© 2015 IJSRSET | Volume 1 | Issue 6 | Print ISSN : 2395-1990 | Online ISSN : 2394-4099
Themed Section: Engineering and Technology

# Trend Analysis in Social Networking using Opinion Mining A Survey

**Saurin Dave*, Prof. Hiteishi Diwanji**
Information Technology Department, L. D. College of Engineering, Ahmedabad, Gujarat, India

## ABSTRACT

The popularity of social media in recent years has increased a lot and created new opportunities to study the interactions of different groups of people. Because of this it comes to possible to study the user's opinion. Two popular topics in the study of social networks are community detection and finding trends. Trend Analysis is an ongoing field of research in Data mining field, May referred as an "Opinion Mining" by some researchers. Trend Analysis is the computational treatment of opinions, sentiments and subjectivity of text. The related field to Trend Analysis contains finding emotions in a statement which attracted researchers recently are discussed. The main target of this survey is to give nearly full image of Trend Analysis along with the community detection and their related fields with brief details. The main contributions of this paper include the overall analysis to bridge the concept between trend analyses with community detection. To find out the anomaly detection in a community with using the concept of trend analysis is a difficult task.

**Keywords**: Trend Analysis, Sentiment Analysis, Opinion Mining, Data Mining, Trend Detection, Anomaly Detection

## I. INTRODUCTION

Trend analysis can also consider as an Opinion Mining which is a computational study of people's opinions, attitudes and emotions toward an entity. The entity can represent individuals, events or topics. These topics are most likely to be covered by reviews. These expressions may interchangeable according to different authors.

However, some researchers stated that Opinion Mining and Trend Analysis have slightly different notions. Opinion Mining extracts and analyses people's opinion about an entity while Trend Analysis identifies the sentiment expressed in a text then analyses it. Therefore, the target of Trend Analysis is to find out opinions, identify user's review they express, and then classify their polarity as shown in the Figure - 1.

Trend Analysis can be considered a classification process as illustrated in Fig. 1. There are three main classification levels according to[1]: document-level, sentence-level, and aspect-level. Document-level aims to classify an opinion document as expressing a positive or negative opinion or sentiment. It considers the whole document a basic information unit (talking about one topic). Sentence-level aims to classify trends expressed in each sentence. The first step is to identify whether the sentence is subjective or objective. If the sentence is subjective, Sentence-level will determine whether the user's sentence expresses positive or negative opinions.

### A. Feature selection

Trend Analysis task is considered a sentence classification problem. Firstly in "Sentence Classification" problem is to extract and select text features. Some of the current features are :

Terms presence and frequency: These features are individual words and their frequency counts. It either gives the words binary weighting (zero if the word appears or one if otherwise) or uses term frequency weights to indicate the relative importance of features. Parts of speech (POS): To find adjectives, as they are important indicators of opinions in any particular review.

Opinion words and phrases: These are words commonly used to express opinions including good or bad, like or hate, etc. On the other hand, some phrases express opinions without using opinion words. For example: cost me an arm and a leg.

Negations: the appearance of negative words may change the opinion orientation like not good is equivalent to bad. That means any positive word can reflect the negative meaning if negation happens before it.

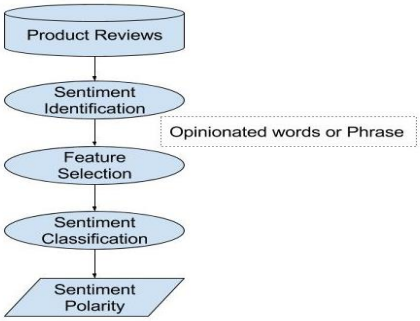


**Figure 1** Sentiment analysis process on product reviews [1]

### B. Sentiment classification

Sentiment Classification techniques can be divided into machine learning approach, lexicon based approach and hybrid approach. The Machine Learning Approach (ML) applies the famous Machine Learning algorithms. The Lexicon-based Approach relies on a sentiment lexicon, a collection of known and precompiled sentiment terms which can be implemented by defining the set of rules. It is divided into dictionary-based approach and corpus-based approach which use statistical or semantic methods to find sentiment polarity or sometimes it may interpret as a sentiment score. The hybrid Approach combines both approaches and is very common with sentiment lexicons playing a major role in the majority of methods.

### C. Applications of Sentiment Analysis

One very popular way of finding the review which is "Word of Mouth" is the way of giving information from one person to another person. By giving the right reviews, it helps the customers in taking decisions. Word of Mouth gives the information about the reactions, opinions or attitudes of consumers about products, business or services that they share with the society. It imparts the information on the basis of social networking and trust. Most of the people in their social network depend on families, friends, and others. It is also indicated by research that people believe on the opinions of other people quickly that exist outside their social networks such as online reviews which is very surprising. Hence, this is where Trend Analysis comes into study. As the online review sites, blogs, social networking sites provide huge amount of opinions, this helps in making decision making process easier for us.

### D. Community Detection

Because of the prevalence of social networks, community detection on these networks has become an important research topic in these days. Using community detection, useful metadata about large scale networks can be captured. These communities represent relationships between entities and allow us to examine patterns that emerge in social media, publications, and a multitude of other types of networks. Community detection also allows for easy visualization of networks and their structure. So, from this it is easy to identify to understand the interest and hates of different communities. This kind of study is going to be very helpful for the market researchers.

## II. LITRETURE REVIEW

In [1], the author has discussed regarding all the basic concepts of the Sentiment Analysis and various techniques to implement it.

