International Journal of Applied Engineering Research ISSN 0973-4562 Volume 14, Number 1 (2019) pp. 118-125


# A Novel Optimizer Technique for Sentiment Polarity in Social Web Environment

**K.K. Uma**[1] and **Dr. K. Meenakshisundaram**[2]

*[1]Research Scholar, [2]Associate professor, Department of Computer Science, Erode Arts and Science College, Erode, Tamilnadu, India.*

**Abstract**

Sentiment analysis is a process that automates mining of attitudes, opinions, views and emotions from text, speech, tweets and database sources. The sentiment trend analysis is a significant research component of the social media big data. It attracts many researches which could support many real-world applications. Most of the traditional methods are mainly uses the standard emotion thesaurus, supervised learning and statistical methods. These methods are not scalable to the social media big data. So, the evolutionary based methods are necessary for this type of unknown search space problems. Therefore, proposed a grey wolf-based sentiment trend analysis method which is specifically for the social web big data sentiment trend analysis problem. In this work, a new technique based on optimization is introduced for sentiment trend analysis (i.e.) Grey Wolf Optimization (GWO) algorithm is used in the area of sentiment analysis. Social web data are collected in the form of Resource Description Framework (RDF) triples, and then GWO algorithm is used to digitize the collective RDF triples. User's sentiment scores are calculated for the computed positions of wolves concerning the sentiment words with Opinion Lexicon. Based on the wolf best positions sentiment scores are computed to analyze the trends of the user's sentiments. At the end, sentiment trend of the online user is analyzed by time. Experimental results carried on real-world Airbnb datasets and demonstrate the excellent performance of the proposed grey wolf-based sentiment trend analysis method on the sentiment trend analysis.

**Keywords:** Sentiment Trend Analysis, Opinion Mining, Grey Wolf Optimization, Semantic Mining, Unknown Search Space Problems.

## I. INTRODUCTION

People opinion is most significant for many decision-making processes. Few years before, people asked opinion to their friends and family; consumer asked opinion to other consumers who already purchased or used the product or service. An organization made decisions based on the opinions of the consumers by conducting surveys, focusing groups or opinion polls. But now, Due to increased number of usages in social media, gathering information or opinion is so easy. Large numbers of opinion available on the web. From the past two decades the field has grown rapidly to become one of the most active research areas in NLP. NLP is researched in data mining, Web mining, and information retrieval.

Sentiment Analysis (SA) is the computational study of people's opinions, attitudes and emotions towards an entity. The entity can be individuals, events or topics. These topics most are to be covered by reviews. Sentiment Analysis is known as Opinion Mining (OM). However, some researchers stated that sentiment analysis and opinion mining have a little difference. Sentiment Analysis includes many tasks few are sentiment extraction, sentiment classification, subjectivity classification, summarization of opinions or opinion spam detection, among others. It focusses to analyse people's sentiments, attitudes, opinions emotions, etc. towards various things such as, products, individuals, topics, organizations, and services. Sentiment Analysis identifies the sentiment expressed in a text then analyzes it while opinion mining extracts and analyzes people's opinion about an entity. Hence, the aim of sentiment analysis is to find opinions, recognize the sentiments expression, and then classify their polarity.

Sentiment Analysis is an application of data mining that extracts and analyzes the unstructured data automatically. Sentiment Analysis is a Natural Language Processing (NLP) technique that automatically extracts the unstructured data in proper context and classify these into different categories like positive, negative, neutral etc. The unstructured data are the kind of the opinion, sentiments, attitude, emotions, views etc. Opinion Extraction and Sentiment classification are the major tasks of sentimental analysis. Opinion Extraction is extracting the opinionated phrases, in proper context, from free text and Sentiment classification is classifying opinionated phrases based on sentiment orientation. Machine learning based techniques and lexicon-based techniques are the two techniques involved in Sentimental analysis. In sentiment classification Various machine learning algorithms are used. Supervised and unsupervised learning algorithm are the two types of machine learning algorithm. In Lexicon based techniques, sentiment words in sentiment dictionary that sentiment words are used for sentiment classification.

Compared to the conventional sentiment analysis, a sentiment trend analysis includes a continuous sentiment analysis that combines many discrete, analyzed sentiments, and other additional information. The sentiment trend analysis provides much more valuable information, since it gives a trend of sentiments, as well as the sentiments of each day. In addition, the sentiment trend analysis has the ability to provide latent information by inferences of sentiments between two different times, while the conventional sentiment analysis approaches do not. Hence a hybrid sentiment trend analysis method is proposed, by using a nature-inspired optimization algorithm, and a well-defined, online lexical resource.

This work suggests a new technique for sentiment trend analysis using Grey Wolf Optimization (GWO) algorithm. Here, Facebook portal data is used for analyzing. Initially, 'feeling' tag properties on each user's Facebook wall are collected, and then represented them in the form of Resource Description Framework (RDF) triples. GWO algorithm is used to extract the trend of each user's sentiments from the accumulated social web data. By using the leadership hierarchy such as position of wolves in the GWO algorithm, we can quickly consider both the past and current sentiments of each user. Next, we apply two sentiment scores such as positive and negative scores supported by opinion lexicon to the collected triples, and calculate a total sentiment score, based on updated positions new score values are calculated with respect to each word. Based on the computed total sentiment scores considering both past and current sentiments, we are able to analyze the user's sentiment trend. Finally, experiments are conducted to evaluate the proposed method. Experimental results show that this technique is effectively analyzes the user's sentiment trend.

