I²SDS
The Institute for Integrating Statistics in Decision Sciences

Technical Report TR-2008-3
April 18, 2008

Modeling Latent Sources in Call Center Arrival Data

Joshua Landon
Institute for Integrating Statistics in Decision Sciences
The George Washington University, Washington, DC

Fabrizio Ruggeri
CNR-IMATI, Milano

Refik Soyer
Department of Decision Sciences
The George Washington University, Washington, DC

M. Murat Tarimcilar
Department of Decision Sciences
The George Washington University, Washington, DC

Modeling Latent Sources in Call Center Arrival Data

Joshua Landon
Institute for Integrating Statistics in Decision Sciences
The George Washington University, Washington, DC

Fabrizio Ruggeri
CNR-IMATI, Milano

Refik Soyer*
Department of Decision Sciences
The George Washington University, Washington, DC

M. Murat Tarimcilar
Department of Decision Sciences
The George Washington University, Washington, DC

ABSTRACT

In this paper, we discuss issues that arise in the analysis of call center arrivals that are mostly linked to individual ads. More specifically, we consider the case where there is no complete linkage between the calls and the advertisements that led to the calls. The ability to model and infer such latent call arrival sources is important from a marketing as well as an operations point of view since knowledge of the linkage improves forecasting performance of the model. We pose this as a missing data problem and develop a data augmentation algorithm for the Bayesian analysis. We implement the proposed algorithm to simulated and actual call center arrival data and discuss its performance.

*Contact author: Department of Decision Sciences, Funger Hall 415, The George Washington University, Washington, DC 20052, USA. E-mail: soyer@gwu.edu

1. Introduction and Overview

The focus of the previous work in call center arrival modeling was primarily on forecasting models for optimal scheduling and staffing of telephone operators in call centers [Gans et al. (2003)]. Some of the work involved use of time series models such as ARIMA processes and transfer function models as in Andrews and Cunningham (1995); queuing models as in Jongbloed and Koole (2001), doubly stochastic Poisson models as in Avramidis et al. (2004). More recently, Bayesian nonhomogeneous Poisson process (NHPP) models have been considered by Soyer and Tarimcilar (2008) and by Weinberg, Brown and Stroud (2007).

The focus of the previous work was to model the call arrival (demand) process based on aggregate arrival data. An exception to this is the work by Soyer and Tarimcilar (2008) where the authors considered modeling of the call center arrival process for evaluating efficiency of advertisement and promotion policies to develop marketing strategies for call centers. As noted in the recent comprehensive review of the literature in call center research by Aksin, Armony and Mehrotra (2007), there is an important interface between operations and marketing components of call centers. However, most of the previous research has failed to emphasize this interface.

In this paper we consider an extension of the Bayesian call center arrival models of Soyer and Tarimcilar (2008) where there is no complete linkage between the calls and the advertisements that led to the calls. In reality it is not uncommon to have cases where the customer does not know the ad s/he is responding to. The ability to model and infer such latent call arrival sources is important from a marketing as well as an operations point of view since knowledge of the linkage improves forecasting performance of the models.

In what follows, we first present the modulated Poisson process model considered for call arrivals by Soyer and Tarimcilar (2008) and introduce a model that takes into account the issue of incomplete linkage. This is done in Section 2. This model enables us

to formulate the incomplete linkage problem as a missing data problem and provides us with a framework to infer the unknown sources of the calls. In Section 3 we introduce a data augmentation algorithm to develop a Bayesian analysis of the model and discuss how posterior and predictive distributions are obtained. We illustrate how advertisement specific as well as aggregate call arrival predictions can be made using the model. The implementation of the data augmentation algorithm is illustrated with two examples in Section 4 where we discuss the performance of the algorithm and its accuracy.

2. Modulated Poisson Process Model with Latent Variables

Following Soyer and Tarimcilar (2008) we define $N_i(t)$ as the number of calls arrived during a time interval of length t as response to the ith advertisement and $\boldsymbol{Z}_i$ as a $p \times 1$ vector of covariates that describe the characteristics of the ith advertisement. Typically, the covariate vector $\boldsymbol{Z}_i$ will consist of media expense (in \$'s), venue type (monthly magazine, daily newspaper etc.), ad format (full page, half page, color etc.), offer type (free shipment, payment schedule etc.) and seasonal indicators.

To reflect the fact that effectiveness of advertisement i is a function of time, the authors assumed that $N_i(t)$ is described by a nonhomogeneous Poisson process (NHPP) with intensity function $\lambda_i(t)$ and noted that the effectiveness of the advertisement, that is, its ability to generate calls, decreases by time. In order to consider the effect of covariates on the call volume intensity a modulation of the NHPP is considered following Cox (1972). The modulated NHPP model has the intensity function

$$\lambda_i(t, Z_i) = \lambda_0(t)\, e^{\boldsymbol{\beta}' Z_i}$$

where $\lambda_0(t)$ is the baseline intensity function and $\boldsymbol{\beta}$ is a vector of parameters.