In [2], Trend analysis finds the phrases in a text or document that contains some sentiment. There may be some objective facts or subjective opinions in the text. It is compulsory to distinguish between them. SA helps in determining the entities and subject from text towards which sentiment is directed. Sentiments are categorized as objective (facts), positive (represents a state of gladness, happiness, pleasure or satisfaction) or negative (represents a state of sorrow, regret, sadness or disappointment). On the basis of degree of polarity, a score can be given to the sentiments. In this paper author has discussed regarding to the Negation phrase identification and accordingly finds out the sentiment score and concludes POS (Part-of-Speech) as a negative or positive. But not discussed regarding negative adjectives and adverbs. Opinion words and phrases such as "like", "nice", "hate", "I'd suggest that..." are words

or phrases that convey positive or negative opinions. Statistical-based or Lexicon-based are the main approaches which identify the semantic orientation (positive or negative) or polarity of opinion words.

In [3], the author has defined a new type of "in-disguise" anomaly, the community based anomaly. In addition to that there are only six possible types of community based anomalies in evolutionary networks: grown, shrunken, merged, split, born, and vanished communities. Here, the author has not mentioned that how to implement this concept and also not defined which algorithm should use to find out the type of it.

In [4], Anomaly detection is an important problem with multiple applications. As real-world networks are constantly changing, there has been a shift in focus to dynamic graphs, which evolve over time. a user's perspective, people are able to post their own content through various social media. This will be helpful for the data analytics.

In each paper they have discussed either regarding the Trend Analysis or Community Detection. So, the hybrid approach of Community Detection in Trend Analysis is completely missing which can be the future area of interest for the researchers.

## III. METHODS AND MATERIALS

### A. Datasets

From a user's perspective, people are able to post their own content through various social media, such as forums, micro-blogs, or online social networking sites. From a researcher's perspective, many social media sites release their application programming interfaces (APIs), prompting data collection and analysis by researchers and developers. For instance, Twitter currently has three different versions of APIs available, namely the REST API, the Search API, and the Streaming API. With the REST API, developers are able to gather status data and user information; the Search API allows developers to query specific Twitter content, whereas the Streaming API is able to collect Twitter content in real time. Moreover, developers can mix those APIs to create their own applications. Hence, sentiment analysis seems having a strong fundament with the support of massive online data [2]. In these days twitter will be the most important platform to understand the people's opinion. Many researchers are working on this to make it as accurate as possible.

### B. Approaches for Sentiment Analysis

Depending on the perspectives of the different persons doing the sentiment analysis, the approach can be keyword-based, concept-based, lexical affinity based, or discourse driven.

Keyword-based Approach: In Keyword-based approach, main task is the construction of word lexicons. So that, text can be classified in to affect category on the basis of presence of affect words like "happy", "awesome", "sad", "bored".

Concept-based Approaches: The concept-based approaches use web ontologies and semantic networks to achieve semantic text analysis. Thus, these approaches help the system in extracting the conceptual and affective information from natural language opinions. These approaches mainly rely on implicit meaning or feature associated with natural language concepts. So, these approaches are better than the approaches which use keywords and word co-occurrence counts. Concept-based approaches can detect the sentiments better than syntactical techniques. These approaches can also find multi-word expressions even the expressions don't convey any emotion explicitly. The concept-based approaches mainly rely on the knowledge bases.

Lexical Affinity: Lexical affinity approach is slightly more advanced than keyword-based approach. This approach assigns a probabilistic 'affinity' to arbitrary words for a particular emotion rather than simply detecting affect words in the text.

Discourse Structures: In discourse structures approach, discourse relations between text components are used as features for classification.

## IV. RESULTS AND DISCUSSION

Trend analysis is a process of extracting information from user's opinions. The decisions of the people get affected by the opinions of other people. Today, if any person wants to buy a product or wants to watch a movie then he or she will first search the reviews and opinions about that product or movie on social media, blogs etc.

As there is a huge explosion of user's opinion on social media like Twitter, Facebook and other user forums, then identification of sentiment becomes very difficult from this huge data manually. Many times people's choice keeps changing and may happen that one may behaves completely opposite to the community towards to which they belongs. To find out such behavior in these days very essential but accuracy is still a major issue. So, there is a need of trend analysis system along with the community detection. In Trend analysis it is required to generate a proper rule set to analyze the emotions in a text.

## V. CONCLUSION

Community detection and sentiment analysis are two important topics in the study of social networks. Sentiment Analysis is very crucial factor for finding the trends. Each community behaves differently from each other. So, study the behaviour of community is a major issue. Sentiment analysis or opinion mining is a field of study that analyses people's sentiments, attitudes, or emotions towards certain entities. This paper tackles a fundamental problem of sentiment analysis, sentiment polarity categorization. While these are generally treated as separate issues, so in future more study should be adopted an integrative approach that enabled granular sentiment analysis on the level of individual communities. Finally we required to make a rule based sentiment analysis system and afterwards it needs to integrate with the community detection based approach. So, we want a system which can able to detect community based on trend analysis.

## VI. REFERENCES


[1] Walaa Medhat, Ahmed Hassan, Hoda Korashy, "Sentiment analysis algorithms and applications: A Survey". (27 May 2014), Volume 5, Issue 4, December 2014. http://dx.doi.org/10.1016/j.asej.2014.04.011

[2] Xing Fang* and Justin Zhan "Sentiment analysis using product review data" Journal of Big Data (2015) 2:5, DOI 10.1186/s40537-015-0015-2

[3] Zhengzhang Chen, William Hendrix, Nagiza F. Samatova, "Community-based anomaly detection in evolutionary networks", © Springer Science+Business Media, LLC 2011

[4] Stephen Ranshous, Shitian Shen, Danai Koutra, Steve Harenberg, Christos Faloutsos3 and Nagiza F. Samatova, "Anomaly detection in dynamic networks: a survey", Volume 7, May/June 2015, © 2015 The Authors. WIREs Computational Statistics published by Wiley Periodicals, Inc.