This paper is organized as follows. Section 2 represents related research of this work. Section 3 describes the proposed algorithm for sentiment analysis. Section 4 explains the simulation results and Section 5 concludes this paper.

## II. RELATED WORKS

In the past decade of years, various soft computing meta-heuristic algorithms have been applied in the field of sentiment analysis. Gohil et al (2018) has proposed sentiment analysis of Health Care. Tools available for sentiment analysis of twitter health care research are analyzed by reviewing existing studies, then method would work best was determined. Kwon et al (2017) have introduced sentiment trend analysis in social web environments. Ant Colony Optimization (ACO) algorithms and SentiWordNet were used for sentimental analyzing. Here, user generated social data was collected in the form of RDF (Resource Description Framework) triples, and then ACO algorithm was used to digitize the amassed RDF triples. Pheromone values were computed to extract trends of the user's sentiments by the use of ACO algorithm equations. User's overall sentiment score were calculated with the computed pheromone values and sentiment scores obtained from SentiWordNet. At last the trend of user's sentiments was analyzed by time.

Kotelnikov and Pletneva (2016) have introduced text sentiment classification based on a genetic algorithm and word and document co-clustering. Here a positive and negative sentiment expressed in the text was automatically recognized. Krishna et al (2013) has introduced polarity trend analysis of public sentiment on youtube. In this paper sentiment analysis of the YouTube comments related to popular topics were performed using machine learning techniques. Korkontzelos et al (2016) has analyzed the effect of sentiment analysis on extracting adverse drug reactions from tweets and forum posts. Medhat et al (2014) has surveyed sentiment analysis algorithms and applications. Aim of this survey is to give full image of SA (Sentiment Analysis) techniques. Sahayak *et al (2015)* discussed about a paradigm to extract the sentiment from a famous micro blogging service, Twitter, where users post their opinions for everything. They have also discussed about the sentiment analysis with the Machine learning algorithms. In this work machine learning algorithms were used to classify the sentiment tweets in the twitter. Tree kernel and Parts of Speech were used in this research work.

Suchdev *et al (2014)* analysed people's sentiments in their tweets and it makes the companies to bring up profits in their products by increasing the quality and the recommendation given by the customer's tweets. Positive and negative tweets were noted and further things were proceed according to that records. In this paper they have discussed about the sentiment summarization system which takes the documents that have to be analysed as input, and generates a detailed document summarizing the opinions in the input documents.

Sentiment analysis with Hadoop was done by Ingle *et al (2015)*. They proposed that analysis of the sentiments of Twitter users through their tweets in order to extract what they think using Hadoop which will process large amount of data and clustering is done faster in Hadoop. Contextual polarity of text can be determined by sentiment analysis. In this process, they have made the twitter data in the form of clusters with the help of Hadoop. Li *et al (*2010) mainly focuses on the document level sentiment classification. In his work, he added sentiment classification is to classify a text according to the sentimental polarities it contains in it. Favourable or unfavourable is the best example for it. Sentiment classification achieved state-of-art performance. Word level, phrase level, sentence level and document level were the four different levels in which sentiment classification was performed. Term-counting approaches (lexicon-based) and machine learning approaches (corpus-based) were the two kinds of document level in the sentiment classification.

Zhang *et al (2013)* work proposed that directly at the user sentiment state of twitter, the unique medium, and applies set pair analysis method for trend analysis. They started with set pair contact degree, then based on set pair affective computing model to make comparison with the size relationship of same degree, difference degree, opposition degree of the emotion, to design the user sentiment trend analysis model. Secondly, they analysed the sentiment among the owners. One user's sentiment orientation threshold as prerequisite for user behaviour prediction was also analysed.

This research gave better results. Guha et al (2016) has proposed a weakly supervised approach (named, TweetGrep) lets the data analyst easily define a topic by few keywords and adapt a generic sentiment classifier to the topic – by jointly modelling topics and Sentiment Analysis of Topical Tweets. Madhoushi et al (2015) has surveyed Sentiment analysis techniques in recent works. Rejanimol et al (2017) has introduced SA of Movie Reviews Based on Twitter Data. Collomb et al (2014) has studied and compared sentiment analysis methods for reputation evaluation.

Nielsen (2011) has proposed a new evaluation of a word list for sentiment analysis in microblogs. Rosenthal et al (2017) has continued SemEval task 4: Sentiment analysis in Twitter. Fang and Zhan (2015) have introduced sentiment analysis using product review data. Aim of this work was to tackle the problem of sentiment polarity categorization i.e. the fundamental problems of sentiment analysis. Santos and Gatti (2014) have introduced deep convolutional neural networks for sentiment analysis of short texts. Hasan et al (2018) has proposed machine learning-based sentiment analysis for twitter accounts.

Mirjalili et al (2014) have introduced Grey wolf optimizer which was inspired by grey wolves. This algorithm is based on the leadership hierarchy and hunting mechanism of grey wolves in nature. Four types of grey wolves were used such as alpha, beta, delta, and omega. Also, the hunting was described by separating that process into three categories. Precup et al (2017) has introduced an easily understandable grey wolf optimizer and its application to fuzzy controller tuning. Yassien et al (2017) has applied GWO to the 0/1 Knapsack Problem. The knapsack problem in networks is investigated in this paper. To find the best solution new algorithm is introduced that maximizes the total carried value without exceeding a known capacity using Grey Wolf Optimization (GWO) and K-means clustering algorithms.

