Influence Tracker using Sementic Analysis

Dhobale Madhuri Pandurang[1] Lande Bhakti Dadabhau[2] Nighot Rutuja Dattatray[3]
Wagh Rupali Kailas[4]
[1,2,3,4]Department of Computer Engineering
[1,2,3,4]JCOE University of Pune, India

Abstract— The problem of increasing influence spread has been widely studied in social networks, because of its extreme number of applications in determining critical topics in a social network for information dissemination. In survey, all the method are static in nature, which are designed for social networks with a constant set of links. However, no of forms of social interactions are flexible in nature, with relatively short periods of interaction. Any influence spread may happen only during the period of interaction, and the probability of spread is a function of the corresponding interaction time. In such cases, it may be useful to consider the influential nodes based on the run time interaction patterns. Alternatively, one may wish to find the most likely starting points for a given infection pattern. We will propose methods which can be used both for reduction of information spread, as well as the backward tracing of the source of influence spread. The LDA (Latent Dirichlet Allocation), Sentiment Analysis and greedy algorithms are used. We will present practical results implement the effectiveness of our approach on a number of real data sets.

Keywords: Data Mining, Influential Nodes Tracking, Social Network, Greedy Algorithm, Sentiment Analysis

I. INTRODUCTION

Social Media and Micro blogging platforms like Facebook, Twitter, LinkedIn used for spreading encapsulated news and trending topics across the globe at a rapid pace. A topic becomes trending if more and more users are contributing their opinion and judgments, thereby making it more valuable for online perception[3].sentiment analysis of twitter determine the polarity and inclination of vast population towards specific topics, items or entity.

Now a days, the application of data analysis can easily decided during public election, movie promotion, brand agreement and many other fields.Twitter data is used as a whole for analysis by the application of sentiment analysis and a study of different algorithms and methods that help to track influence and impact of a particular user/brand active on the social network.

Social media influence is a marketing term that describes an individual's potential to affect other people's thinking mind in a social online community. The more influence a person has, the more plead that individual has to companies or other individuals who want to promote an idea or sell a product.

At its most basic level, by using twitter followers, Facebook friends or any other social media the influence can be calculated. A more through analysis is required, however, to determine how a person connect with social connections, how those people connections are and the level of trust between the person and their connections.

II. RELATED WORK

Scalable Influence Maximization for Prevalent Viral Marketing [6] in Large-Scale Social Networks, Influence maximization, defined by Kempe, Kleinberg, and Tardos (2003), is the problem of finding a small set of seed nodes in a social network that maximizes the spread of influence under certain influence cascade models. The scalability of influence maximization is a key factor for enabling prevalent viral marketing in large-scale online social networks. Prior solutions, such as the greedy algorithm of Kempe et al. (2003) and its improvements are slow and not scalable, while other heuristic algorithms do not provide consistently good performance on influence spreads. In this paper, we design a new heuristic algorithm that is easily scalable to millions of nodes and edges in our experiments.

Influential Neighbours Selection for Information Diffusion [8] in Online Social Networks However, this model is usually unrealistic in online social networks since we cannot typically choose arbitrary nodes in the network as the initial influenced nodes. From the point of view of an individual user who wants to spread information as much as possible, a more reasonable model is to try to initially share the information with only some of its neighbour's rather than a set of any arbitrary nodes; but how can these neighbour's be effectively chosen?

Mining Social Networks for Viral Marketing [2] these models allow us to design "viral marketing" plans that maximize positive word-of-mouth among customers. Social network models have been descriptive, rather than predictive: they are built at a very coarse level, typically with only a few global parameters, and are not useful for making actual predictions of the future behaviour of the network. In the past, this was largely due to lack of data the networks available for experimental study were small and few, and contained only minimal information about each node for find out viral news in social media.

III. IMPLEMENTATION

We are calculate topic influence and user influence and also rank a node in the form of positive, negative and neutral rating for node by using sentiment analysis which can helping to predict a node in future.

