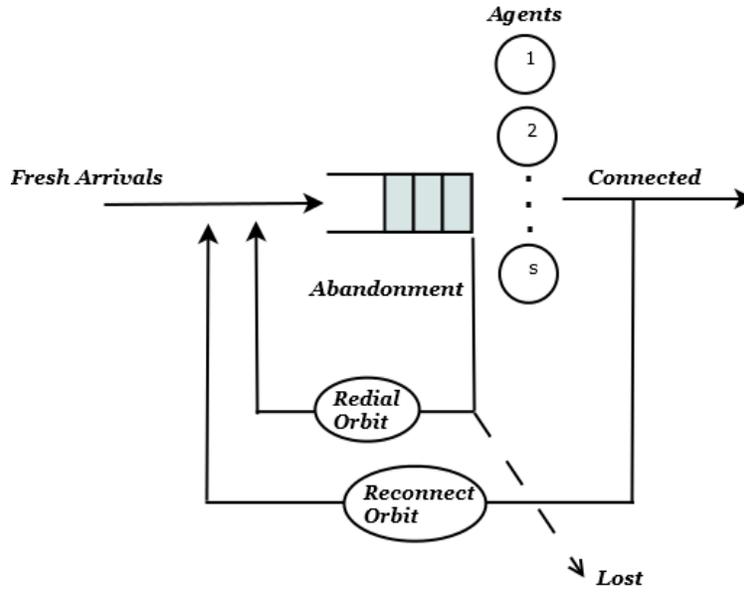


Figure 1: Queueing diagram of a call center

According to the model description, the total volume is influenced by the service level in the call center, since a bad service level leads to more abandonments, which leads to a larger number of redials. Thus, the total call volumes depend on the staffing decisions. To illustrate this, we show one simulation example. We simulate the real situation in a call center by following the diagram in Figure 1, and we generate a call center data set of 100 days. We vary the number of agents per day. For simplicity, the fresh arrival rate is set to be 10 calls per minute every day. Since the sum of independent Poisson random variables is again Poisson distributed, F_i (i.e., the fresh volume in day i) is then Poisson distributed with rate $10 \cdot 60 \cdot 24 = 14400$. We take B , H , Γ_{RD} and Γ_{RC} to be exponentially distributed in our simulation. An example of this simulated data set is shown in Table 1. The total call volume and fresh call volume of each day are plotted in Figure 2 for a 100-day time interval. In this example, we set $p = 0.5$ and $q = 0.2$.

Most interestingly, Figure 2 shows that not only the total volumes are much higher than the fresh volumes (as they should), but that they exhibit much higher variability. If a manager would be able to see the fresh volume, it will be easy to predict the future volume, since they are just Poisson realizations with constant rate. However, since the manager can not identify who the caller is in the data set, he will only see the total volume in the data set. The extra fluctuations of the total volumes are caused by two effects: (i) different service levels will lead to different abandonment rates, and that will eventually lead to differences in number of redials, (ii) effect (i) is *amplified* by the reconnects, since reconnects will worsen

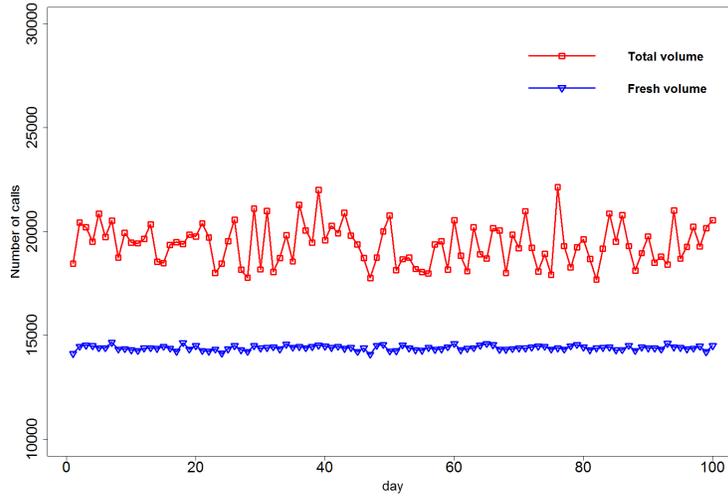


Figure 2: Call volumes from a simulation example

Table 1: An example of simulated data set

Arr time	Pick time	Dep time	Result	day
0.000	0.000	0.697	Answered	1
0.093	0.093	1.071	Answered	1
0.116	0.116	0.486	Answered	1
0.278	0.278	4.650	Answered	1

the service levels by generating extra workload to the system. On the other hand, the fresh volume has less variability and is independent of the staffing decisions. We will show our approach to estimate it as well as redial and reconnect probabilities in the next section.