The cumulative intensity (mean-value) function of the modulated NHPP is

$$\Lambda_i(t, \boldsymbol{Z}_i) = \Lambda_0(t) e^{\boldsymbol{\beta}' \boldsymbol{Z}_i} \tag{2.1}$$

where $\Lambda_0(t) = \int_0^t \lambda_0(s)ds$. Given $\Lambda_0(t)$, $\boldsymbol{\beta}$, and $\boldsymbol{Z}_i$, the distribution of $N_i(t)$, is a Poisson model with

$$P(N_i(t) = n|\Lambda_0(t), \boldsymbol{\beta}, \boldsymbol{Z}_i) = \frac{(\Lambda_0(t)e^{\boldsymbol{\beta}'\boldsymbol{Z}_i})^n}{n!}\, exp(-\Lambda_0(t)\, e^{\boldsymbol{\beta}'\boldsymbol{Z}_i}).$$

In the modulated NHPP model (or PIM) if we assume a parametric form for $\Lambda_0(t)$ and a parametric prior for $\boldsymbol{\beta}$ then we can do a fully parametric Bayesian analysis. Soyer and Tarimcilar (2008) use the *power law model*

$$\Lambda_0(t) = \gamma\, t^{\alpha} \tag{2.2}$$

with intensity function $\lambda_0(t) = \gamma\alpha t^{\alpha-1}$, where $\alpha > 0$, $\gamma > 0$ and point out that values of $\alpha < 1$ implies that the effectiveness of ads deteriorate with time.

2.1 A Latent Variable Model

In reality it is not uncommon to have cases where 100 % linkage between the ads and the calls may not exist for all calls. As noted by Soyer and Tarimcilar (2008), one strategy to deal with this problem is to treat these "unassigned" calls as if they were generated by the same call arrival process and to assume no covariate information for the unassigned calls. More specifically, for all unassigned calls, the authors assume that the intensity function is given by

$$\Lambda_0^u(t) = \delta_u \Lambda_0(t) = \delta_u \gamma\, t^{\alpha},$$

where δ_u is a random component which rescales the baseline intensity function to reflect the behavior of all the unassigned calls.

However, this approach does not enable us to infer the ad sources of the unassigned calls and thus rather limited. Alternatively, we can treat this as a missing data problem and introduce latent variables as will be discussed in the sequel.

Assume that we have m ads with starting times $T_1, \ldots, T_m$. Calls are recorded as number of arrivals in the intervals $I_j = (t_{j-1}, t_j]$ where n_{ij} is the number of calls associated with ad i in the interval j. In addition to the n_{ij} linked calls in interval j we observe the number of calls can not be linked to any ad. Let u_j denote the number of calls that can not be linked to any individual ad in interval I_j.

Given the above setup we still have a NHPP for ad $N_i(t)$ with cumulative intensity function

$$\Lambda_i(t) = \Lambda_0(t - T_i)e^{-\boldsymbol{\beta}'\boldsymbol{Z}_i}1_{[T_i,\infty)}(t). \tag{2.3}$$

and given $\Lambda_i(t)$ we have the independent increments property and $N_i(t)$'s are independent of each other. We define latent variables Y_{ij}, $j = 1, \ldots, m$ to denote the unobserved number of calls in interval I_j that was intended for ad i. Let $\underline{Y}_j$ denote the vector of such latent variables for time interval j, that is, $\underline{Y}_j = (Y_{1j}, \cdots, Y_{mj})$. Furthermore, define p_{ij} as the probability that any unlinked call in interval j has arrived as response to ad i. Thus, we assume that the latent vector $\underline{Y}_j$ follows a multinomial distribution denoted as $\underline{Y}_j \sim Mult(u_j;\, p_{1j}, \ldots p_{mj})$, where $u_j = \sum_{i=1}^{m} Y_{ij}$ and u_j is observed for all intervals. Also, the latent vectors $\underline{Y}_j$'s are conditionaly independent random variables.

3. Bayesian Analysis of the Latent Variable Model

In the modulated NHPP model with missing links we assume a Dirichlet prior on $\underline{p}_j = (p_{1j}, \ldots p_{mj})$ with parameters $(\alpha_{1j}, \alpha_{2j}, \ldots, \alpha_{mj})$ which is independent across the intervals. Thus, for the jth interval we assume a Dirichlet prior as

$$\pi(\underline{p}_j) \propto p_{1j}^{\alpha_{1j}-1} \cdots\, p_{mj}^{\alpha_{mj}-1}. \tag{3.1}$$

It follows from (3.1) that

$$E[p_{ij}] = \frac{\alpha_{ij}}{\sum_{k=1}^{m} \alpha_{kj}}. \tag{3.2}$$

We can specify the prior parameters as

$$\alpha_{ij} \propto exp\{-\delta(t_j - T_i)\} 1_{[T_i,\infty)}(t_j) \tag{3.3}$$

to reflect the fact that larger number of calls during the early phases of the life of an ad will be followed by a decrease over time with $\alpha_{ij} = 0$ for inactive ads.