The literature study shows that there are various algorithms applied for sentiment trend analysis, many of them based on search behaviours and hunting. But there is no swarm intelligence approach is applied for sentiment trend analysis that mimicking the leadership hierarchy of grey wolves, well known for their pack hunting. Based on this motivation Grey wolf algorithm proposed by Mirjalili et al (2014) is applied for sentiment trend analysis which is inspired by the grey wolf's leadership hierarchy and find out its capabilities in solving standard and real-life applications.

## III.SENTIMENT TREND ANALYSIS

The mood of a person may vary from time to time and environment. It reflects their daily day today life and also in the social media. This research mainly focusses on the user's sentiment about an event in the social media whether their comments and sentiment reflect the positive, neutral or negative impact on the web. In this paper, grey wolf optimizer is applied to analyse the impact of the user sentiments over the web data. Grey wolf optimizer is a new meta heuristic inspired by grey wolves, it is based on leadership hierarchy and the hunting mechanism of the grey wolves. The leadership hierarchy is proposed based on four types of wolves' like alpha, beta, delta and omega. First step in the sentiment analysis is to collect the user generated data from the social web and convert the social data into RDF format (subject, predicate, object). Second step is to set the grey wolf to initial value w (0) and Compute the updated grey wolves' positions for each sentiment word. Third step is to combine overall grey wolves position and compute overall sentiment score S by date. Last step is the visualize and personalize sentiment trend.

### *A. Grey Wolf Optimization*

Grey wolf optimization is a swarm intelligent technique, which mimics the leadership hierarchy of wolves are well known for their group hunting. Mirjalili et al put forwarded the GWO technique. Grey wolf belongs to Canidae family and they exist as a social unit called a pack. Grey wolves are categorized into four types based on hierarchy. They are alpha, beta, omega and delta. The alpha (α) is an oldest member and leader of the pack. It is mainly responsible for decision making. The orders of the dominant wolf should be followed by the pack. The Betas (β) are subordinate wolves which help the alpha in decision making. An advisor of alpha and discipliner for the pack is the beta. The lower ranking grey wolf is Omega (Ω) which has to submit all other dominant wolves. If a wolf is neither an alpha or beta nor omega, is called delta. Delta (δ) wolves dominate omega and reports to alpha and beta. In GWO algorithm, □, □ and □. Guides the hunting process. The □ solutions follow these three wolves. During hunting, the grey wolves encircle prey. Mirjalili *et al* (2014) work is considered for clustering with challenging problems. The mathematical model of the encircling behaviour is given below.

$$D = |C.X_p(t) - A.X(t)| \qquad (1)$$

$$X(t+1) = X_p(t) - A.D \qquad (2)$$

Where 't' is the current iteration, A and C are coefficient vectors, $A=2a.r_1-a$, $C=2.r_2$ where components of a are linearly decreased from 2 to 0 over the course of iterations and r1, r2 are random vectors between 0 and 1. $X_p$ is the position vector of the prey, and X refers to the position vector of a grey wolf.

Step 1: Initialize the grey wolf population $X_i$ (i = 1, 2... n) and coefficients A, and C

Step 2: Calculate the fitness of each search agent

$X_\alpha$=the best search agent

$X_\beta$=the second-best search agent

$X_\delta$=the third best search agent

Step 3: while (t < Max number of iterations)

For each search agent update the position of the current search agent.

Step 4: Update $\alpha, A$ $and$ $C$, and Calculate the fitness value for all search

Grey Wolf Optimization Algorithm

Summarizing the search process starts with creating a random population of grey wolves (candidate solutions) in the GWO algorithm. Over the course of repeated procedures, probable position of the prey is estimated by alpha, beta, and delta wolves. Each candidate solution amends its distance from the prey. The parameter a is diminished from 2 to 0 in order to emphasize exploration and exploitation, respectively. Candidate solutions changes its direction from the prey when A > and converge towards the prey when A< 1. Finally, the GWO algorithm is ended by the satisfying the convergence criterion.

### *B. Proposed Grey Wolf Optimizer for Sentiment Trend Analysis*

This system consisting of four steps: (a) RDF triple-based data collection, (b) Wolfs position updation using GWO algorithm, (c) computation of overall sentiment score, and (d) sentiment trend analysis.

#### *1. RDF Data Collection and Representation*

The social web data have many web attributes, few of them are user's personal information, their location and posted time, images, videos, comments and sentiments. The collected web data is represented using the RDF triples in the form of (s, r, and o). where 's' represents the subject, 'r' indicates predicate and 'o' represents objects. Generally, in sentiment analysis predicate words are user described context for example, about the event (i.e) opinion or sentimental comments about the event.

#### *2. Opinion Lexicon*

Opinion Lexicon (Liu, 2012) is a collection of positive and negative English words used for sentiment analysis. An opinion lexicon is a library of opinion words such as good, excellent, poor, and bad which are used to indicate polarity. Around 6800 words were collected and tabulated and each word is compared with other words in the table and each word is assigned a score based on the comparative study.

#### *3. Wolfs Position Updation using GWO Algorithm*

To apply the best search agent position update mechanism to the collected RDF triples, the following equations are used in the method.

$$D = |C.X_p(t) - A.X(t)| \tag{3}$$

$$X(t+1) = X_p(t) - A.D * \left(\frac{|SRO|}{|SR|}\right) \tag{4}$$

where X indicates the position vector of the wolves (i.e) user posts and $X_p$ is the position vector of the prey at time't' iteration. $|SR|$ is the number of triples whose subject $s \in S$ and predicate $r \in R$ are the same, and $|SRO|$ is the number of triples whose subject $s \in S$ and predicate $r \in R$ and object $o \in O$ are the same. In this problem each search agent is a new RDF triple. Each RDF triple impacts the wolf position, which exploits and explore the user's sentiments. Over the course of repeated procedures the probable position of the prey is estimated by, alpha, beta, and delta wolves.

