Evaluation of effectiveness of an advertisement campaign by use of Twitter

Vasiliy Kuznetsov,[a] Aleksey Semyonov[b]

[a] WIUT—Westminster International University in Tashkent, 12 Istiqbol St., Tashkent, 100047, Uzbekistan

vkuznetsov@wiut.uz

[b] WIUT—Westminster International University in Tashkent, 12 Istiqbol St., Tashkent, 100047, Uzbekistan

asemyonov@wiut.uz

Abstract. The paper addresses the topic of marketing campaign evaluation by means of statistical analysis of publicly available data from popular micro-blog service Twitter. The research is aimed at finding out the parameter or combination of parameters e.g. average frequency, peek frequency, total number of posts, number of positive/negative posts etc. of tweets mentioning the product/company of interest providing best correlation with sales. In particular current paper addresses movie industry by correlating first weekend box office revenue of all movies premiered in the period from September 2012 till end of March 2013 in USA with the amount of tweets mentioning the movie titles during Monday to Sunday in the week of first night of a movie. Moreover analysis is given for "positive tweets" and "negative tweets" search options of Twitter and their effect on the prediction. Additionally the paper outlines major general issues associated with data collection and analysis using Twitter.

Keywords: data mining, Twitter, social networks, advertisement campaign evaluation.

1 INTRODUCTION

Today many businesses are highly dependent on proper marketing mix in order to promote their products, increase sales and survive in highly competitive world. Modern advertisement campaigns are generally expensive and require utilization of various tools for monitoring their effectiveness. This paper is aimed at proposing a feasible tool for advertising campaign effectiveness evaluation by means of analyzing data from popular micro-blog Twitter.

2 EXISTING METHODS OF ADVERTISEMENT CAMPAIGN EVALUATION

Nowadays marketing specialists employ quite a number of tools for evaluating the effectiveness of elements of marketing mix. Such tools can be categorized as following (N.D.2013):

- Direct measures – methods based on the retrospective analysis of changes in sales:
    - Historical Sales Method – a method based on multiple regression analysis between the expenditures on the advertisement and sales over corresponding periods.
    - Experimental Control – a method based on comparing normal level of sales in a number of places (cities) with after-campaign figures and the level of sales in so-called "test cities", not affected by the campaign.
- Indirect measures – methods based on interviews/questionnaires aimed to identify to which extent potential customers perceived the ad message:

- Exposure to Advertisement – interview based approach based on determining the number of persons received (heard, watched) the ad message
- Attention or Recall of Advertising Message Content – lab-based evaluation of how good affected people can recall the message/content of an advertisement.
- Brand Awareness – a method comparing the level of awareness of the brand or product among population affected by the campaign.
- Comprehension – a method based on interviewing potential clients with the aim to evaluate how well they comprehended the message from an advertisement.
- Attitude Change – interview- or questionnaire-based method evaluating the change of attitude towards the product.
- Action – another interview-based method questioning people on how ready they are to make a purchase based on the advertisement message they hear.

Moreover the topic of evaluation of advertisement campaign effectiveness was well reflected by various researches. Drossos et al. (2007) studied the effectiveness of SMS advertisement campaigns and described a set of possible determinants of successful mobile marketing communication, while Robinson, Wysocka and Hand (2007) studied the effectiveness of banner ads in the Internet. Pavlou and Stewart (2000) analyzed criteria for evaluation of effectiveness of interactive advertisement. Various authors like Kinnucan and Forker (1986), Kinnucan, Chang, and Venkateswaran (1993), Reberte et al. (1996) studied effectiveness of marketing means for particular products and areas, in this case fluid milk market in USA. On the other hand authors like Ward and Myers (1979) and Kinnucan and Venkateswaran (1994) analyzed broader markets and introduced time dimension and seasonal variations as the parameters for such evaluation. Finally other researches e.g. Lutz, MacKenzie and Belch (1983) and Mehta (2000) were more inclined towards psychological and behavioral approaches and reviewed attitude toward the ad as a mediator of advertising effectiveness.

3 PROPOSED SOLUTION

It can be stated that the major problems of existing evaluation tools are backward-orientation of sales-based methods and expensiveness and subjectivity of indirect evaluation methods. Therefore it was decided to propose an alternative method that should be able to provide a company with on-time feedback on how effective an advertisement campaign is and at the same time be as reliable and feasible as possible. Such method could be based on frequency of mentioning of a particular brand/product in social electronic media channels such as Facebook, Twitter, various blogs and user review providers.