Priors for other parameters can be specified independent of $\underline{Y}_j$'s and $\underline{p}_j$'s. For example, the prior for γ can be specified as a gamma distribution as $\gamma \sim Gamma(a_\gamma, b_\gamma)$ and a multivariate normal prior can be used for the covariate parameter vector $\boldsymbol{\beta}$. An independent gamma prior can be chosen for α as $\alpha \sim Gamma(a_\alpha, b_\alpha)$.

The Bayesian analysis of the model requires use of a data augmentation step in the MCMC setup. More specifically, by introducing the independent latent vectors $\underline{Y}_j$ for each of the intervals I_j, $j = 1, \ldots, n$, we can design a Gibbs sampler to draw samples from the posterior distribution of all unknown quantities.

Note that given $\underline{Y}_j$'s for all intervals, we know that $(n_{ij} + Y_{ij})$'s are independent Poisson's with parameters $\Delta_{ij} = \Lambda_i(t_j - T_i) - \Lambda_i(t_{j-1} - T_i)$ and

$$\Lambda_i(t) = \gamma t^\alpha e^{-\boldsymbol{\beta}' \boldsymbol{Z}_i}. \tag{3.4}$$

The (conditional) likelihood based on data from all intervals is given by

$$\prod_{j=1}^{n} \binom{u_j}{Y_{1j} \cdots Y_{mj}} \Big\{ \prod_{i=1}^{m} \frac{\Delta_{ij}^{n_{ij}+Y_{ij}}}{(n_{ij} + Y_{ij})!} \, e^{-\Delta_{ij}} \, p_{ij}^{Y_{ij}} \Big\}. \tag{3.5}$$

In implementation of the Gibbs sampler we also draw from the full conditional posteriors of $\underline{Y}_j$'s. The full conditional posterior of $\underline{Y}_j$ is given by

$$\propto \binom{u_j}{Y_{1j}\cdots Y_{mj}} \Big\{ \prod_{i=1}^{m} \frac{\Delta_{ij}^{n_{ij}+Y_{ij}}}{(n_{ij}+Y_{ij})!}\ p_{ij}^{Y_{ij}} \Big\}. \tag{3.6}$$

Alternatively, we can look at the full conditional posterior of Y_{ij} which is given by

$$\propto \frac{u_j!}{Y_{ij}!\ (u_j - \sum_{k=1}^{m-1} Y_{kj})!}\ \frac{\Delta_{ij}^{n_{ij}+Y_{ij}} \Delta_{mj}^{u_j - Y_{ij}}}{(n_{ij}+Y_{ij})!\,(n_{mj}+u_j-\sum_{k=1}^{m-1} Y_{kj})!}\ p_{ij}^{Y_{ij}}\ p_{mj}^{u_j-Y_{ij}},$$

where the normalizing constant is

$$\sum_{Y_{ij}=0}^{u_j} \frac{u_j!}{Y_{ij}!\ (u_j - \sum_{k=1}^{m-1} Y_{kj})!}\ \frac{\Delta_{ij}^{n_{ij}+Y_{ij}} \Delta_{mj}^{u_j - Y_{ij}}}{(n_{ij}+Y_{ij})!\,(n_{mj}+u_j-\sum_{k=1}^{m-1} Y_{kj})!}\ p_{ij}^{Y_{ij}}\ p_{mj}^{u_j-Y_{ij}}.$$

With a Dirichlet prior on $\underline{p}_j$ with parameters $(\alpha_{1j}, \alpha_{2j}, \ldots, \alpha_{mj})$ the full conditional posterior of $\underline{p}_j$ will be also a Dirichlet distribution with parameters $(\alpha_{1j}+Y_{1j}, \ldots, \alpha_{mj}+Y_{mj})$.

If the prior for γ is a gamma then the full conditional of γ is a gamma. More specifically, if we have $\gamma \sim Gamma(a,b)$ then we can easily show that the full conditional is a gamma with shape parameter $\Big[a + \sum_{i=1}^{m}\sum_{j=1}^{n}(n_{ij}+Y_{ij})\Big]$ and scale parameter $\Big[b + \sum_{i=1}^{m}(t_n - T_i)^{\alpha}\, exp(\boldsymbol{\beta}'\boldsymbol{Z}_i)\Big]$ where t_n is the end point of the last interval.

