Magnet News:
You Choose the Polarity of What You Read

Júlio Reis
UFMG, Brazil
julio.reis@dcc.ufmg.br

Pollyanna Gonçalves
UFMG, Brazil
pollyannaog@dcc.ufmg.br

Pedro Olmo
UFMG, Brazil
olmo@dcc.ufmg.br

Raquel Prates
UFMG, Brazil
rprates@dcc.ufmg.br

Fabrício Benevenuto
UFMG, Brazil
fabricio@dcc.ufmg.br

Abstract

When was the last time you read a newspaper and was bombarded with articles you would rather not see? Current news media shows massive number of news every day. But from tragedies to happy stories, people might want to choose to read only those articles that fit their current mood. The purpose of this project is to present the *Magnet News*, a Web tool in which users can choose if they want to see positive or negative news. Our system monitors the sentiment of news from important newspapers using the SentiStrength, a sentiment method proposed by literature that has been proved to be efficient in previous analysis.

Introduction

Every day, online newspapers bombard their readers with many types of news, from tragedies (e.g. plane crash, murder stories, death of celebrities etc) to inspirational stories (e.g. selfless acts of kindness, people crossing the limits set by their disabilities etc). It is natural that this plethora of news is able to trigger a wide range of emotions in their readers, such as happiness, anger, sadness, euphoria, melancholy etc. In fact, the "Mood Management Theory", proposed by D. Zillmann and J. Bryant (Zillmann and Bryant), says that individuals' access to media is highly capable of change their mood states. But do we always want to alter our mood state when reading online news? Or, should we at least choose to which mood we would like to change to? We believe that the emotional and psychological states of readers that should dictate the type of news they would like to read, and not the contrary.

In a previous experimental research, Biswas (Biswas, Riffe, and Zillmann) proved Zillmann's theory, showing that when people have to choose the news they want to read, they are directly influenced by the mood of the place in which they had been eariler. In this experiment, authors placed men and women in bad and good environments and then provided them with magazine articles. Results showed that those who were previously exposed to a bad environment were drawn to good news. It also showed that people in a good mood read more bad news than people in a bad mood.
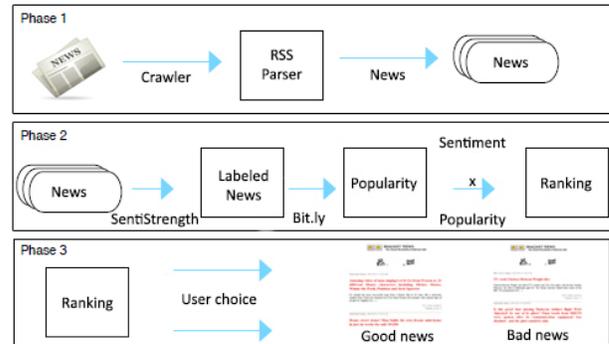
Figure 1: Magnet News running showed in three phases

Based on these facts, this work proposes the *Magnet News* (available at http://www.magnet-news.dcc.ufmg.br), an online newspaper where readers can choose the polarity of the news they would like to read. *Magnet News* has a simple interface and allows users to select news from popular online newspapers on the Web. For now, *Magnet News* gets its news from DailyMail and the BBC News, but it can easily scale up and process news from other popular sources.

The remainder of this paper is organized as follows. We first describe the strategy of collecting the news as well as our strategy to infer their popularity. Next, we discuss the used method for sentiment detection. Finally, we present a demonstration of the system and discuss future work.

System Architecture

In order to easily understand the system architecture of our system, we present Figure 1. Each process that is executed in all three phases will be described next.

News Sources

Our proposed online newspaper selects articles from two of the biggest online journals, DailyMail Online and BBC News Online. The DailyMail is a British tabloid newspaper and has more than one million readers daily. The BBC News Online is also a famous online newspaper, considered one of the most popular news website in the UK with around 14 million global readers every month. So, because of this, we consider both as good sources for our tool.

Again, we emphasize that our system's design allows us to collect news from other sources easily. Every day, our crawler collects the latest news from each RSS of the considered sources. A RSS (Rich Site Summary), also called "feed", is a format for delivering regularly changing web content and is used by most of news-related sites and others online publishers syndicate. Whereas these feeds can be easily accessed, the only arduous process of crawling news is to parse the RSS of each source, since they could be slightly different. After collecting the RSS, our database consists of information from each article, such as title, description and text.