3.1 Methodology

For the current research correlation between the frequency of mentioning a particular movie title in micro-blog Twitter and the resulting box-office revenue at the first weekend of showing was utilized. Movie industry was mainly selected due to limited advertisement campaign period (usually active campaign starts at the week of premiere as was mentioned by Asur and Huberman (2010)) and availability of precise revenue data.

The dataset collected is constituted by tweets mentioning a title of any movie premiered in the period from September 2012 till end of March 2013 in USA. For each particular movie title all tweets were collected for the period of Monday to Sunday of the week of the premiere. Tweet collection was done by the help of standard Twitter API with the title of a movie used as the search term. The search was done three times for each title in order to collect so-called "attitudes" – the special feature of Twitter allowing getting neutral (no additional flags), positive (':)' flag appended to query) and positive (':(' flag appended to query) tweets. All movies premiered in the period were included into analysis to consider both very popular and out of favor titles. Overall for the period it was collected over 5 millions tweets for 195 movie titles.

4 ANALYSIS

The data contains information about the movie weekend revenues in the first week of release as well as counts on number of tweets related to the movies one week prior to the premier. Tweets are categorized by weekdays (Monday, Tuesday, Wednesday, Thursday, Friday, Saturday, Sunday) and by attitude (neutral, positive, negative). Since the data for box-office revenue was obtained for US only it was decided to filter the tweets collected correspondingly. Unfortunately only 4% of the tweets collected had geo-tags included therefore making it impossible to reliably filter the data based on the location. As a result it was decided to filter by the language (IsoLanguageCode field) and exclude non-English tweets from the dataset used for final analysis. After some sanity checks, removing duplicates and filtering by language the total number of tweets remaining totaled in 3.4 millions.

The study applies the Ordinary Least Squares method in order to identify, which category of tweets can be used as a best predictor for total first weekend revenue. The following types of models are investigated:

Table 1. Empirical Model

Predictor variables	**Model** **(outcome variable R – total weekend revenue)**
dayp = number of positive tweets on a particular day dayn = number of neutral tweets on a particular day	$R = \beta_1 \times dayp + \beta_2 \times dayn$

4.1 Results

The results of the analysis show that the empirical specification chosen well explains the behavior of the responsiveness of the total weekend revenue to the changes in the number of positive and neutral Saturday tweets. Adjusted R squared being 0.72 shows that two dependent variables being "Number of positive Saturday tweets" and "Number of neutral Saturday tweets" account for 72% of variance.

The coefficient of "Neutral Saturday Tweets" being significant at 1 percent level of significance shows that on average an increase in one tweet increases the total revenue by 851.29 dollars, while the coefficient of "Positive Saturday Tweets" being significant at 1 percent level of significance shows that on average an increase in one tweet increases the total revenue by 7967.85 dollars.

Table 2. Determinants of Total Weekend Revenue from tweeter (Dependent variable: Total revenue)

	Coefficient beta
Saturday Tweet Neutral	851.29*** (97.22)
Saturday Tweets Positive	7967.85*** (475.33)

Note: *, **, *** - significant at 10, 5, and 1 percent, respectively. The terms in parentheses indicate the standard errors.

Numbers of tweets made during the weekdays also significantly predict the "Total Weekend Revenue", but weaker model in terms of variance explanation are produced.

Table 3. Determinants of Total Weekend Revenue by day

Predictor variables	**Adjusted R squared**	**Beta coefficients**
Friday Tweets Neutral	0.45	1183**
Friday Tweets Positive		4456**
Thursday Tweets Neutral	0.667	1159**
Thursday Tweets Positive		10886**
Wednesday Tweets Neutral	0.591	645**
Wednesday Tweets Positive		18064**
Tuesday Tweets Neutral	0.583	1029**
Tuesday Tweets Positive		19887**
Monday Tweets Neutral	0.487	406*
Monday Tweets Positive		31990**
Total Weekday Tweets Neutral	0.610	313**
Total Weekday Tweets Positive		3030**

Note: *, **, *** - significant at 10, 5, and 1 percent, respectively.

4.2 Discussion of Results

The study found that number of positive and neutral tweets made during the week preceding the premier of the movies can be used as predictors of total first weekend revenue figures. The best model in terms of variance explanation is obtained using the numbers of tweets during the first day of the weekend. The model is:

$$R = 851.29 * a + 7967.85 * b \ , \qquad (1)$$

where R is the Revenue, a is the number of Saturday neutral tweets and b is the number of Saturday positive tweets.