Draws from full conditionals of α and $\boldsymbol{\beta}$ can be obtained as before using adaptive rejection sampling [see for example, Gilks and Wild (1992)] or Metropolis-Hastings [see for example, Chib and Greenberg (1995]. The full conditional of $\boldsymbol{\beta}$ is given by

$$\propto \prod_{j=1}^{n} \Big\{ \prod_{i=1}^{m} \Delta_{ij}^{n_{ij}+Y_{ij}}\, e^{-\Delta_{ij}} \Big\} \pi(\boldsymbol{\beta}) \tag{3.7}$$

and the full conditional of α is given by

$$\propto e^{-\gamma \sum_{i=1}^{m} (t_n - T_i)^{\alpha}\, exp(\boldsymbol{\beta}' Z_i)}\ \pi(\alpha) \prod_{j=1}^{n} \Big\{ \prod_{i=1}^{m} [(t_j - T_i)^{\alpha} - (t_{j-1} - T_i)^{\alpha}]^{n_{ij}+Y_{ij}} \Big\}. \tag{3.8}$$

It is possible to use the model for prediction of call volume at different time periods generated by a single ad or by a group of ads that are active at the time. It is important to note that once the posterior distribution of α, γ and $\boldsymbol{\beta}$ is available we can provide these predictions for any time interval both for an individual ad as well for a group of ads. Given the values α, γ and $\boldsymbol{\beta}$ our inferences about call arrivals in any time interval are independent of latent variables Y_{ij}'s. In other words, assessment of Y_{ij}'s are crucial for making correct inferences about parameters of the modulated NHPP model but they do not directly provide information about the arrivals.

Given the posterior sample $\left\{\alpha^{(g)},\, \gamma^{(g)},\, \boldsymbol{\beta}^{(g)}\right\}_{g=1}^{G}$ from the joint posterior distribution using the Gibbs sampler, we can make call arrival predictions. For a single ad i we can approximate the probability of n calls in the interval $(s,t]$, $s<t$, for advertisement i as

$$P(N_i(t)-N_i(s)=n) \simeq \frac{1}{G}\sum_{g=1}^{G} P(N_i(t)-N_i(s)=n|\alpha^{(l)},\, \gamma^{(l)},\, \boldsymbol{\beta}^{(l)}, \boldsymbol{Z}_i), \quad (3.9)$$

where $n=0,1,\,2,\ldots,$ and $P(N_i(t)-N_i(s)=n|\alpha,\,\gamma,\,\boldsymbol{\beta},\boldsymbol{Z}_i)$ is given by the Poisson model.

If we consider m ads, then, given $\alpha,\gamma,\,\boldsymbol{\beta}$, and $\boldsymbol{Z}_i$'s, we have m conditionally independent NHPPs, $N_i(t)$'s. Thus, the cumulative number of calls generated by the m ads

$$N(t)=\sum_{i=1}^{m} N_i(t) \quad (3.10)$$

is a superposition of m independent NHHPs implying that $N(t)$ is a NHPP with the cumulative intensity function

$$\Lambda(t,\,\boldsymbol{Z}) = \sum_{i=1}^{m} \Lambda_i(t,\boldsymbol{Z}_i)\,1(t-T_i), \quad (3.11)$$

where T_i is the issue date of the ith ad, $1(\bullet)$ is the indicator function. The posterior predictive distribution of cumulative number of calls can be obtained using a Monte Carlo integral approximation as in (3.9) by replacing $N_i(t)$ with $N(t)$ and using (3.11) as the cumulative intensity.

4. Numerical Illustrations

4.1 Example using Simulated Data

We consider data simulated from a modulated nonhomogeneous Poisson process with cost of the advertisng as the single covariate and with baseline cumulative intensity is a power law function. Thus, the cumulative intensity function for advertisiment i is given by

$$\Lambda_i(t, Z_i) = \gamma\, t^{\alpha} e^{\beta Z_i}$$

where β is a scalar and Z_i is the cost of the ith advertisement. Data was generated for 10 different ads starting at the same time assuming $\gamma = 10$, $\alpha = 0.5$ and $\beta = 0.1$. The costs of the ads changed between 1 an 10 units and 20 time intervals were generated for each ad.

After the complete data was simulated, in each time interval it was assumed that there were certain calls whose source ads were unknown. These latent calls for each interval j, that is, u_j's were simulated independently assuming a binomial distribution and Y_{ij}'s were then generated using the multinomial distributions $Mult(u_j;\, p_{1j}, \ldots p_{mj})$. In choosing the multinomial probabilities it was assumed that p_{ij}'s were proportional to the cost of the ads implying that ads with high costs and thus high call arrival intensities are more likely to have latent calls.

Once the data was generated, the Bayesian methodology of Section 3 was applied using diffused priors for all the parameters. More specifically, we assumed Dirichlet priors for $\underline{p}_j$'s with parameters $\alpha_{ij} = 1$ for all $i = 1, \ldots, 10$ and $j = 1, \ldots, 20$. Prior for β

was assumed to be normal with mean 0 and variance 100. Similarly, diffused gamma priors were specified for the parameters α and γ.