#### *4. Computation of Overall Sentiment Score*

To compute each user's overall sentiment score *S*, the following equation is used.

$$S = \sum_{i=1}^{n} SC_i \tag{5}$$

$$SC = \sum_{i=1}^{n} (P(X_i) \,.\, (p_i - n_i)) \tag{6}$$

where n is the number objects in *O* for a predicate *P*, $SC_i$ is a sentiment score for object $o_i$ $(o_i \in O)$, $X_i$ is a position value for the RDF triple (user, Feeling, $o_i$). $P(X_i)$ is the position region among the other wolves and sum of all $P(X_i)$ always equal to 1. $(1 < i < n)$. $P(X_i)$ measures how much each $o_i$ affects overall sentiments. The terms $p_i$, $os_i$, and $n_i$ are positive, objective, and negative scores of the object $o_i$, respectively.

The value *S* approaches a positive value (or has a positive direction), if the user feels better than before, and vice versa. Note that we easily obtain the positive, objective, and negative scores from opinion lexicon. Each score has a range [-1, 1], and the sum of these three scores always equals 1. For example, the word 'happy' has 0.875 positive, 0.125 objective, 0 negative scores from Opinion Lexicon (Liu, 2012). In the proposed method, we only use two sentiment scores, positive and negative scores. The reason is that objective scores for the words are likely to mean neutral sentiments, and thus they are not helpful to decide each polarity of the words for sentiment analysis.

#### *5. Sentiment Trend Analysis*

By performing linear regression analysis, the direction of the user's sentiment is positive, and he gradually gets better feelings with time. Thus, we are aware of the user's sentiment trend, even though the user posts articles with various sentiments (feelings). In addition, we can analyze the sentiment trend of each user, without considering which emotional words the user uses in the same situation. The following algorithm updates the position values and overall sentiment scores, and then performs the sentiment trend analysis in the social web.

| **Algorithm** |
|---|
| **Input:** The collected RDF triples in the form of (*s, r, o*)<br>**Output:** Overall sentiment scores (*S*) table by date |
| **Method:**<br>1: construct a $m \times n$ matrix;<br>2: **for** all ($m_i$, $n_j$) elements in the $m \times n$ matrix<br>3: **if** ($j^{th}$ emotion word is used in $i^{th}$ RDF triple)<br>4: mark 'X' on ($m_i$, $n_j$) element;<br>5: **end if**<br>6: **end for**<br>7: **for** all $i^{th}$ column<br>8: number successively from 1 to $\|o_i\|$;<br>9: **end for**<br>10: update each $X_i$ with the given a at $m_i$;<br>11: extract a $p \times n$ sub-matrix from the $m \times n$ matrix;<br>12: compute overall sentiment scores S for each date;<br>13: generate overall sentiment scores values table by date;<br>14: visualize for sentiment and do trend analysis; |

In the algorithm, we use a two-dimensional matrix to mark which emotion word is used in the post (lines 1-4), and compute |SR| and |SRO| values for the given $o_i$. Next, the algorithm updates the position values $X_i$ by using the proposed (3) and (4).

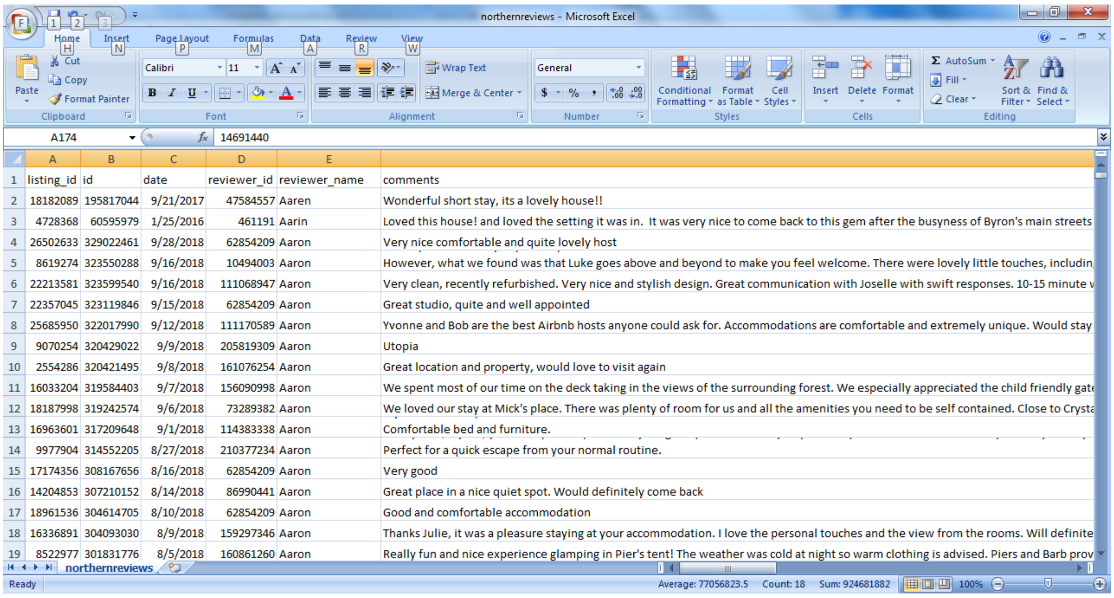
International Journal of Applied Engineering Research ISSN 0973-4562 Volume 14, Number 1 (2019) pp. 118-125
© Research India Publications. http://www.ripublication.com

To compute the overall sentiment scores by date, we then extract a p x n sub-matrix, whose rows and columns consist of p distinct dates and n distinct emotion words, respectively. Then, we compute overall sentiment scores for each date, referencing the obtained sentiment scores from opinion lexicon, and generate an overall sentiment score table by date. Finally, based on the generated *S* table, we are able to visualize and perform personal sentiment trend analysis on the social web.