3 ESTIMATION MODEL

Many call center data sets are similar to the simulated data set that we showed in Table 1; customer identity information is not available. Thus, in such call center data sets for N days, we would only know A_i, C_i ($i = 1, \dots, N$), which stands for the number of abandoned calls in day i and the number of connected calls in day i , respectively. We denote T_i as the total number of calls in day i . We define $r_i := A_i/T_i$, which is the abandonment percentage of day i . With A_i and C_i given, r_i and T_i can be obtained. We denote the number of fresh calls in day i as F_i . Our goal is to estimate F_i, p and q .

By definition, we know that the expected number of total calls is the sum of expected fresh arrivals, expected number of redials and expected number of reconnects, i.e.,

$$\mathbb{E}T_i = \mathbb{E}F_i + \mathbb{E}C_i \cdot q + \mathbb{E}A_i \cdot p. \quad (1)$$

Also, since a call is either answered or abandoned, we know that $T_i = A_i + C_i$. Insert this equation into equation (1), we obtain

$$\mathbb{E}F_i = (1 - q)\mathbb{E}C_i + (1 - p)\mathbb{E}A_i. \quad (2)$$

To estimate F_i, p and q , we start with the simple case where $F_i \sim \text{Pois}(F^*)$, i.e., each day has the same arrival rate of fresh calls, but where F^* is unknown. Note that, by this assumption, we ignore the intraweek arrival pattern in the call center data set. We will extend our model to address this pattern in section 5.

For given data points $A_1, \dots, A_N, C_1, \dots, C_N$, we consider the following minimization problem to estimate p, q and F :

$$\min_{q,p,F} \sum_{i=1}^N |(1-p)A_i + (1-q)C_i - F|, \quad (3)$$

where the objective function is the sum of absolute deviance (SAD). In this paper, we refer to \hat{p} and \hat{q} as estimated value of p and q by using our model, respectively, and p^* and q^* as the true value of p and q , respectively.

Property 1 The triple (p^*, q^*, F^*) is a local minimizer of minimization problem (3) on F ; in other words, given $p = p^*, q = q^*$, F^* is the minimizer of problem $\min_F \sum_{i=1}^N |(1-p^*)A_i + (1-q^*)C_i - F|$, when the sample size is large enough.

Property 2 The triple (p^*, q^*, F^*) is not the only local minimizer of problem (3). More specifically, all local minimizers are of form $(p', q', F^* - \frac{q' - q^*}{1 - q^*} F^*)$, where the following equation holds

$$(1 - p^*)q' - (1 - q^*)p' + (p^* + q^*) = 0, 0 \leq p', q' \leq 1. \quad (4)$$

Property 2 indicates that (p^*, q^*, F^*) is not the only minimizer of (3). In fact, it is not the global minimizer. Because the errors measured by SAD are scaled errors, in the sense that if we choose large number for \hat{p} and \hat{q} , the error would be smaller. An extreme example that indicates this scaling problem is letting $\hat{p} = 1$ and $\hat{q} = 1$, and SAD would always be 0 by choosing $\hat{F} = 0$. Therefore, we introduce the following minimization problem, which uses the Weighted Absolute Percentage Error (WAPE) instead of SAD as the objective function to remove this scaling problem

$$\min_{q,p,F} \frac{\sum_{i=1}^N |(1-p)A_i + (1-q)C_i - F|}{\sum_{i=1}^N |(1-p)A_i + (1-q)C_i|}. \quad (5)$$

To obtain (\hat{p}, \hat{q}) via minimization problem (5), we do the following: we solve (5) on a grid of p and q . For given values of p and q , the denominator of WAPE is a constant, solving (5) is simply minimizing the SAD over F . Thus, for any given combination of p and q , we can obtain

$$\text{WAPE}_{(p,q)} := \min_F \frac{\sum_{i=1}^N |(1-p)A_i + (1-q)C_i - F|}{\sum_{i=1}^N |(1-p)A_i + (1-q)C_i|}.$$

We choose \hat{p} and \hat{q} such that they lead to minimum WAPE, i.e., $(\hat{p}, \hat{q}) = \underset{0 \leq p, q < 1}{\operatorname{argmin}} \text{WAPE}_{(p,q)}$.