It is important to note that the data available to us for analysis consists of the number of calls n_{ij}'s with known sources, that is, advertisements, as well as the number of calls u_j with latent sources for each interval. Thus, the following results are obtained using the data augmentation type algorithm presented in Section 3.

In Figure 1, we present the posterior distribution of the shape parameter α of the baseline intensity function on the left. As we can see the posterior distribution is concentrated around (0.45, 0.6) which captures the actual value of 0.5. Posterior distribution of the cost parameter β is illustrated on the left in Figure 1. The posterior density is concentrated in the region (0.06, 0.12) which again includes the actual value of 0.1. The posterior distribution of the scale parameter γ, which is not shown here, is concentrated in the (8, 12) interval and peaked at the actual value of $\gamma = 10$.

All these results suggest that the Bayesian approach presented in Section 3 to deal with latent advertisement sources seems to be performing satisfactorily. We can see this more explicitly by analyzing the posterior distributions of Y_{ij}'s for different intervals and by comparing the posterior inferences with the actual values of Y_{ij}'s that are known to us from the simulation. In Figure 2, we present the posterior distributions of Y_{ij}'s for time interval 8 where, for example, label "Yi8a1" denotes the posterior distribution for ad 1.

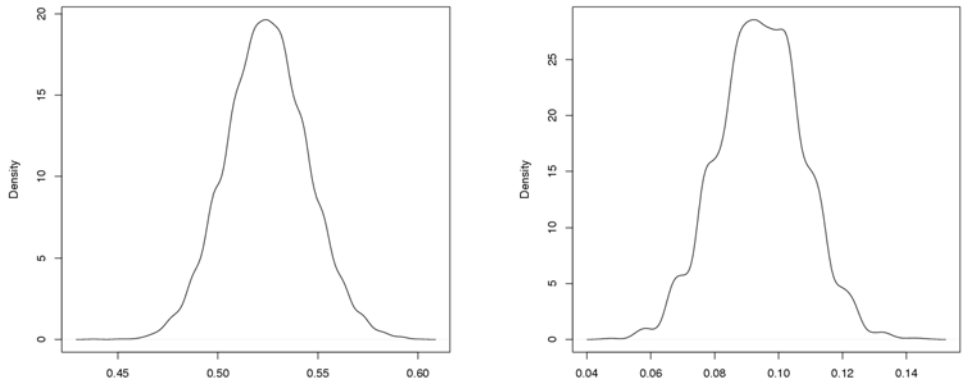

Figure 2. Posterior distributions of α (left panel) and β (right panel).

