nTopic Organic Traffic Study

1 Abstract

The objective of this study is to determine whether content optimization solely driven by nTopic recommendations impacts organic search traffic from Google. The study used 3 randomly applied treatments of nTopic modified content, random keyword modified content, and unmodified content. nTopic modified content saw an organic traffic lift of 17.5%, random keyword modified content saw a 10% drop and unmodified content saw drop in traffic of 15%. This indicates that nTopic optimized content can create increases in organic search traffic that cannot be explained solely by content freshness measurements or random fluctuations.

2 Experiment Methodology

We first describe the process in which we collected the data for our study.

1. Randomly select 200 pages from a single website: we use a single website to prevent domain authority from influencing the effectiveness of content modifications.
2. Identify top keyword referring traffic to each of 200 selected pages
3. Randomly apply treatments to each of the 200 pages
    - **Treatment 1**: Insert 50 recommended words from nTopic Keyword Recommendations for Keyword and Page Content via Paid API
    - **Treatment 2**: Insert 50 random words from dictionary to page
    - **Treatment 3**: Make no modification to the content
4. Record 3 weeks worth of organic google traffic for each keyword prior to transformation
5. Record 3 weeks worth of organic google traffic beginning 1 month after the launch of transformation
6. Record shift in organic Google traffic for each page
7. Record average shift for each transformation method

Figure 1 shows some summary plots of the experiment data. While the empirical distributions of the treatment traffic shifts seem to indicate nTopic performs better on average, there is much higher variability in the less visited sites which is clear in the scatter plot. Therefore, simply calculating the percent shift for each page and analyzing the empirical distributions is not sufficient.

Properly assessing the impact of the treatments requires a more principled analysis of the data. We give the details of this analysis in the next section followed by a discussion of the results and model validation considerations. It should be noted that even comparing these noisy percent lift histograms using the non-parametric Mann and Whitney (1947) U test shows the nTopic to be significantly more effective than the other two approaches.

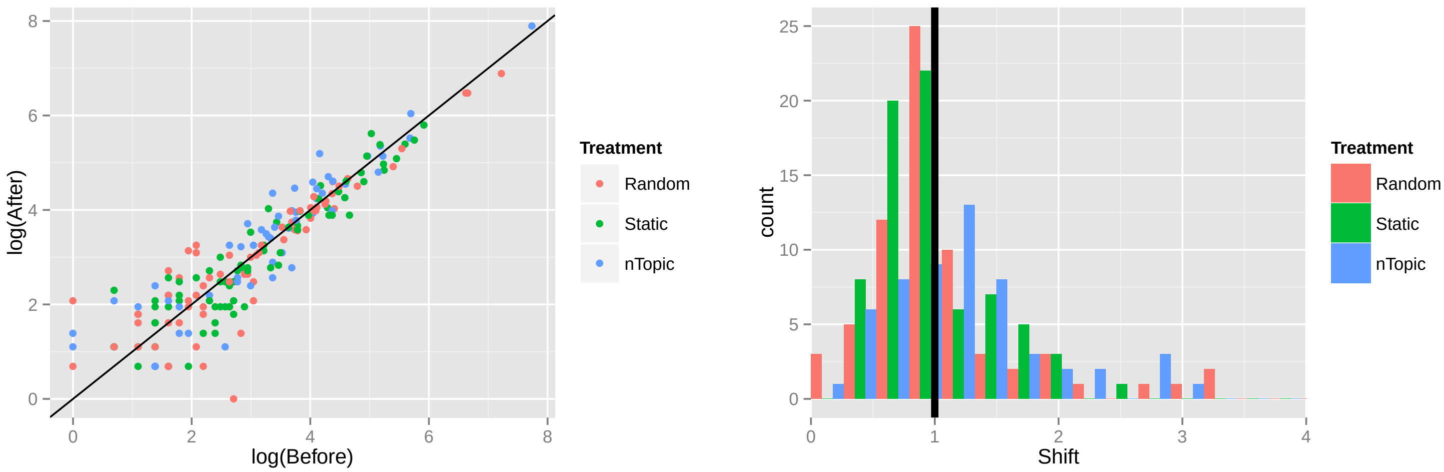

Figure 1: Left: Scatter plot of the amount of organic search traffic before and after on the log scale the treatments. Right: Histograms of observed percent increase (or decrease) in organic search traffic for each treatment.

3 Poisson Random Effects Model for Search Traffic

First, we assume that visits to a particular page follow a Poisson process with some rate λ. This can be characterized by the time between visits to the website being independent and exponentially distributed with rate λ where one then counts the cumulative events. This is a standard assumption in server traffic modeling and many other applications where we count the number of events up to some time (Arlitt and Williamson, 1997). A consequence of this assumption is that the number of events (visits to the website) over any period of time will be a Poisson distribution. Furthermore, the number of events over any non-overlapping periods of the same length will have independent Poisson distributions with the same mean.