Ranking Strategy

In order to define which news should be highlighted in our Web tool, we rank them by a relation between its sentiment score and its popularity. In the next section, we will describe our strategy of get the popularity of each news article. Now, let's assume that $S(a)$ is the sentiment score of an article a, and $P(a)$ as the normalized popularity of that same article. So, the relation between the sentiment score and the popularity of the article a, $R(a)$, can be described by the equation below:

$$R(a) = S(a) * P(a) \qquad (1)$$

With this strategy, we can order the articles that will appear at *Magnet News* by its sentiment score and popularity.

Link Popularity

In this work, we infer the popularity of an article using Bit.ly[1]. Bit.ly is a well-used URL shortening service that shortens billions of links per month (Antoniades et al.). The service's API provides the possibility of checking the total number of clicks that a shorten link has received. What is interesting here is that the service keeps the same shorten URL for requisitions of the same long URL. For example, if a user requests to shorten www.dummy.com, Bit.ly will return a shortened URL such as http://bit.ly/dummyID. But, if another user tries to shorten the same URL, the system will also return http://bit.ly/dummyID. This is important for us since the global number of clicks in a shorten URL provides a good estimation of popularity.

Sentiment Score Method

Sentiment analysis has been a hot topic in studies which analyze Web media (Gonçalves et al.). Thus, in order to be able to label the news articles collected by our system as bad or good news, we used SentiStrength, a machine-learning-based algorithm that detects sentiments in texts in a scale ranging from -5 (very negative) to +5 (very positive). SentiStrength was chosen because it was shown in a previous analysis (Gonçalves et al.) to be the method with better accuracy and coverage. SentiStrength method is available for testing in iFeel (Araújo et al.).

As we have presented before, our daily collection generated some information about each article, such as title, description, and its text. As processing all these information

[1]http://bit.ly

(a) Good news   (b) Bad news

Figure 2: Screen snapshots of the system when users select to read the good news

could increase the execution time of the method for sentiment detection, it would be interesting to identify which fields would really be relevant to identify an articles sentiment. So, in order to analyze the correlation among the articles title and text, we conducted an experiment with 300 collected articles. The Pearson's correlation (p) calculated was equal to 0.2656, implying that there is a low linear correlation between the two variables. Although this correlation is positive, it is not very strong. So, we chose to label each articles polarity by analyzing the sentiment of its text corpus, and not only its title.

Demo Scenarios

In this section, we present demonstrations of the *Magnet News* execution. Figure 2(a) shows the execution of the system when a user chooses to read the good news and bad news, respectively.

As we can see, besides presenting a list of news with the polarity selected, the system also shows the popularity (total number of clicks) and the polarity (the result of SentiStrength method) of each article. By publishing *Magnet News*, we hope that people could simply use it in order to select the type of news they want to read based on their current mood. As a natural extension of this tool, we would like to add more online news such as NewYorkTimes, CNN, Wall Street Journal and so on.

References

Antoniades, D.; Polakis, I.; Kontaxis, G.; Athanasopoulos, E.; Ioannidis, S.; Markatos, E. P.; and Karagiannis, T. 2011. We.b: The web of short urls. In *20th International Conference on World Wide Web*, 715–724. ACM.

Araújo, M.; Gonçalves, P.; Benevenuto, F.; and Cha, M. 2014. ifeel: A system that compares and combines sentiment analysis methods. In *WWW (Companion Volume)*. International World Wide Web Conference (WWW'14).

Biswas, R.; Riffe, D.; and Zillmann, D. 1994. Mood influence on the appeal of bad news. *Journalism and Mass Communication Quarterly* 71(4):689–696.

Gonçalves, P.; Araújo, M.; Benevenuto, F.; and Cha, M. 2013. Comparing and combining sentiment analysis methods. In *1st ACM Conference on Online Social Networks (COSN'13)*.

Zillmann, D., and Bryant, J. 1985. Affect, mood, and emotion as determinants of selective exposure. *Selective Exposure to Communication* 157–190.