Enhancing Brand Visibility by Identifying Influencers

Mr. Anmol Jain, Mr. BhavyaSanghavi, Mr. RishabhRathod, Prof. KhushaliDeulkar

Department of Computer Engineering,
University of Mumbai, India

Abstract - In this digital age, the influence of social media has crossed all bounds. Thus, it is necessary to understand and analyze the importance of social media in driving public thoughts. On the contrary, social media has provided a much needed impetus to brands in getting a valuable insight into customer preferences.
Today, it is necessary for a brand to stay socially active, relevant and connected. Many brands face the problem in understanding and properly implementing social media into existing business policy. It is necessary for the brand to reach out to the influencers, which magnify the brand visibility to the widest possible audience. Precisely, our model accounts for the influence of the blogger, the standing of the blog site which published a specific blog post, and the impact of the post itself. This strategic information brings the brands closer to their social customers with an integrated solution. Further, our algorithm considers the Google page rank, the in-links and the alexa analysis that refine the existing models further.
Our project gives a statistical analysis of various bloggers that can be used for brand monitoring, brand enhancing and brand advertising with the help of the calculated influence score.

Index Terms:**Blog analysis, Social Media Influence**

I. INTRODUCTION

The explosive growth of Social Media has provided millions of people the opportunity to create and share content on a scale barely imaginable in the first half of this decade. Humungous participation in these social networks is reflected in the opinions and reviews driven by the public that are constantly posted and discussed in social sites. Given the widespread popularity of the content, it is natural to target one's messages to highly connected people who will propagate them further in the social network. Particularly, in Twitter, which is one of the fastest growing social networks on the Internet, and thus the focus of advertising companies and celebrities eager to exploit this vast new medium. As a result, opinion, ideas, and products compete with all other content for the scarce attention of the user community. In spite of the abrupt fashion with which all these communications take place, certain issues capture considerable public attention, thus rising to the top in terms of popularity and contributing to new trends and to the public agenda of the community. How this happens in a world where crowdsourcing dominates is

still an unresolved problem, but there is considerable agreement on the fact that two aspects of information transmission seem to be important in determining which content receives attention.

II. PROBLEM FORMULATION

Social media influence matters on numerous levels. For businesses, online influence signals better engagement with customers and a smarter social media strategy. On an individual level, influence of social media can be a sign of advanced digital skills and an ability to create interesting social media content.

To break this down further, a better social media influence of the company social accounts can improve their brand's visibility in the online community and increase the reach of their social messaging. This is an implicit purpose behind investing valuable resources into a social media team or a social media manager. Considering the importance of social media, we have targeted the blogosphere. Our model aims to run across the blogosphere to generate a statistical report about the influence of the bloggers in various categories with the help of an influence score. This score is then used for various brand enhancement tasks.

III. PROPOSED SOLUTION

The proposed solution comprises of the following aspects:

- To develop a metric that generates an influence score by considering various social media parameters.
- The metric uses Facebook, Twitter, back-links, page rank and alexa statistics. This data is stored in a database.
- The algorithm is then run across this data to obtain an influence score.
- The scores range from 1 to 100, with greater scores corresponding to a higher ranking and strength of one's social influence.
- Along with the score, it provides a detailed report consisting of the comparison between different influencers.
- Further, it is run over the blog posts to derive the gist of the post that culminates into POS tagging.
- These tags can be used later to find similar posts and blogs.

- With the score and the report, it can be used for brand monitoring, brand advertising and generating insights into customer preferences.
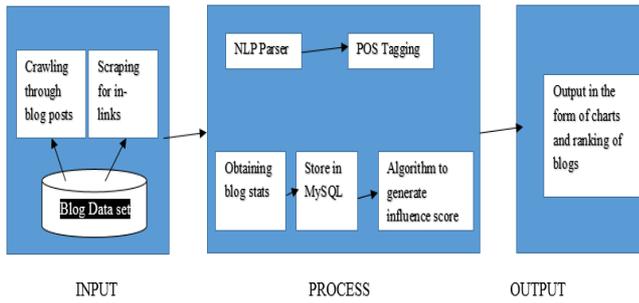
IV. ARCHITECTURE DESIGN

Figure 1: Architecture Design

A. INPUT

- **Crawling through blog posts**
  The tool runs across the data set of blogs. It does so to obtain the blogger's social media statistics like Twitter, Facebook. In addition, it also scrapes through to obtain Alexa statistics.
- **Scrapping for in-links**
  It also crawls through the web to derive the number of in-links as well as the Google Page Rank of the blog.

B. PROCESS

- **NLP Parser**
  It is run over the blog posts to derive the gist of the post that culminates into POS tagging. These tags can later be used to find similar posts and blogs.
- **Storing the Data**
  The data obtained by scraping and crawling is stored in the MySQl database through Xampp. This a dynamic database and changes as the data changes.
- **Algorithm to generate influence score**
  The algorithm is run on the data obtained and stored. The algorithm generates an influence score for each blogger.