It is also evident from the results that number of positive tweets has a far greater effect on total weekend revenue in the models investigated compared to number of neutral tweets. These results might seem logical as those potential spectators who rate the movie positively are more likely to see it, but it can bear some important implications for the promoters of the movie. Positive tweets might influence the opinion of other bloggers and potential spectators, as in all models increase in revenue with the increase in number of positive tweets by 1 is far greater than the price of one movie ticket. The highest rate of change obtained relates to number of Monday positive tweets, which indicates that an increase in one positive tweet of Monday increases the revenue by 31,990 US dollars. The same types of models applied for other weekdays show a gradual decline in corresponding beta coefficients. Such a result might mean that number of positive tweets earlier during the week influences potential spectators more than positive tweets made later during the week. From the psychological point of view it can be explained by a "Halo" effect, which continues to create and attract more spectators throughout the week.

5 FURTHER RESEARCH

Regarding the continuation of the current research it should be mentioned the following:

- Other sources should be analyzed in a similar manner - social networks, blogs, user reviews etc.
- Consider other areas/industries – as it was stated above movie industry was mainly selected due to limited duration of ad campaigns and availability of box office revenue data. Therefore the next step is to conduct similar research over other non-related areas e.g. retailing, luxury products, and products of international corporations to allow generalizing the applicability of the concept.
- More intelligent filtering of data – it was observed that some portion of the tweets collected were not directly related to the movie title considered, they simply contained the words of the query issues, but in absolutely different meaning. This was very significant for short titles with commonly used words (e.g. "Admission" or "The Call"). In such cases additional words were added to the query in order to filter unrelated results. However an approach based on natural language processing is preferable in order to be able to filter out such accidental records from the dataset.

6 REFERENCES

Asur, S., & Huberman, B. A. (2010, August). Predicting the future with social media. *In Web Intelligence and Intelligent Agent Technology (WI-IAT), 2010 IEEE/WIC/ACM International Conference on* (Vol. 1, pp. 492-499). IEEE.

Drossos, D., Giaglis, G. M., Lekakos, G., Kokkinaki, F., & Stavraki, M. G. (2007). Determinants of effective SMS advertising: an experimental study. *Journal of Interactive Advertising*, *7*(2), 16-27.

Kinnucan, H. W., & Venkateswaran, M. (1994). Generic Advertising; and the Structural Heterogeneity Hypothesis. *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie*, *42*(3), 381-396.

Kinnucan, H. W., Chang, H. S. C., & Venkateswaran, M. (1993). Generic advertising wearout. *Review of Marketing and Agricultural Economics*, *61*(03).

Kinnucan, H., & Forker, O. D. (1986). Seasonality in the consumer response to milk advertising with implications for milk promotion policy. *American Journal of Agricultural Economics*, *68*(3), 562-571.

Lutz, R. J., MacKenzie, S. B., & Belch, G. E. (1983). Attitude toward the ad as a mediator of advertising effectiveness: determinants and consequences. *Advances in consumer research*, *10*(1), 532-539.

Mehta, A. (2000). Advertising attitudes and advertising effectiveness. *Journal of Advertising Research*, *40*(3), 67-71.

N.D., (2013), Methods of measuring advertising effectiveness. *MBA knowledge base*. Retrieved November, 2012, from www.mbaknol.com

Pavlou, P. A., & Stewart, D. W. (2000). Measuring the effects and effectiveness of interactive advertising: A research agenda. *Journal of Interactive Advertising*, *1*(1), 62-78.

Reberte, C., Kaiser, H. M., Lenz, J. E., & Forker, O. (1996). Generic Advertising Wearout: The Case of the New York City Fluid Milk Campaign. *Journal of Agricultural and Resource Economics*, 199-209.

Robinson, H., Wysocka, A., & Hand, C. (2007). Internet advertising effectiveness. *International Journal of Advertising*, *26*(4), 527-541.

Ward, R. W., & Myers, L. H. (1979). Advertising effectiveness and coefficient variation over time. *Agricultural Economics Research*, 31(1), 1-11.

Vasiliy Kuznetsov is the Course Leader for BSc Business Information Systems for Westminster International University in Tashkent (WIUT). He received BS in Business Computing from WIUT and MBA from Moscow Business School. His research interests include data-mining and analysis of non-structured data.

Aleksey Semyonov: is a Head of Quality Enhancement in Westminster International University in Tashkent. He received his Mphil in Psychology and Education from University of Cambridge, and BSc in Computer Science from University of Westminster. His research interests include psychology, education, decision making, behavior modeling.