## IV. EXPERIMENTAL RESULTS

The data set from the AIRBNB is taken for analysis [31]. In this data set information about various cities around the world were given by the customers in their feedback. By analyzing this dataset one can understand the tourism information about those cities. Out of the available data sets in the data base, the data set of the city Northern Rivers of New South Wales, Australia is taken for analysis.

In this data set detailed listings, detailed calendar listings, detailed reviewed data of the listings, summary of information(over the period of time is taken) and metrics for the listings, summary review data and listing id which is used for time-based analysis of the city Northern Rivers is available for analysis. Out of those single user's comments are taken for analysis. Data for the past three years starting from 2015 is taken for proving the experimental result. User sample data is given in the following screen shot in figure 1 .



| | A | B | C | D | E | |
|---|---|---|---|---|---|---|
| 1 | listing_id | id | date | reviewer_id | reviewer_name | comments |
| 2 | 18182089 | 195817044 | 9/21/2017 | 47584557 | Aaren | Wonderful short stay, its a lovely house!! |
| 3 | 4728368 | 60595979 | 1/25/2016 | 461191 | Aarin | Loved this house! and loved the setting it was in. It was very nice to come back to this gem after the busyness of Byron's main streets |
| 4 | 26502633 | 329022461 | 9/28/2018 | 62854209 | Aaron | Very nice comfortable and quite lovely host |
| 5 | 8619274 | 323550288 | 9/16/2018 | 10494003 | Aaron | However, what we found was that Luke goes above and beyond to make you feel welcome. There were lovely little touches, includin |
| 6 | 22213581 | 323599540 | 9/16/2018 | 111068947 | Aaron | Very clean, recently refurbished. Very nice and stylish design. Great communication with Joselle with swift responses. 10-15 minute v |
| 7 | 22357045 | 323119846 | 9/15/2018 | 62854209 | Aaron | Great studio, quite and well appointed |
| 8 | 25685950 | 322017990 | 9/12/2018 | 111170589 | Aaron | Yvonne and Bob are the best Airbnb hosts anyone could ask for. Accommodations are comfortable and extremely unique. Would stay |
| 9 | 9070254 | 320429022 | 9/9/2018 | 205819309 | Aaron | Utopia |
| 10 | 2554286 | 320421495 | 9/8/2018 | 161076254 | Aaron | Great location and property, would love to visit again |
| 11 | 16033204 | 319584403 | 9/7/2018 | 156090998 | Aaron | We spent most of our time on the deck taking in the views of the surrounding forest. We especially appreciated the child friendly gate |
| 12 | 18187998 | 319242574 | 9/6/2018 | 73289382 | Aaron | We loved our stay at Mick's place. There was plenty of room for us and all the amenities you need to be self contained. Close to Crysta |
| 13 | 16963601 | 317209648 | 9/1/2018 | 114383338 | Aaron | Comfortable bed and furniture. |
| 14 | 9977904 | 314552205 | 8/27/2018 | 210377234 | Aaron | Perfect for a quick escape from your normal routine. |
| 15 | 17174356 | 308167656 | 8/16/2018 | 62854209 | Aaron | Very good |
| 16 | 14204853 | 307210152 | 8/14/2018 | 86990441 | Aaron | Great place in a nice quiet spot. Would definitely come back |
| 17 | 18961536 | 304614705 | 8/10/2018 | 62854209 | Aaron | Good and comfortable accommodation |
| 18 | 16336891 | 304093030 | 8/9/2018 | 159297346 | Aaron | Thanks Julie, it was a pleasure staying at your accommodation. I love the personal touches and the view from the rooms. Will definite |
| 19 | 8522977 | 301831776 | 8/5/2018 | 160861260 | Aaron | Really fun and nice experience glamping in Pier's tent! The weather was cold at night so warm clothing is advised. Piers and Barb prov |

**Fig 1** Sample Data Set from AIR BNB

Sentiment analysis from the user's review decides the no customer's visit to the city for tourism. If the user feedback is positive during a particular period then no of visitors is increased to a considerable amount for tourism or business to the city Northern Rivers of New South Wales similarly and vice versa if the user feedback is negative the no of visiting person has been dwindled. If there are no reviews no of visitors follows the same pattern up to the observation of new trend. and moreover, due to the popularity and fame of airbnb it is necessary to check out the trending booking instantly.

It is surprise to learn the latest trend and predictions on the global scale to manage the listing to offer tip top customer service.The customer rating and reviews are the key to promote to retain more guests to grow more over and over. Approximately 150 million users across 190 countries has been attracted that shows an authentic and affordability to make right decisions to reveal interesting trends. The customer with travel experience thrive to discover an interesting trending spaces to lead to an end to end travel and hotel booking movement.

Each collected data is represented in the form of RDF triples. In this work, used the framework where the pre-processor is applied to the raw sentences which make it more appropriate to understand RDF Triples. It is a generic way for modelling of information in web resources mainly used in knowledge application process.it precisely defines a binary relationship between two resources called subject and predicate well identified by URL s.