Property 2 also tells us that the minimum WAPE will be attained in a line on the grid of p and q . This means that we cannot estimate p, q and F simultaneously. This is not surprising, since in such data sets, we have two known parameters (namely, A_i and C_i), with three degrees of freedom (namely, p, q and F_i). On the other hand, it also indicates that when one variable is given, we can uniquely determine the other two. In a call center, there are different ways to estimate the reconnect probability. For example, we can ask agents to do some polling (e.g. for one whole day), we staff enough agents, so that almost all calls are handled, and we ask each agent to record each connected call's customer name or identity, then by the end of the day, we can calculate how many customers have called back. For the redial probability, this is more difficult to do, since abandoned customer's information is often not recorded. Using polling to determine the number of fresh calls is also difficult. Because the number of fresh calls is not stable over time, due to presence of trend and seasonality.

In this estimation model, we only consider redial and reconnect in the same day of the fresh call. We will motivate this assumption when we analyze the redial and reconnect behaviors from a real call center data set.

4 SIMULATION EXPERIMENT

In this section, we test our estimation model (5) in the data sets generated by discrete-event simulation. The data generation procedure is the same as described in Section 2. We generated five data sets using different parameter settings. The parameters are shown in Table 2. In each example, we vary the values of all different parameters. Note that we take B, H, Γ_{RC} and Γ_{RD} to be exponentially distributed in the simulation. However, since A_i and C_i are realizations which can be obtained from the data, how B, H, Γ_{RC} and Γ_{RD} are distributed becomes irrelevant in our estimation model (5). Consequently, we can extend our estimation model to call centers where these variables have general distributions.

Table 2: Parameters of the simulation experiments

Example	p	q	F^*	$1/\mu$	θ	δ_{RD}	δ_{RC}
1	0.5	0.2	4	10	2	20	100
2	0.5	0.2	10	4	2	20	100
3	0.5	0.2	10	4	2	5	10
4	0.7	0.3	4	9	3	20	100
5	0.7	0	4	9	3	20	100

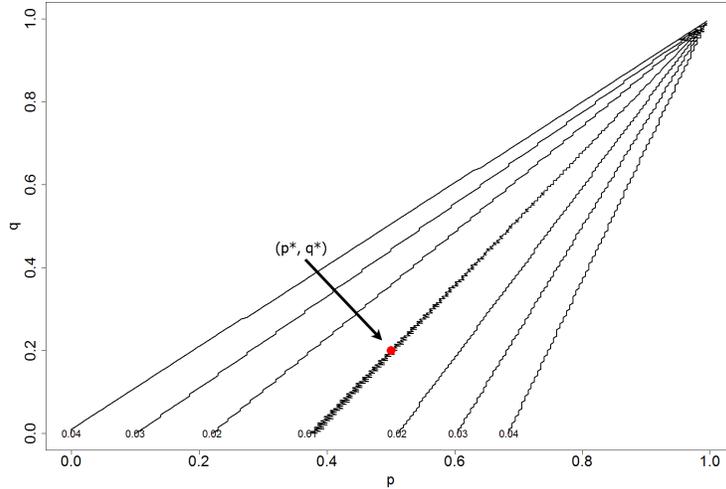


Figure 3: Values of WAPE on the grid of p and q for simulated data example 1, with the red points standing for (p^*, q^*)

Figure 3 illustrates the combinations (p, q) for which $WAPE_{(p,q)} = \gamma$, where γ is varied as 0.01, 0.02, 0.03 and 0.04, in addition to the optimum (p^*, q^*) . Other simulation examples (not shown here) gave similar graphs as in Figure 3. These figures all confirm that if the reconnect probability is known, we can estimate the radial probability accurately.