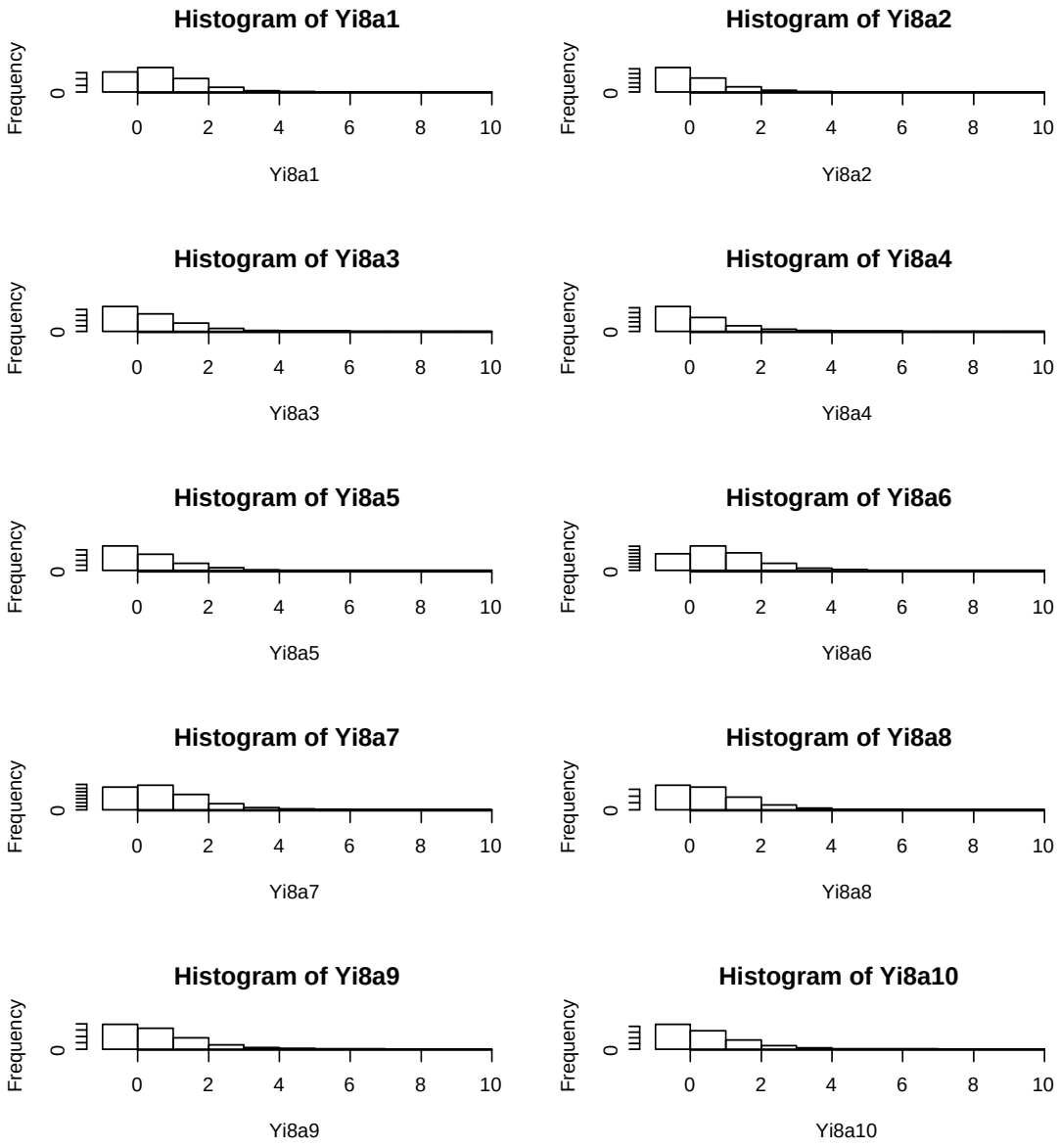

Figure 2. Posterior distributions of Y_{ij}'s $(j = 8,\ i = 1, \ldots, 10)$ for interval 8.

A comparison of the posterior means, medians and modes with the actual values are illustrated in Table 1 for two different time intervals. As we can see from the table, the posterior point estimates are close to the actual values of Y_{ij}'s in most cases. Similar results were obtained for other time intervals.

Interval 2	Ad 1	Ad 2	Ad 3	Ad 4	Ad 5	Ad 6	Ad 7	Ad 8	Ad 9	Ad 10
Mean	1.48	1.92	1.68	1.98	2.10	1.95	2.06	2.07	1.89	1.86
Median	1	2	1	2	2	2	2	2	1	1
Mode	0	1	0	1	1	1	1	1	0	0
Actual Value	0	2	2	2	1	0	2	5	2	3
Interval 8	Ad 1	Ad 2	Ad 3	Ad 4	Ad 5	Ad 6	Ad 7	Ad 8	Ad 9	Ad 10
Mean	1.14	0.71	0.86	0.73	0.83	1.40	1.22	1.11	1.04	0.95
Median	1	0	1	0	1	1	1	1	1	1
Mode	1	0	0	0	0	1	1	0	0	0
Actual Value	1	0	0	1	0	2	1	2	1	2

Table 1. Comparison of posterior inferences with actual Y_{ij}'s for intervals 2 and 8.

4.2 Example using Actual Call Arrival Data

We also applied our model to some actual call center arrival data similar to what is used in Soyer and Tarimcilar (2008) to see how the approach works with large number of advertisements over a long period of time. Data is on weekly call arrivals generated as response to advertisements appeared in print media and includes information on the cost of advertisement as well as type of promotion being offered in the ad. The data is available for 84 advertisements with different starting times, T_i's, over a period of 72 weeks. Thus, at a given week there can be large number of advertisements that are available and potentially generating calls. The actual data has 100 % linkage between the ads and the calls. In order to see the performance of our approach, we assumed that in each interval the ad sources were not known for some of the calls and we simulated the u_j's and Y_{ij}'s as we did in Section 4.1.

An example of the data is shown in Table 2 for some of the weeks between 1 and 25. We note that during the first week there are only 2 active ads and there was only 1 call out of 24 which can not be linked to one of the two active ads. Based on the N_{ij} column, there were 7 calls for ad 1 and 17 calls for ad 2. Similarly, during week 5 there

were 7 active ads and 3 calls could not be associated with any of these. As expected as the ad gets *older* it starts generating less calls. For example, we see that during week 25 no calls were received for ads 1, 2 and 3 which are the oldest among the 18 that are active at that time.

Week j	Active Ads	u_j	N_{ij}
1	2	1	$(7,\ 17)$
2	3	0	$(2, 14, 5)$
5	7	3	$(1, 1, 1, 12, 18, 8, 16)$
10	13	5	$(0, 3,\ 0, \ldots\ldots, 3)$
25	18	3	$(0, 0,\ 0, \ldots, 7,\ 24, 3)$

Table 2. Example of data with unknown call sources.

It is important to note that as the number of active ads gets larger over time the dimension of the multinomial distribution of $\underline{Y}_j$ gets larger. For example, in week 2 we had only 3 active ads whereas this number was 18 on week 25. Thus, in week 25 the 3 calls with unknown sources could be in reponse to any of these active 18 ads.