The data sample of mentioned dataset in the proposed work details the review set of customers. The comparative trend analysis using different optimization gives detailed insight about the working nature of optimization in given dataset. The result is useful in turning the knowledge into decision. The following are the obtained sentiment trends by the Ant Colony Optimization and Grey Wolf Optimization.

International Journal of Applied Engineering Research ISSN 0973-4562 Volume 14, Number 1 (2019) pp. 118-125

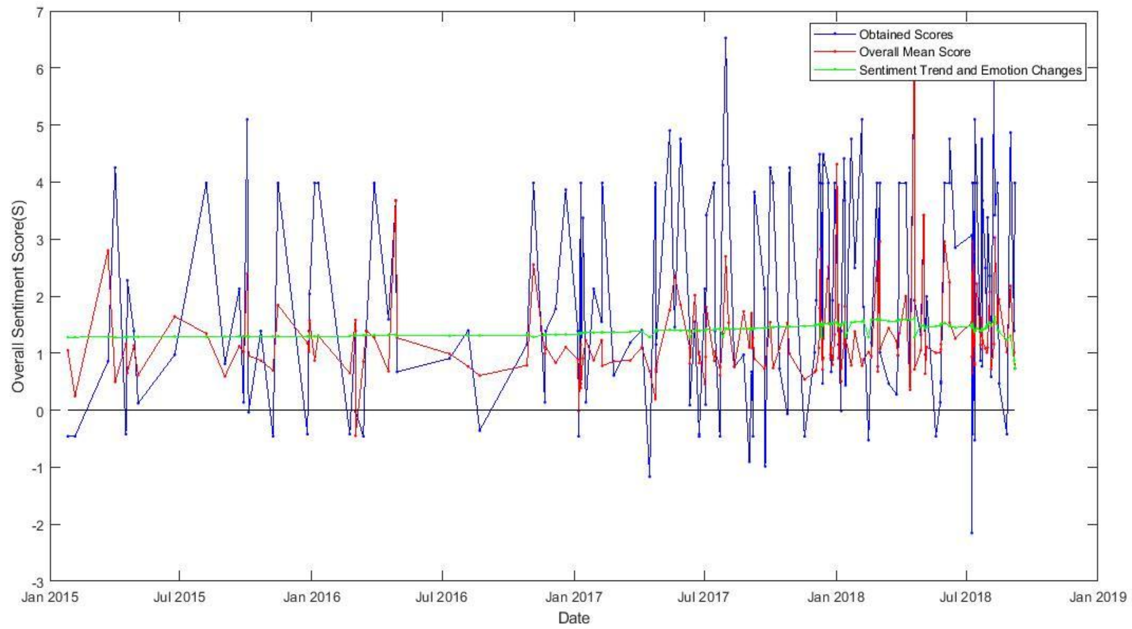




**Fig 2** Sentiment Analysis by Ant Colony Optimization

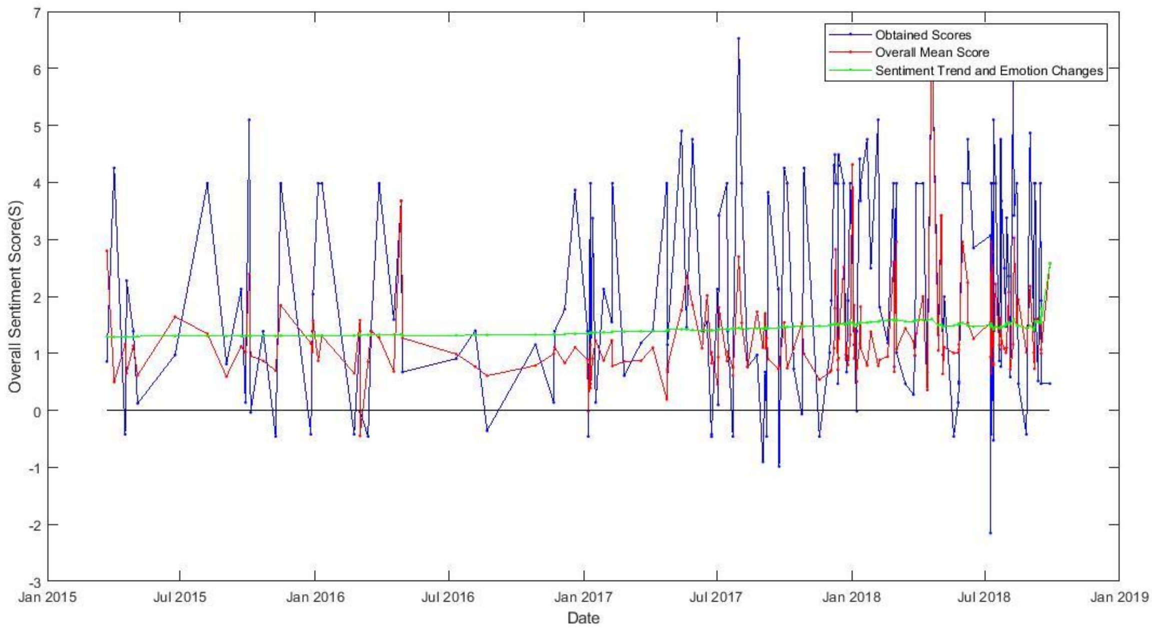



**Fig 3** Sentiment Analysis using Grey Wolf Optimization

Overall Sentiment scores are found out from word cloud. The following figure 4 shows some of the examples of user's emotions. Word Cloud is nothing group of words in an electronic image used in particular series of electronic text. Each word is of different size in such a way how frequently they are used in the text. From the given word cloud, we can analyze how many positive words, negative words and neutral words written by the user, shown in Figure 4. Word clouds stresses the most important concepts within text by making the most frequently used words larger fonts than less frequently used words in smaller fonts. The arrangement of words is not of much importance. Therefore, no sentiment is linked with the particular color. **Sentiment analysis** is necessary for understanding the context behind a word cloud. Following figure 4 is the word cloud that is taken for analyzing in this experiment.