5 ANALYSIS OF REAL CALL CENTER DATA

In this section, we analyze real call center data to understand the radial and reconnect behaviors. The call center data is obtained from a call center company called Vanad Laboratories in the Netherlands. The data set consists of call arrival records to the municipality of Rotterdam in the Netherlands. The calls are recorded from 1st April 2012 to 29th September 2012. On Sundays, the call center is closed. On Saturdays, the arrival volume is quite low, i.e., 5508 total call records for 26 Saturdays. Therefore, we may ignore the

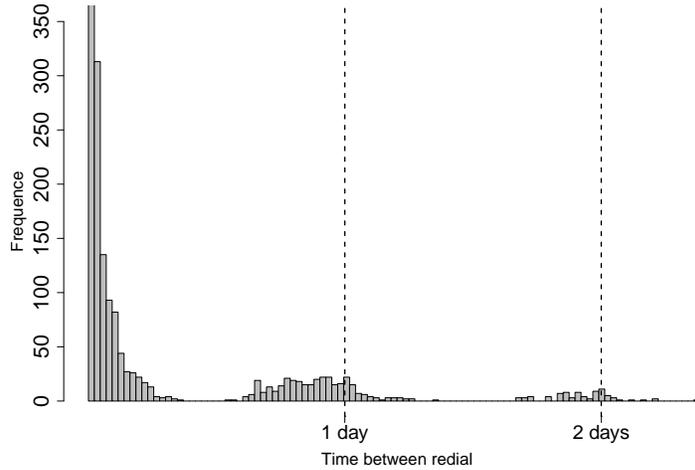


Figure 4: The histograms redial time for Router A

weekends call data, and focus only on the weekdays. We also remove the week which consists of one or few days of holidays. This leaves us with 22 weeks of data. Each call record consists of seven attributes, i.e., call arrival date, arrival time, caller’s phone number, router name, agent number, time that the call is answered and the time that the call is hanged up. Approximately 20% of the caller’s phone numbers is unidentified, since some callers set their phone number to be invisible by the call receivers. We assume that each caller is identified by its phone number. There are different routers that can be selected by an caller. The selection of router is done by customer during Interactive Voice Response (IVR) unit. After the customer has made the selection, his call will be distributed by an Automatic Call Distributor (ACDs). Each router represents one or multiple types of questions that a customer may have. We will focus our study on two specific routers which are referred to as router A and router B.

5.1 Redial and Reconnect Behavior

For this data set, we do have the caller-identity information, which allows us to we can follow customer and see whether he called back or not. In Figures 4, we plot the histograms of time-to-redial for router A, i.e., Γ_{RD} . For time to reconnect and router B, we obtain a similar figure. We can see that both for redial and reconnect, most of the customers call back in the same day as their fresh calls. A small fraction of the customers redial or reconnect one or two days later after the fresh call. Therefore, in our model, it is sufficient to assume that the redials and the reconnects arrive in the same day as the fresh call, i.e., customer who calls again one or more days later will be regarded as another fresh call. Some descriptive statistics are shown in Table 3.

Table 3: Descriptive statistics

Router	Total volume	Fresh volume	Redials	Reconnects
A	41624	36515	2142	2967
B	28526	23782	1117	3627

After removing the unidentified calls, we use following formulas to calculate redial and reconnect probabilities

$$p^* = \frac{\sum_{i=1}^N RD_i}{\sum_{i=1}^n A_i}, \quad (6)$$

$$q^* = \frac{\sum_{i=1}^N RC_i}{\sum_{i=1}^n C_i}, \quad (7)$$

where RD_i is the number of redials in day i , and RC_i is the number of reconnects in day i .

The probabilities p^* and q^* of each weekday are shown in Table 4 and 5.

Table 4: Real redial and reconnect probabilities of each weekday for Router A.

	Mon	Tue	Wed	Thu	Fri
p^*	0.52	0.52	0.46	0.49	0.43
q^*	0.08	0.08	0.07	0.09	0.08

Table 5: Real redial and reconnect probabilities of each weekday for Router B.

	Mon	Tue	Wed	Thu	Fri
p^*	0.42	0.40	0.38	0.38	0.39
q^*	0.15	0.13	0.14	0.15	0.12

From Tables 4 and 5 we see that the redial and reconnect probabilities are different, i.e., redial probability is usually larger than the reconnect probability. Intuitively, this makes sense, since an abandoned customer has higher urge to call back than an answered customer. For different routers, redial probability has more fluctuations than reconnect probability. However, within the same router for every weekday both probabilities are stable, only except for the redial probability for router A on Friday. Therefore, it is sufficient to have two parameters for all weekdays together for redial and reconnect probabilities of each router.