Let each page be indexed by i and the number of visits to that page in the first 3 week period be denoted x_i. We denote the number of visits to the same page during the second 3 week period as y_i. Since these pages are all the same website, it is reasonable to assume that the λ_i are each drawn from some population of visit rates. We assume based upon the discussion above that for each i

$$x_i \sim \text{Poisson}(\lambda_i) \text{ where } \log(\lambda_i) \sim \text{N}(\mu, \sigma^2).$$

This is known as a generalized random effects model for Poisson data where there distribution of means is a log normal distribution. We recommend the work by Chib and Carlin (1999) for more information on these types of models.

Since our inferential goal is to understand the effect of each treatment– indexed by k– on the page arrival rate, we assume that the arrival rate for each page is multiplied by a common factor β_k. Specifically, we assume that if page i had treatment k,

$$y_i \sim \text{Poisson}(\beta_k \lambda_i).$$

We now have a complete model specification for our data with the goal of understanding the difference between β_1, β_2, and β_3 which correspond to the multiplicative effects of the random keywords, static, and nTopic treatments respectively.

4 Results

We used the JAGS software to perform model analysis in a Bayesian setting on the parameters $\boldsymbol{\beta}$, μ, and σ. This approach is extremely effective for fitting custom, flexible models and for performing model validation as demonstrated in the next section (Plummer, 2003). Since this is a Bayesian approach, we need to specify our prior understanding of the variables of interest with probability distributions. Our goal is objective inferences, so we chose priors that represent having no prior information about the parameters.

JAGS allows us to take draws from the posterior distribution which represents our understanding of the parameters in our model after we observe the data. We ran the JAGS Gibbs sampler for 10,000 adaptive iterations, then for an additional 10,000 burn-in iterations; we then stored the final 10,000 draws from the posterior distribution to do inference on $\boldsymbol{\beta}$. Figure 2 shows the posterior summaries. The first clear insight is that the posterior distribution for the **nTopic shift is clearly and significantly larger than the other two effects.** Furthermore, the random keywords and static treatments are quite similar to each other and both are below 1 indicating those treatments actually reduce the rate in site visits.

5 Model Validation

There are two major assumptions that must be validated in this analysis. First, the choice of random effects distribution could play a major role. However, the JAGS software allowed us to quickly check many different random effects distributions. We then considered a semi-parametric finite mixture of distributions which can approximate a wide class of distributions (Fraley and Raftery, 2002). In all of these cases, the results were the same. Furthermore, the log normal assumption is quite reasonable because its heavy tailed features accurately depict the fact that most visiting rates will be around some common region with a few extreme outliers.

The second is that within a given treatment, we assume the new visiting rate is some multiplicative shift of the old. One way to test this assumption is to perform a second analysis where we first remove random data from the second period in each treatment and treat it as missing. In the Bayesian framework, those missing observations are treated as parameters to estimate and their estimates are called posterior predictive distributions. Figure 3 shows boxplots for the posterior predictive distributions for 5 observations removed randomly from each treatment. We see that the held out data is within the estimated range, and hence our model does not appear to have any systematic bias as a predictive tool. While this is not a conclusive test of model correctness, it is strong evidence that the results are valid.

