Analyzing Tablet Usage Share During the December 2012 Holiday Season

Chitika, Inc.

February 6, 2013

Abstract

This paper explains the analysis procedure used by Chitika Insights in producing our post-Christmas market share report, using tablet share as a case study. During the post-Christmas time period, we observed marked changes in Apple iPad® and Amazon Kindle Fire® traffic share. We utilized a Holt-Winters forecasting model to understand the impact of these changes quantitatively, as well as inform our prediction of subsequent changes in the marketplace. Our predictions were validated by the market share values we observed during the month of January 2013.

1 Introduction

Early in January, Chitika released market share numbers based on traffic patterns observed in our advertising network during December 2012. At the time, we were interested in reporting on the changes in the network during the several days after Christmas in particular, based on our hypothesis that trends during the time period would provide insight into consumer behavior during the holiday season.

During the period extending from Christmas Day to December 27, inclusive, we observed a larger than expected increase in the traffic share for the Amazon Kindle Fire®, and a larger than expected decrease in the traffic share for the Apple iPad® (all versions). Although these results were in part larger than expected due to bias inherent in looking at change across only three days, we nevertheless only expected a return to about 80% market share for the iPad. The strong performance of the Kindle Fire is perhaps the most notable aspect of our original report.

In this paper we discuss our use of the Holt-Winters forecasting method, also known as triple exponential smoothing, to model the consumer tablet market. This model also formed the basis for our prediction of longer-term market share rates.

The data we collected for January 2013 confirmed our predictions for the tablet market.

2 Data Sources

Chitika is an advertising network that serves billions of online advertisements globally per month. Chitika collects data from user agent information provided by Web browsers when we serve an ad impression. We maintain a database internally that is used to identify devices as accurately as possible based on data contained in the user agent. This database is constantly being refined as new devices and software are released.

For the market share report for the post-Christmas time period, the data drew from the time period of December 1 to December 27, since we were interested in quantifying changes in the Chitika network on Christmas and the two days following in comparison with December 1 through Christmas Eve.

For the present paper, we consider data collected from ads displayed in the United States and Canada over the time period of December 1 to January 27.

3 Data Summary

A summary of the traffic share data from the post-Christmas tablet device report is shown below:

Device	Share: December 1-24 (%)	Share: December 25-27 (%)	Difference (%)
Amazon Kindle Fire	4.48	7.51	3.03
Samsung Galaxy Tablet	3.01	4.39	1.38
Google Nexus	1.12	2.04	0.92
Microsoft Surface	0.23	0.40	0.17
BlackBerry Playbook	0.70	0.68	-0.02
Apple iPad	86.0	78.86	-7.14

The above table gives the values in market share changes that we published in the post-Christmas market share report. Note, however, that it does not describe absolute changes in traffic, but only relative share.

4 iPad Forecast Model

In order to better understand the meaning of the data observed in the Chitika network during the post-Christmas period, we build a model of tablet market share change using the Holt-Winters method. We provide a brief overview of the Holt-Winters method below.

4.1 The Holt-Winters Method

The Holt-Winters method for time-series forecasting is an extension of the simple exponentially-weighted moving average [2, 3]. The exponentially-weighted moving average (EWMA) applies a smoothed statistic to a time series. The form of the model is similar to a simple moving average, which constructs a smoothed statistic by taking a fixed subset of the time series into account at each point. The EWMA, however, weights past data so that the most recent data has a greater effect on the moving average. The simple exponential smoothing equation can be written as

$$s_{i+1} = \alpha x_i + (1 - \alpha)s_i, \tag{1}$$

where $\alpha (0 < \alpha < 1)$ is a smoothing constant defining the weight of the previous data point and x_i is the actual data at a time interval i. Because of the recursive relationship between

s_{i+1} and s_i in equation 1, we can also write s_{i+1} as

$$s_{i+1} = \alpha x_i + \alpha(1-\alpha)x_{i-1} + \alpha(1-\alpha)^2 x_{i-2} + \cdots . \tag{2}$$

Equation 2 makes it clear that exponential smoothing gives the most weight to the most recent observation x_i and less weight to earlier observations. That is, the exponential smoothing approach is based on the idea that the parameters of the stochastic process generating the time series may be changing over time, and so the most recent data points provide the most reliable estimate of the current value of the parameters.

The exponentially-weighted moving average by itself does not, however, do a good job of handling trend in the time series. In order to model the trend of the time series, if present, we can use a second expontential smoothing to model the influence of a trend factor on the signal, which we modeled as s_i above. That is, we can describe the unsmoothed value of the trend as the difference between the current and previous unsmoothed smoothed signal,

$$t_i = \beta(s_i - s_{i-1}) + (1-\beta)t_{i-1}. \tag{3}$$

Thus the trend tells us at any point how much the smoothed signal changed in the previous time interval.