5.2 Intra-week Seasonality

In model (5), we made the assumption that each day has the same fresh call arrival rate. This is an unrealistic assumption in a real call center. As illustrated in the paper by ? as well as other papers, call center data would show strong intra-week patterns. To make our model applicable in call center data with intra-week seasonality, we make adjustments to model (5). To this end, we assume that the weekly total fresh calls distributed to each day of the week in a multiplicative way, i.e.,

$$\mathbb{E}F_i = \mathbb{E}WF_{w_i} \cdot \beta_{d_i},$$

where w_i and d_i are the week number of day i and the weekday of day i , respectively, $d_i \in \{1, 2, 3, 4, 5\}$, $w_i = 1, 2, \dots, n$, where n is the number of weeks. WF_{w_i} is the total number of fresh calls of week w_i . Thus, β_{d_i} can be interpreted as the proportion of calls on weekday d_i than in the whole week. A key assumption of this multiplicative model is that β_{d_i} does not depend on the week number (relaxation of this assumption is beyond the scope of this paper, see also section 7). Therefore, our estimation model changes to

$$\begin{aligned} \min_{p, q, \beta} & \frac{\sum_{i=1}^N |(1-p)A_i + (1-q)C_i - WF_{w_i} \cdot \beta_{d_i}|}{\sum_{i=1}^N [(1-p)A_i + (1-q)C_i]} \\ \text{s.t.} & \sum_{d_i} \beta_{d_i} = 1, \\ & WF_{w_i} = WA_{w_i}(1-p) + WC_{w_i}(1-q), \quad w_i = 1, 2, \dots, n, \end{aligned} \quad (8)$$

where WA_{w_i} and WC_{w_i} are the total number of abandoned calls and total number of connected calls in week w_i , respectively. The intuition behind model (8) is that the daily fresh call volume is proportional to the weekly total fresh call volume.

Solving minimization problem (8) is difficult even on the grid of p and q , noting that the objective function is not differentiable. Also, unlike model (5), in model (8), we have a hard constraint on variables β . Therefore, we reformulate problem (8) into a linear programming (LP) problem on a grid of p and q , which is easier to solve. The corresponding LP problem of model (8) can be written as

$$\begin{aligned}
 & \min_{\beta, Z^+, Z^-} \sum_{i=1}^N (Z_i^+ + Z_i^-) \\
 & \text{s.t.} \quad \sum_{d_i} \beta_{d_i} = 1, \\
 & \quad Z_i^+ - Z_i^- = \frac{|(1-p)A_i + (1-q)C_i - WF_{w_i} \cdot \beta_{d_i}|}{\sum_{i=1}^n [(1-p)A_i + (1-q)C_i]}, \text{ for all } i, \\
 & \quad WF_{w_i} = WA_{w_i}(1-p) + WC_{w_i}(1-q), \text{ for all } i, \\
 & \quad 0 \leq p, q, \beta \leq 1, \\
 & \quad Z_i^+, Z_i^- \geq 0 \text{ for all } i.
 \end{aligned} \tag{9}$$

Note that the second constraint is linear, since when p and q are given, term $\sum_{i=1}^N [(1-p)A_i + (1-q)C_i]$ is a constant. One can show that problem (8) and (9) are equivalent. We will omit the proof here.

6 RESULTS

In subsection 5.2 we adjusted our estimation model to the real call center data set with multiplicative intraweek patterns. We simplified the original optimization problem (8) to a LP problem (9). Then we solved this LP problem for given p and q . We apply this method to the Vanad Laboratories data set, and plot the points with different value of WAPE on a grid of p and q , with grid size of 0.01. The estimation results of router A is in Figure 5. One could see from this figure that the minimum WAPE is attained in thicker lines which are wider than the ones from the simulated data sets. Estimation result for router B (not shown here) gave similar graph as in Figure 5. Therefore, instead of obtaining a line of minimum WAPE, for this data set, we have a region of minimum WAPE. This could be caused by multiple reasons; (i) the redial and reconnect probabilities of each weekday are slightly different, while we assume them to be the same every weekday; (ii) extra error is introduced when we model seasonality.