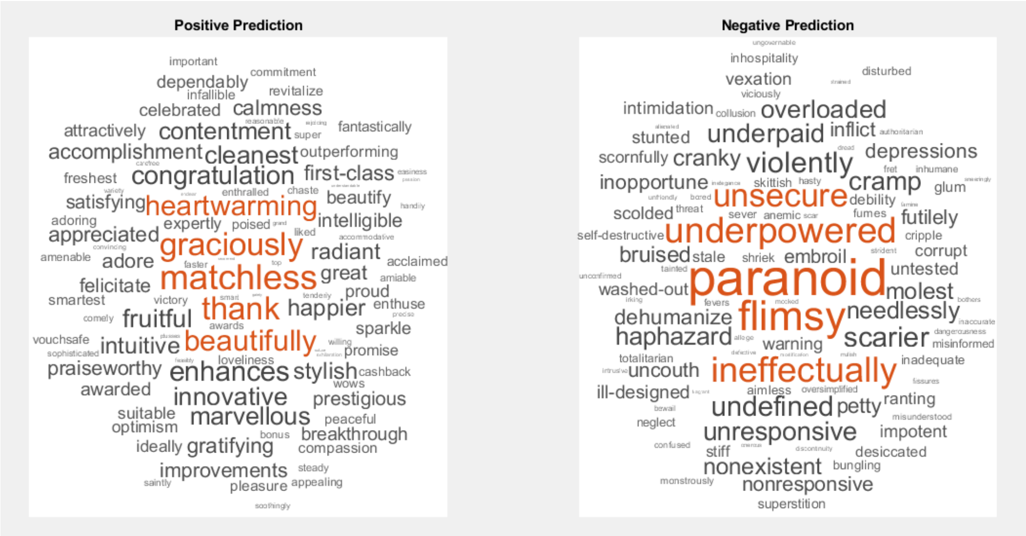



**Fig 4.** Word Cloud representation for the Obtained Sentiment Score

In the figure 3 Red line indicates mean swing of the sentiment scores over the period of time. Blue line indicates obtained sentiment scores for the posts written at a particular time. Green line indicates sentiment trend changes with respect to emotions change in the user. During the period of the Jan 2016 to Jan 2017 the no of posts is written minimum. But even though the sudden drop in the obtained scores was found exactly by proposed method and drop in sentiment trend analysis. The score zero indicates neutral. If the obtained sentiment score has positive scores then the emotions said to be positive and if the obtained sentiment score negative scores then the emotions said to be negative. Over all the emotions of the user is always positive and slightly negative at some times. Sentiment trend analysis is a straight line on positive axis which means that user almost is positive and goes slightly down when the obtained scores is negative at times. By having a close look of figure 2 and 3 around July 2018 the green color curve which depicts the sentiment analysis is almost flat in the figure 2 Ant Colony Optimization but those variations by obtained scores are very rapidly followed by the followed by variations in sentiment trend analysis in figure 3 by grey wolf optimization and after that downwards trend is followed by ant colony optimization in figure 2 and not able to pick up increase in obtained scores but grey wolf optimizations algorithm does that to perfection and in a faster manner in figure 3. From this we can understand that convergence rate of grey wolf optimization technique is faster than ant colony optimization algorithm.

## V. CONCLUSION

Sentiment trend analysis method using GWO algorithms. Here, user generated social data are collected in the form of RDF triples and then to digitize the amassed RDF triples, GWO algorithm are used. By using the proposed equations in GWO algorithm, position values are computed to extract trends of the user's sentiments. Then, user's overall sentiment score is computed with the computed position values and sentiment scores obtained from Opinion Lexicon. Finally, the trend of user's sentiments is analyzed by time. Since Grey wolf Algorithm is meta heuristic and exploits the leadership and hunting abilities of wolves it performs the search techniques better than the ant colony optimization algorithm. More over ant colony algorithm uses changes in probabilistic distribution change at every iteration and we cannot estimate the convergence time of ant colony optimization. The above disadvantages were overcome by grey wolf algorithm since also the limitation is rectified in proposed grey wolf optimizer in which the grey wolves have natural leadership mechanism. Further, this algorithm has only few input parameters and easy to implement, which makes it superior than earlier ones. For the verification of the proposed method, experiments are conducted. Also, to analyze the performance of this proposed system comparison is made with ACO algorithm.

## REFERENCES

[1] Liu, B. (2012). Sentiment analysis and opinion mining. Synthesis lectures on human language technologies, 5(1), 1-167.

[2] Collomb, A., Costea, C., Joyeux, D., Hasan, O. and Brunie, L., 2014. A study and comparison of sentiment analysis methods for reputation evaluation. Rapport de recherche RR-LIRIS-2014-002.

[3] dos Santos, C. and Gatti, M., 2014. Deep convolutional neural networks for sentiment analysis of short texts. In Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers (pp. 69-78).

[4] Fang, X. and Zhan, J., 2015. Sentiment analysis using product review data. Journal of Big Data, 2(1), p.5.

[5] Faris, H., Aljarah, I., Al-Betar, M.A. and Mirjalili, S., 2017. Grey wolf optimizer: a review of recent variants and applications. Neural Computing and Applications, pp.1-23.