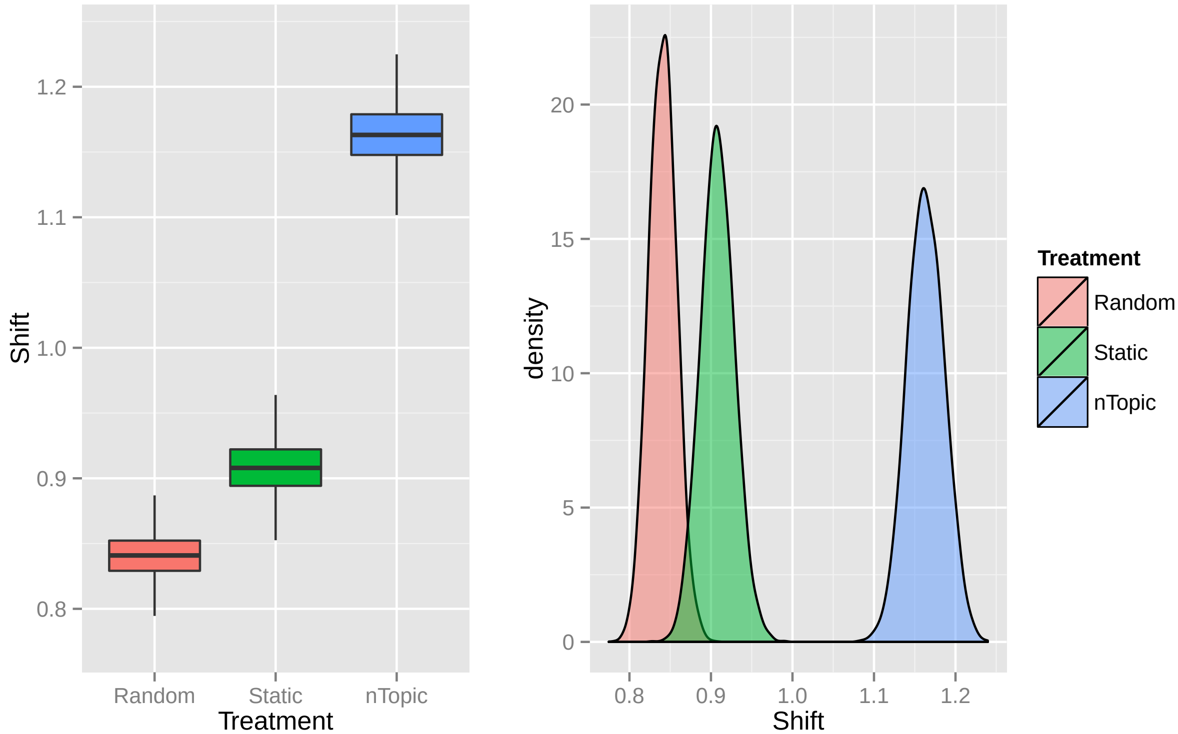

Figure 2: Posterior summaries of the shifts β_1, β_2, and β_3 corresponding to random, static, and nTopic treatments respectively. On the left, boxplots of the posterior parameter distributions are presented. On the right, kernel smoothed density estimates of the posterior are shown.

6 Conclusions

We have shown strong evidence that insertion of nTopic recommended keywords can increase organic search traffic. This increase is significantly better than a static webpage or random fresh content. It should be noted that

- We **cannot** conclude that topic modeling or content relevancy is used in Google's algorithm.
- We **cannot** determine the exact mechanism by which nTopic recommended keywords increase Google traffic.
- We **can**, however, provide evidence that insertion of nTopic recommended keywords increase Google traffic.

References

Arlitt, M. F. and Williamson, C. L. (1997). Internet web servers: workload characterization and performance implications. *IEEE/ACM Trans. Netw.*, 5(5):631–645.

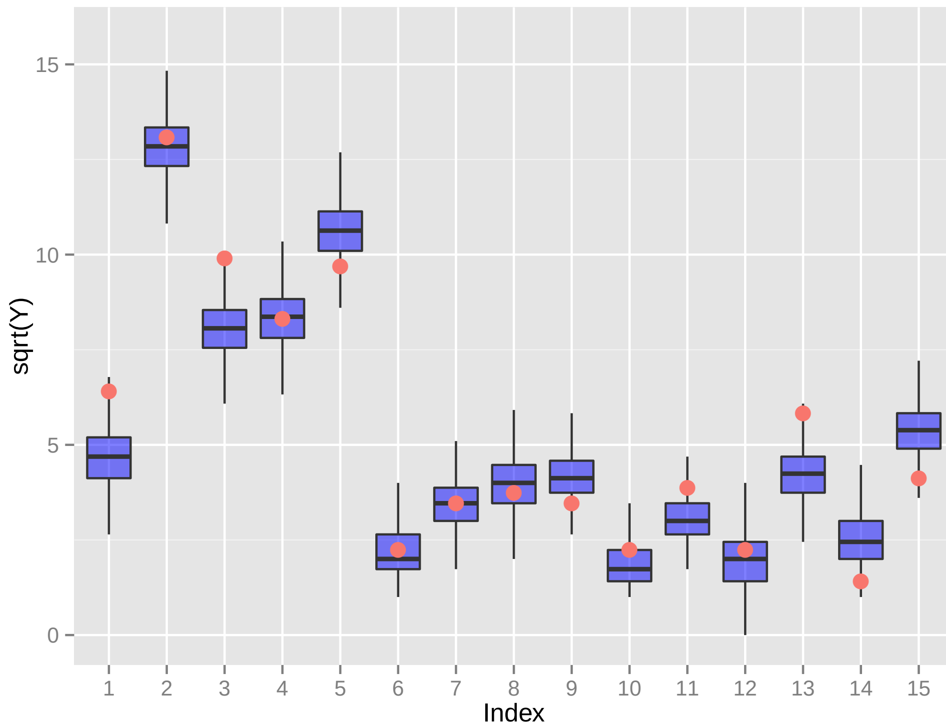

Figure 3: Posterior predictive estimates for the excluded after treatment data in the validation phase with the actual observed rates shown in red.

Chib, S. and Carlin, B. P. (1999). On mcmc sampling in hierarchical longitudinal models. *Statistics and Computing*, 9:17–26.

Fraley, C. and Raftery, A. E. (2002). Model-based clustering, discriminant analysis, and density estimation. *Journal of the American Statistical Association*, 97(458):611–631.

Mann, H. and Whitney, D. (1947). On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Stat.*, 18:50–60.

Plummer, M. (2003). Jags: A program for analysis of bayesian graphical models using gibbs sampling.