In this case, the smoothed signal itself can be described in a similar manner to the simple exponential smoothing in equation 1, but now we also consider the trend,

$$s_i = \alpha x_i + (1-\alpha)(s_{i-1} + t_{i-1}). \tag{4}$$

To construct a forecast from this result, we take the most recent smoothed value, s_i, and add the last smoothed trend for each additional time step h; that is,

$$x_{i+h} = s_i + h t_i. \tag{5}$$

The variance of the prediction error e_i at a point s_i can be estimated for double exponential smoothing by

$$\mathrm{Var}(e_i(k)) = \mathrm{Var}(e_i)\left[1 + \sum_{j=1}^{k-1}(2\alpha + (j-1)\alpha^2)^2\right], \tag{6}$$

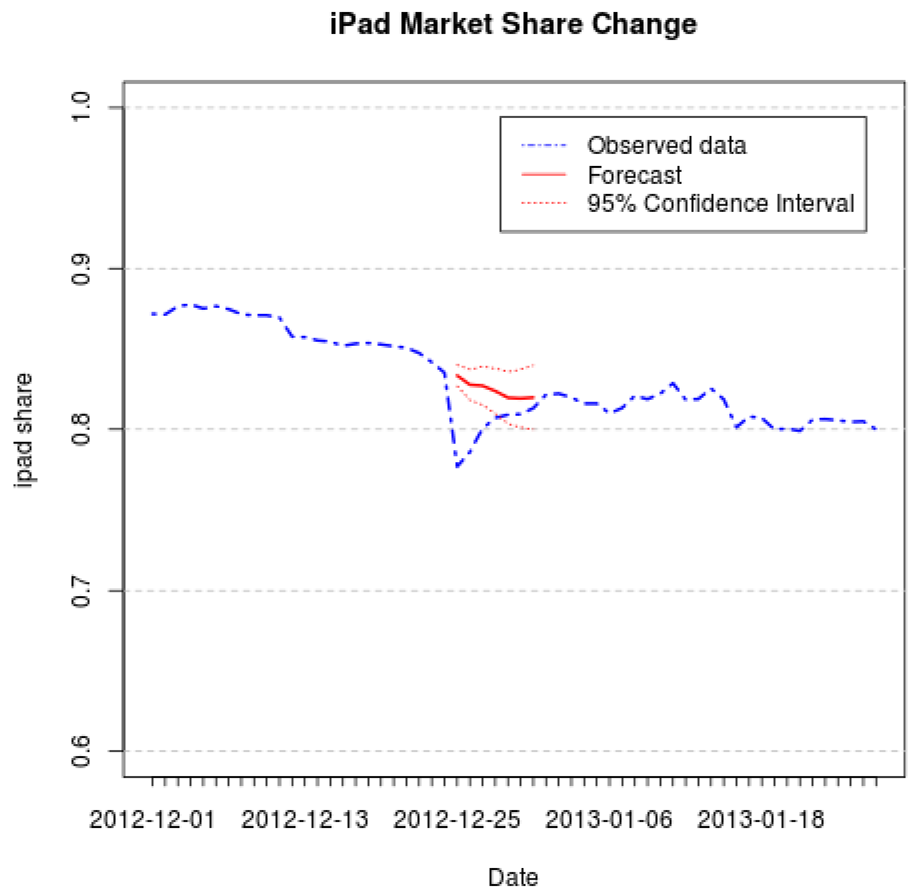
which allows us to compute a confidence interval [1].

To describe the seasonality in the time series, we model yet a third quantity with exponential smoothing. For the case of an additive seasonality, which we used in the current study, the line of reasoning proceeds in a similar fashion to the above. For details on calculating the variance of the prediction errors for the Holt-Winters model with additive seasonality, see [1].

4.2 Tablets Forecast

Chitika applied a Holt-Winters forecast model to the December tablet share data in order to understand what a reasonable prediction of short-term changes in the tablet market might look like, given the information we had from December 1-24. Model-fitting was conducted

Figure 1: Change in iPad Market Share



in `R`, using a triple exponential smoothing method, with an additive model for seasonality. The seven day forecast for iPad share is shown as the solid line in figure 1 below, along with the 95% confidence intervals for the forecast.

Note that figure 1 suppresses the zero point on the y-axis in order to make the trend changes more explicit.

5 Analysis

The Holt-Winters model for iPad share change predicts a declining market share for the iPad, continuing the trend observed in the December data through Christmas Eve. The observed data was well outside the confidence intervals of the forecast, leading us to conclude that the dip observed in iPad market share after Christmas was unlikely to be due to chance.

Since the model suggested that the changes we observed in the days following Christmas were non-random, we investigated the underlying data from the Chitika network to see if the dip could be accounted for by changes in the network itself. Our investigation did not yield evidence of anomalous events in the network.

6 Conclusion

Our analysis of the tablet market share in the post-Christmas time frame was ultimately confirmed by the traffic data observed in the Chitika network over the course of January 2013. This example provides justification for the forecasting approaches used at Chitika to understand shifts in the online world.

In addition, initial reports that Apple sales may have been lower than expected during the holiday further suggests that the dip observed during the three days post-Christmas was an accurate description of the marketplace during that time (see for example `http://bgr.com/2013/01/18/ipad-sales-weak-q4-2012-294090/`).

The case study above gives an overview of how Chitika uses short-term forecast models to understand the significance of changes in our network data, using the specific example of iPad market share based on traffic in our network and the Holt-Winters method for forecasting. Although Holt-Winters is not our exclusive methodology for the analysis of time series data—we do not, for example, discuss ARIMA or regression models here—this paper is meant to provide information on data analysis techniques used by the Chitika Insights team in more detail.

References

[1] B.C. Archibald. Parameter space of the Holt-Winters' model. *International Journal of Forecasting*, 6:199–209, 1990.

[2] Charles C. Holt. Forecasting seasonals and trends by exponentially weighted moving averages. *Office of Naval Research Memorandum*, 52, 1957.

[3] P.R. Winters. Forecasting sales by exponentially weighted moving averages. *Management Science*, 6:324–342, 1960.