In Figures 5, the true parameter lies close to the line where minimum WAPE is attained. In Table 6, we compute the real redial and reconnect probabilities using formulas (6) and (7). Assume the real reconnect probability p^* can be calculated accurately via polling method which we described in section 3, our estimated redial probability is shown in Table 6 under the column $\hat{p}|q^*$. One can see that our estimation of redial probability for router A is approximately 0.05 higher than the real redial probability, while for router B, our estimation is about 0.09 higher. For a call center with $r = 20\%$, i.e., 20% of all calls are abandoned, 0.09 error in redial probability would lead to less than 2% error in estimating the number of fresh calls. For these two routers, which have much less abandonments than 20%, 0.09 would lead to even less errors. Therefore, maximum of 0.09 error in our estimation of redial probability is acceptable. In Table 6, we also compare the real number of fresh calls F_i^* with our estimation of the number of fresh calls \hat{F}_i by using \hat{p} and q^* , which can be obtained via $\hat{F}_i = (1 - \hat{p})A_i + (1 - \hat{q})C_i$. To obtain F_i^* , we remove all the redials and reconnects in the data set, and we assume that the unidentified customers have the same behaviors as identified customers. We use the WAPE of the fresh calls as the error measurement, i.e., $WAPE_F$, which

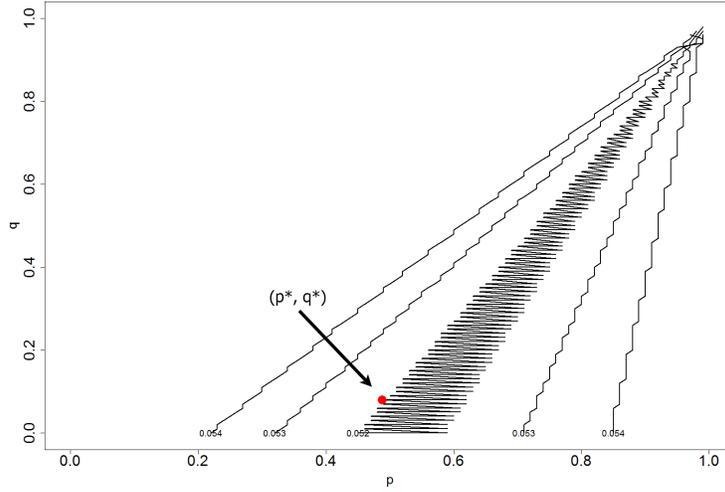


Figure 5: Values of WAPE on the grid of p and q for Router A

is defined as

$$\text{WAPE}_F := \frac{\sum_{i=1}^N |\hat{F}_i - F_i^*|}{\sum_{i=1}^N |F_i^*|}.$$

One can see from Table 6 that the WAPE_F for both routers are quite small, which are both less than 3%. The real fresh calls and the estimated fresh calls of router A are plotted Figure 6. In this figure, our estimations \hat{F}_i are quite close to the real fresh calls F_i^* . For router B, we obtained a similar graph (not shown here) as in Figure 6.

Table 6: Estimation results for both routers

Router	CI_{p^*}	CI_{q^*}	$\hat{p} q^*$	WAPE_F
A	0.488	0.080	0.54	2.6%
B	0.398	0.141	0.49	3.0%

7 Conclusion and further research

In this paper, we proposed a model to estimate the number of fresh calls of a call center, which is hidden under the total number of calls. The fresh calls represent the true demand, since they do not depend on the service levels of the call center, while the total number of calls do. Thus, the number of fresh calls should be used when one makes forecasts. However, direct calculation of the number of fresh calls is difficult in some call centers, since customer identity information is not available in their data. In our model, we try to estimate the redial probability, reconnect probability and the fresh calls simultaneously in call center data without customer identity information by solving a minimization problem. However, we show analytically that although the actual redial probability, reconnect probability and the number of fresh calls is indeed a minimizer of this minimization problem, they are not the only minimizer. In order to have an accurate estimation, one variable needs to be given. We propose a polling method in call centers to calculate the reconnect probability. Once the reconnect probability is given, the other two variables can be estimated. Via simulated data, our estimation results of redial probability and the number of fresh calls are accurate. We also validate our model via two real call center data sets. Our estimation of redial probabilities for both data sets are close to the actual redial probabilities, with errors of less than 0.09. Furthermore, our