[6] Gohil, S., Vuik, S. and Darzi, A., 2018. Sentiment Analysis of Health Care Tweets: Review of the Methods Used. JMIR public health and surveillance, 4(2).

[7] Guha, S., Chakraborty, T., Datta, S., Kumar, M. and Varma, V., 2016, May. TweetGrep: Weakly Supervised Joint Retrieval and Sentiment Analysis of Topical Tweets. In ICWSM (pp. 161-170).

[8] Hasan, A., Moin, S., Karim, A. and Shamshirband, S., 2018. Machine Learning-Based Sentiment Analysis for Twitter Accounts. Mathematical and Computational Applications, 23(1), p.11.

[9] Joshi, H. and Arora, S., 2017. Enhanced grey wolf optimization algorithm for global optimization. Fundamenta Informaticae, 153(3), pp.235-264.

[10] Kohli, M. and Arora, S., 2017. Chaotic grey wolf optimization algorithm for constrained optimization problems. Journal of Computational Design and Engineering.

[11] Korkontzelos, I., Nikfarjam, A., Shardlow, M., Sarker, A., Ananiadou, S. and Gonzalez, G.H., 2016. Analysis of the effect of sentiment analysis on extracting adverse drug reactions from tweets and forum posts. Journal of biomedical informatics, 62, pp.148-158.

[12] Kotelnikov, E.V. and Pletneva, M.V., 2016. Text sentiment classification based on a genetic algorithm and word and document co-clustering. Journal of Computer and Systems Sciences International, 55(1), pp.106-114.

[13] Krishna, A., Zambreno, J. and Krishnan, S., 2013, December. Polarity trend analysis of public sentiment on YouTube. In Proceedings of the 19th International Conference on Management of Data (pp. 125-128). Computer Society of India.

[14] Kwon, K., Jeon, Y., Cho, C., Seo, J., Chung, I.J. and Park, H., 2017, February. Sentiment trend analysis in social web environments. In Big Data and Smart Computing (BigComp), 2017 IEEE International Conference on (pp. 261-268). IEEE.

[15] Madhoushi, Z., Hamdan, A.R. and Zainudin, S., 2015, July. Sentiment analysis techniques in recent works. In Science and Information Conference (SAI) (pp. 288-291).

[16] Medhat, W., Hassan, A. and Korashy, H., 2014. Sentiment analysis algorithms and applications: A survey. Ain Shams Engineering Journal, 5(4), pp.1093-1113.

[17] Mirjalili, S., Mirjalili, S.M. and Lewis, A., 2014. Grey wolf optimizer. Advances in engineering software, 69, pp.46-61.

[18] Nielsen, F.Å., 2011. A new ANEW: Evaluation of a word list for sentiment analysis in microblogs. arXiv preprint arXiv:1103.2903.

[19] Precup, R.E., David, R.C., Szedlak-Stinean, A.I., Petriu, E.M. and Dragan, F., 2017. An easily understandable grey wolf optimizer and its application to fuzzy controller tuning. Algorithms, 10(2), p.68.

[20] Rosenthal, S., Farra, N. and Nakov, P., 2017. SemEval-2017 task 4: Sentiment analysis in Twitter. In Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017) (pp. 502-518).

[21] Saremi, S., Mirjalili, S.Z. and Mirjalili, S.M., 2015. Evolutionary population dynamics and grey wolf optimizer. Neural Computing and Applications, 26(5), pp.1257-1263.

[22] Teeparthi, K. and Kumar, D.V., 2016, December. Grey wolf optimization algorithm based dynamic security constrained optimal power flow. In Power Systems Conference (NPSC), 2016 National (pp. 1-6). IEEE.

[23] Vinothini, J. and Ashok Bakkiyaraj, R., 2012. Grey Wolf Optimization Algorithm for Colour Image Enhancement Considering Brightness Preservation Constraint. In IEEE Conf Comput Vis Pattern Recognit (Vol. 482, pp. 1234-1241).

[24] Xu, H., Liu, X. and Su, J., 2017, September. An improved grey wolf optimizer algorithm integrated with Cuckoo Search. In Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), 2017 9th IEEE International Conference on (Vol. 1, pp. 490-493). IEEE.

[25] Yassien, E., Masadeh, R., Alzaqebah, A. and Shaheen, A., 2017. Grey Wolf Optimization Applied to the 0/1 Knapsack Problem. International Journal of Computer Applications, 169(5).

[26] Zhang, C. and Wang, J., 2013. The sentiment trend analysis of twitter based on set pair contact degree. International Journal of Computer Science Issues (IJCSI), 10(1), p.798.

[27] Li, S., Lee, S. Y. M., Chen, Y., Huang, C. R., & Zhou, G. (2010, August). Sentiment classification and polarity shifting. In Proceedings of the 23rd International Conference on Computational Linguistics (pp. 635-643). Association for Computational Linguistics.

[28] Ingle, A., Kante, A., Samak, S., & Kumari, A. (2015). Sentiment analysis of twitter data using hadoop. International Journal of Engineering Research and General Science, 3(6).

[29] Sahayak, V., Shete, V., & Pathan, A. (2015). Sentiment analysis on twitter data. International Journal of Innovative Research in Advanced Engineering (IJIRAE), 2(1), 178-183.

[30] Suchdev, R., Kotkar, P., Ravindran, R., & Swamy, S. (2014). Twitter Sentiment Analysis using Machine Learning and Knowledge-based Approach. International Journal of Computer Applications, 103(4).

[31] http://insideairbnb.com/get-the-data.html