The data includes information on the cost of each advertisement (in $000) as well as the offer type with three categories: standard, interest bearing installment and free originating and return shipment. The three offer types and these are captured by two dummy variables in the model. More specifically, we used the log of the cumulative intensity function

$$log[\Lambda_i(t)] = \theta + \alpha logt + \beta_1 Z_{1i} + \beta_2 Z_{2i} + \beta_3 Z_{3i} \tag{4.1}$$

where $\theta = log\gamma$ in (3.2), Z_{1i} is cost of the advertisement, Z_{3i} and Z_{4i} are the dummy variables representing the second and third order types. We assumed diffused but proper priors for all parameters. We used the same parameter values for the Dirichlet prior on $\underline{p}_j$'s, for the normal prior on ad cost coefficient β_1, and for the gamma priors on α and γ as in Section 4.1. Priors for β_2 and β_3 were also assumed to be independent with normal distributions with mean 0 and variance 100.

A reasonable way to assess the performance of the proposed Bayesian approach of Section 3 is to compare the posterior and predictive inferences from the model with unknown call sources with those results from the complete model where all the call sources are known. In Table 3, we present a comparison of the posterior means and standard deviations of the unknown parameters from the latent source model and complete source model. We note that in all cases, the posterior inferences are very similar under the latent and complete source models. This suggests that the proposed approach for modeling latent sources performs reasonably well in this case.

	Latent Source Model		**Complete Source Model**	
	Mean	Std	Mean	Std
α	0.4156	0.0074	0.4015	0.0064
θ	2.1928	0.0425	2.0910	0.0337
β_1	0.0303	0.0061	0.0289	0.0045
β_2	0.1945	0.0498	0.1931	0.0345
β_3	0.6350	0.0405	0.7781	0.0387

Table 3: Comparison of Posterior Means and Standard Deviations

In Figure 3, we present the density plots for the posterior distributions of α, β_1, β_2, and β_3 from the latent source model. Note that the posterior α shows that the call generating ability of the advertisements deteriorates by time. As expected, the cost as well as the offer types have positive effect on the call intensity as implied by the support for positive values in the posterior distributions of these coefficients.

As previously discussed in Section 4.1, a more direct assessment of the performance of the Bayesian approach for modeling latent sources can be made by looking at the posterior distributions of Y_{ij}'s and comparing them to the actual values which are known to us.

In Table 4 we show a comparison of the posterior means and the actual values Y_{ij}'s for week 5 where we had 7 active ads and 60 calls 3 of which have latent sources.

We see from the table that Y_{ij}'s for all the ads except the 4th one predicted (or classified) correctly. Again we note that as the number of time intervals (weeks) increase the dimension of the Y_{ij}'s increase and those ads which have been active for a longer time become less likely sources of unidentified calls. Obviously this is due to the fact that the call intensity decreases by time as α takes values less than 1. The table also presents the posterior means of p_{ij}'s for the interval. We note that the prior means for p_{ij}'s, which were chosen as $1/7 = 0.143$ for each ad in the interval, now are revised to posterior mean values shown in the last column. Based on these, the expected posterior probability that a latent source call is generated by any of the first three ads is quite small. On the other hand, the expected posterior probability is lot higher for ads 5, 6 and 7.