C. OUTPUT

- **Report**

The algorithm generates a detailed comparison of the bloggers with their influence score. This report is generated in the form of Google charts and graphs.

V. IMPLEMENTATION

A. ALGORITHM/ METHODS USED

1. Scraping the data set of blogs
2. Collecting the values of social media attributes
3. Normalize the values of all factors
4. Calculate the relative score on a scale of 100
5. Display tags for vlogs used for categorization and SEO (Search Engine Optimization)
6. Data visualization of the scale and statistics

Factors considered for the influence score:

- Google Page Rank
- Backlinks
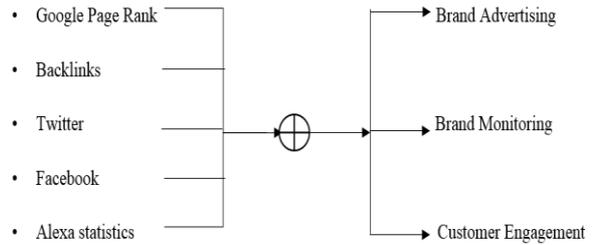- Twitter (retweets, followers, following)

- Facebook (likes)

Figure 2: Overview

VI. RESULTS AND DISCUSSIONS

Comparison of the influence score with Wordpress Social Media Authority Tool

- The Wordpress Social Media Authority tool takes an URL as input and gives a number representing the Social Media Authority.
- This number ranges anywhere in the positive scale.
- Few websites may have the Social Media Authority number rising as far as 6 digits. Thus comparison of different blogs becomes a tedious task due to the large difference in their respective Social Media Authority Number.
- The SMA number takes only social media platforms into consideration like Twitter and Facebook. Other parameters like Google Page Rank and Alexa statistics are not considered.

- o The social influence score calculated for each blog in our project is normalized and is represented on a scale of 100.
- o This score is user-friendly and enables any user to compare the score of different blogs in the dataset.
- o Also, the algorithm does not only depend on the social media platforms. It also considers other parameters like Google Page Rank and backlinks.

Thus, the results are in sync with each other. The influence score is more readable and accounts for more number of parameters thus increases the reliability.

VII. CONCLUSIONS AND FUTURE SCOPE

Our project has built a tool that gives a statistical report capturing the minutiae of the bloggers with a calculated social media influence score. The details and the exhaustive analysis are further used for brand advertising, brand monitoring and reflects customer engagement and thus, enhances the brand visibility.
The future work involves mitigating the risks and hurdles that have come across while executing our problem statement. The social media influence can thus be made efficient by considering the following factors and getting rid of the limitations:

- Enhancing the score by considering other Twitter parameters like mentions, tags, favourites, etc.
- Tackling spam bots that hinder the effective scrapping of data.

- Requesting permission from Facebook users to access parameters like friends, interests, etc.
- Expanding the social media influence to other platforms like pinterest, tumblr, instagram, etc.
- Performing sentiment analysis to comments and posts in order to determine its relevance.

REFERENCES

[1] Shail K Dinkar, Hemant Kumar , Interaction Information Retrieval and Improved Page Rank Algorithm Based on Aceess Duration of Page, www.ijert.org, ISSN: 2278-0181, August(2012)

[2] Martha Sideri, IraklisVarlamis, Blogrank: Ranking on the blogosphere, Athens University of Economics and Business.

[3] Page, L., Brin, S., Motwani, R. &Winograd, T. (1998). The pagerank citation ranking: Bringing order to the web. Technical report, Stanford, USA.

[4] Nakajima, S., Tatemura, J., Hino, Y., Hara, Y., Tanaka, K., (2005), "Discovering Important Bloggers based on Analyzing Weblog Threads", 2nd Annual Workshop on the Weblogging Ecosystem: Aggregation, Analysis and Dynamics, WWW 2005.

[5] Xiaodan Song, Yun Chi, Koji Hino, Belle L. Tseng, Identifying Opinion Leaders in the Blogosphere , CIKM'07, November 6-8, 2007, Lisboa, Portugal.

[6] Xiaodan Song, Yun Chi, Koji Hino, Belle L. Tseng, Summarization System by Identifying Influential Blogs, ICWSM 2007, March 26-28, 2007, Boulder, Colorado, U.S.A.

Anmol Jain, B.E. in Computer Engineering, Dwarkadas J.Sanghvi College of Engineering, Mumbai, India.

BhavyaSanghavi,B.E. in Computer Engineering, Dwarkadas J.Sanghvi College of Engineering, Mumbai, India.

RishabhRathod, B.E. in Computer Engineering, Dwarkadas J.Sanghvi College of Engineering, Mumbai, India.

KhushaliDeulkar, Assistant Professor, Dwarkadas J.Sanghvi College of Engineering, Mumbai, India.