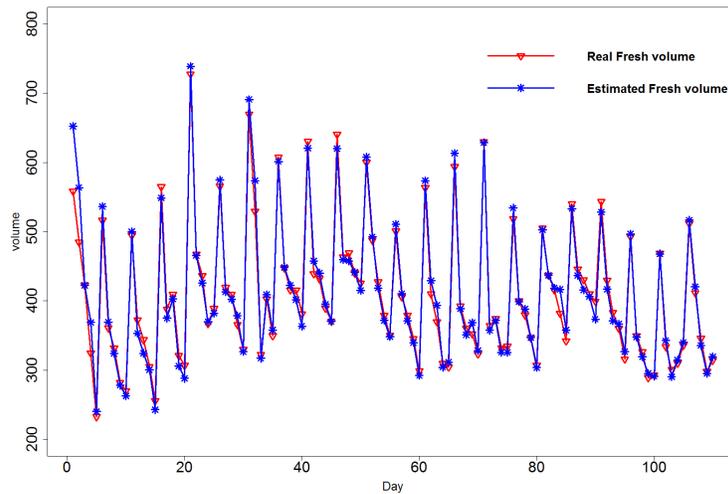


Figure 6: Real number of fresh calls VS. estimated number of fresh calls for router A

estimation of the number of fresh calls are very close to the real number of fresh calls, with the WAPE less than 3%.

In addition to helping call center managers to estimate the true demands of the call centers, this paper also addresses the reconnect customer behavior in call centers. In the data set of Vanad Laboratories, we found that the number of reconnects is significant. Neglecting it will lead to inaccurate prediction of the call volumes, which will cause inaccurate staffing in the call centers. Inspired by these findings in this paper, we propose the following topics for further research:

1. For a call center manager, it would be interesting to know what are the consequences of neglecting reconnects in terms of costs or service levels;
2. In order to make the right staffing decisions, it would be useful to evaluate the service levels of a call center with consideration of the reconnect behaviors;
3. The redial and reconnect behaviors will introduce intraday correlation to the call center data. For example, if a manager saw a busy morning, he would expect a busy afternoon, since some “morning customers” will redial or reconnect in the afternoon. This raises the question how managers should update the agent schedules dynamically when morning realizations are available.

Acknowledgements

The authors would like to thank Vanad Laboratories for sharing the data set with us, helping us to understand the data set, and answering related questions. The authors also thank the municipality of Rotterdam for allowing us to use their data. The authors also thank Jan Kalden for the inspiring discussions. Finally, the authors appreciate professor Avi Mandelbaum and the Service Enterprise Engineering (SEE) Lab for offering another data set, although it is not used in this paper.

REFERENCES

- Andrews, B., and S. Cunningham. 1995. “LL Bean improves call-center forecasting”. *Interfaces* 25 (6): 1–13.
- Avramidis, A. N., A. Deslauriers, and P. L’Ecuyer. 2004. “Modeling daily arrivals to a telephone call center”. *Management Science* 50 (7): 896–908.

- Gans, N., H. Shen, Y. Zhou, K. Korolev, A. McCord, and H. Ristock. 2009. "Parametric stochastic programming models for call-center workforce scheduling". Technical report. Available via <http://faculty.washington.edu/yongpin/Gans-Shen-Zhou-Scheduling.pdf> [accessed March 27, 2013].
- Hoffman, K., and C. Harris. 1986. "Estimation of a caller retrial rate for a telephone information system". *European Journal of Operational Research* 27 (2): 207–214.
- Ibrahim, R., and P. L'Ecuyer. 2013. "Forecasting Call Center Arrivals: Fixed-Effects, Mixed-Effects, and Bivariate Models". *Manufacturing & Service Operations Management* 15 (1): 72–85.
- Jongbloed, G., and G. Koole. 2001. "Managing uncertainty in call centres using Poisson mixtures". *Applied Stochastic Models in Business and Industry* 17 (4): 307–318.
- Koole, G. 2013. *Call Center Optimization*. MG books, Amsterdam.
- Shen, H., and J. Z. Huang. 2008. "Interday forecasting and intraday updating of call center arrivals". *Manufacturing & Service Operations Management* 10 (3): 391–410.
- Steckley, S., S. Henderson, and V. Mehrotra. 2010. "Forecast errors in service systems". *Probability in the Engineering and Informational Sciences* 23 (2): 305.
- Taylor, J. 2008. "A comparison of univariate time series methods for forecasting intraday arrivals at a call center". *Management Science* 54 (2): 253–265.
- Taylor, J. 2012. "Density forecasting of intraday call center arrivals using models based on exponential smoothing". *Management Science* 58 (3): 534–549.