Nowadays, the age of Internet has changed the way people express their point of views, opinions by using the blog online forums, post, product review websites, etc. Number of people are using social network sites to express their feelings, opinion and share views about their lives tweets. Through the online communities, we get an interactive media where consumers inform and influence others through assembly. Social media is generating a large volume of sentiment large data in the form of tweets, status updates, blog posts, comments, reviews, etc. Moreover, social media provides an

opportunity for businesses and marketing by giving a platform to connect with their customers for advertising. People mostly depend upon user generated review over online to a great extent for decision making. For e.g. if someone wants to buy a product or wants to use any service, then they firstly look up its reviews online, discuss about it on social media before taking a decision. The amount of content generated by users is too vast for a normal user to analyse. So there is a need to automate this, various sentiment analysis techniques are widely used. Sentiment analysis inform user the information about the product is satisfactory or not before they buy it. Marketers and firms use this analysis data to understand about their products or services in such a way that it can be offered as per the user's requirements.
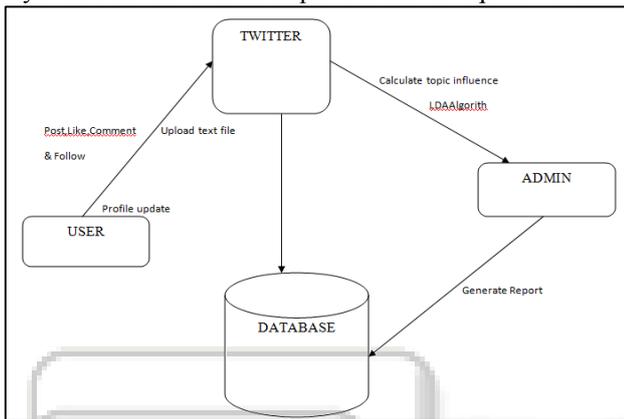

Fig. 1: System Architecture

We are using following algorithm:

A. LDA:

LDA(Latent Drichlet Allocation) is technique or method for finding topics that these document contain.it is topic modelling approach and also used for classification purpose. We are using LDA algorithm to classify the large amount of sample dataset into no of parts for easy to surveying.

B. Sentiment Analysis:

Sentiment analysis is the process of opinion mining text to identify and classify the subjective opinions expressed by the writers. Normally it is used to determine whether the writer's attitude towards a particular topic or product, etc. is positive, negative, or neutral. It is used by businesses to help them understand the social sentiment of their brand, product or services while monitoring online conversations. In the context of a twitter sentiment analysis, at its simplest, sentiment analysis quantifies the mood of a tweet or comment by counting the number of positive and negative words. By subtracting the negative from the positive, the sentiment score is generated.

IV. RESULT AND CONCLUSION

The influence maximization problem in social network is analysed using sample data taken from social tweets. The greedy method is applied for finding influence node tracking in a social network. This heuristic search technique allows to finding a best active node in a network Therefore the best active node is considered as most powerful node which influences another node in a network. When the active node is identified then the next step finding best communication network. Hence the active network is identified. And also finding negative and positive impact for particular node.The practical results achieved better performances.

REFERENCES

[1] W. Chen, Y. Wang, and S. Yang, "Efficient influence maximization in social network," in KDD, 2009, pp. 199– 208.
[2] P. Domingos and M. Richardson, "Mining the network value of customers," in KDD, 2001, pp. 57–66.
[3] D. Kempe, J. Kleinberg, and E. Tardos, "Maximizing the spread of influence through a social network," in KDD, 2003, pp. 137–146.
[4] M. Kimura and K. Saito, "Tractable models for information diffusion in social networks," in PKDD, 2006, pp. 259–271
[5] W.Yu, G.Cong, G.Song, and K.Xie, "Community-based greedy algorithm for mining top-k influential nodes in mobile social networks," in KDD, 2010, pp. 1039–1048.