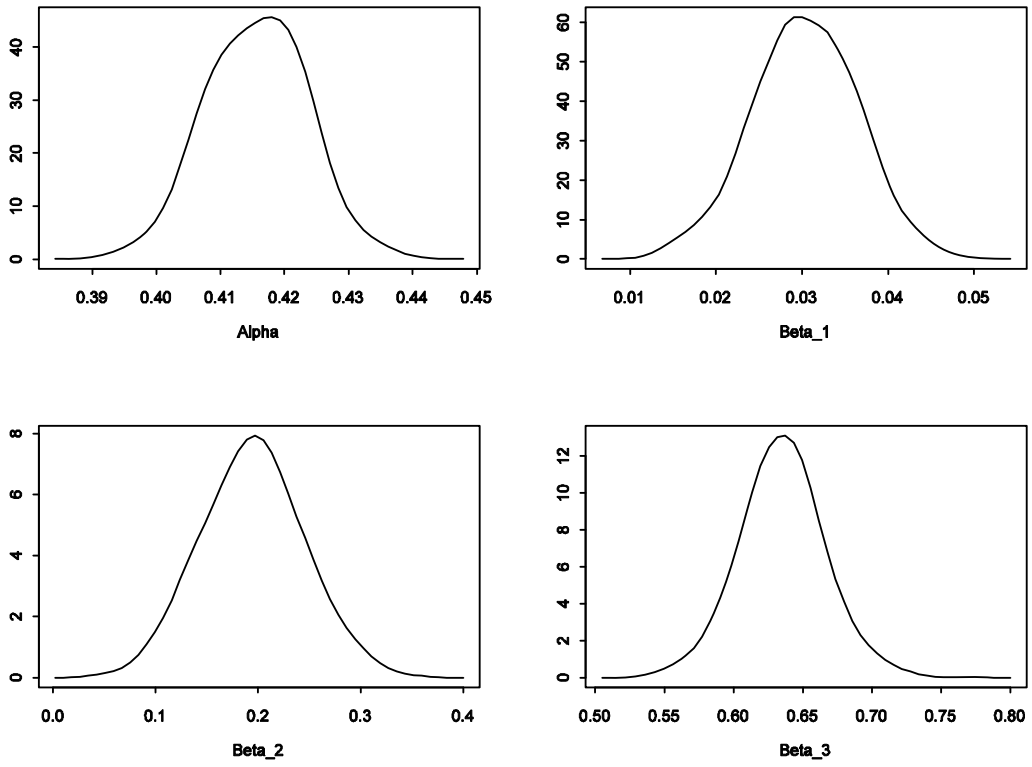

Figure 3. Posterior distributions of parameters in the latent source model.

Ad i	Actual Y_{ij}	$E[Y_{ij}]$	$E[p_{ij}]$
1	0	0.05	0.018
2	0	0.07	0.022
3	0	0.04	0.014
4	0	0.56	0.186
5	1	0.79	0.266
6	1	0.77	0.256
7	1	0.71	0.238

Table 4. Comparison of actual Y_{ij}'s with posterior means for week 5.

Similar conclusions can be reached by looking at the posterior distributions of Y_{ij}'s shown in Table 5. We can see from the table that for ads 1-3, the posterior probability values of the actual Y_{ij} values are very high. For ads 4, 5 and 7 the mode of the posterior distribution is at the actual value. But in these cases the variance is higher than the cases of ads 1-3. Note that we have specified diffused priors for p_{ij}'s which do not take into account the age of the ad. We clearly see that the posterior inferences on p_{ij}'s suggest that we learn about the latent sources from the data. Thus, these results are quite satisfactory. If we use informative priors such as in (3.3), the accuracy will naturally improve. We obtained similar results for the other time intervals.

Ad i	Actual Y_{ij}	$P(Y_{ij}=0)$	$P(Y_{ij}=1)$	$P(Y_{ij}=2)$	$P(Y_{ij}=3)$
1	0	0.946	0.053	0.001	0.000
2	0	0.934	0.066	0.000	0.000
3	0	0.958	0.041	0.001	0.000
4	0	0.544	0.361	0.086	0.009
5	1	0.377	0.461	0.151	0.011
6	1	0.422	0.406	0.154	0.018
7	1	0.420	0.447	0.132	0.001

Table 5. Posterior distributions of Y_{ij}'s for week 5.

5. Concluding Remarks

In conclusion, our experience with the proposed Bayesian approach for modeling latent sources and the corresponding data augmentation algorithm within the Gibbs

sampler have shown lot of promise. The proposed approach provided very close posterior inference results for the model parameters and actual arrivals when it is compared with complete source model results. The inferences about Y_{ij}'s when compared to actual values were found to be reasonably close.

References

Aksin, O.Z., Armony, M. and Mehrotra, V. (2007). The Modern Call-Center: A Multi-Disciplinary Perspective on Operations Management Research. *Production and Operations Management*, Vol. 16, pp. 665-688.

Andrews, B. H. and Cunningham, S. M. (1995). L. L. bean Improves Call-Center Forecasting. *Interfaces*, Vol. 25, pp. 1-13.

Avramidis, A. N., Deslauriers, A. and L'Ecuyer, P. (2004). Modeling Daily Arrivals to a Telephone Call Center. *Management Science*, Vol. 50, pp. 896-908.

Chib, S. and Greenberg, E. (1995). Understanding the Metropolis-Hastings algorithm, *The American Statistician*, 49, 327-335.

Cox, D. R. (1972). The Statistical Analysis of Dependencies in Point Processes. In *Stochastic Point Processes*. Ed. P. A. W. Lewis, pp. 55-66, New York Wiley.

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

Gilks, W. and Wild, P. (1992). Adaptive Rejection Sampling for Gibbs Sampling. *Journal of the Royal Statistical Society, Ser. B*, Vol. 41, pp. 337-348.

Jongbloed, G. and Koole, G. (2001). Managing Uncertainty in Call Centers using Poisson Mixtures. *Applied Stochastic Models in Business and Industry*, Vol. 17, pp. 307-318.

Soyer, R. and Tarimcilar, M. M. (2008). Modeling and Analysis of Call Center Arrival Data: A Bayesian Approach, *Management Science*, Vol. 54, pp. 266-278.

Weinberg. J., L. D. Brown, J. R. Stroud. (2007). Bayesian forecasting of an inhomogeneous Poisson process with applications to call center data. *Journal of the American Statistical Association*, Vol. 102, pp. 1185–1199.