IJCTA, 9(40) 2016, pp. 327-334

Temporal Sentiment Analysis: A Review

Akanksha Puri, Sukhnandan Kaur and Rajni Mohana

ABSTRACT

Natural language processing is a most common area of research that probe how computer understand and manipulate natural language text or speech to do useful things. Sentiment analysis is one of the active research area in natural language processing. Sentiment analysis is process of determining the emotional tone behind a series of words, identifying and categorizing opinions expressed in a entity, in order to derive whether the writer's attitude towards a areas in natural language processing. Sentiment analysis based on temporality is gaining much attention in many real time applications. This manuscript highlights the key concepts of various state-of-art temporal sentiment analysis along with the research gaps. It also put focus on the normalization of various temporal expressions. It covers different temporal expressions and the methods for normalization of these expressions. The main focus is to find the research gaps in temporal sentiment analysis. To facilitate the future work, a discussion of state-of-art resources and methods for temporal sentiment analysis is also provided.

Keywords: Natural language processing, Temporal sentiment analysis, Opinion mining, Temporal expressions, Normalization.

I. INTRODUCTION

Natural language processing is a field of computer science and artificial intelligence appertained with the interactions between computer and natural languages. It is sometimes also known as an AI-complete problem. Natural language processing is a most common area of research that probe how computer understand and manipulate natural language text or speech to do useful things [1].

There are many tasks that come under natural language processing like sentiment analysis, mane-entity-recognization, translation, summarization and many more. Sentiment analysis is one of the active research area in natural language processing and text mining. Opinion mining is another name given to Sentiment analysis. Sentiment analysis is process of determining the emotional tone behind a series of words, identifying and categorizing opinions expressed in a entity, in order to derive whether the writer's attitude towards a areas in natural language processing.

In the recent years, temporal tagging has acknowledged rising study in the area of natural language. Temporal tagging is an area of Sentiment analysis along with natural language. Temporality is the state of actual within or having some relationship with time. It is traditionally the linear sequence of past, present and future. It plays a vital role in sentiment analysis to calculate sentiment strength of any entity. Temporal sentiment analysis [2] is Aggregate sentiment communicate at different points in time and perform trend analysis by looking at how sentiment changes over time.

The remainder of this paper is planned as follows. Section 2 highlights the related work. Section 3 describes the Temporal Sentiment Analysis. Section 4 describes the temporal expressions in document and their normalization. Section 5 discusses the applications of temporal sentiment analysis and components used in temporal sentiment analysis are described in section 6. Section 7 describes the performance metrics. In section 8 Research gap is discussed and finally in section 9 we have concluded the temporal sentiment analysis.

* Department of Computer Science and Engineering, Jaypee University of Information Technology, Solan, Himachal Pradesh, India, *E-mail: akankshapuri1994@gmail.com*

II. RELATED WORK

In a natural language processing, many researchers work on Temporal Sentiment analysis. Tomohiro et al. [3] worked on temporal sentiment analysis and proposed two graph methods: Topic graph and Sentiment graph. They have analyzed the temporal trend of sentiment and topic from text with timestamps. Yoonjung et al. [4] described a framework for sentiment analysis for generating context driven features. They proposed methods for in co-operate two tasks: sentiment generation and Sentiment classification. Some researchers put efforts to linking text sentiment to public opinion time series. Brendan et al. [5] proposed various methods for analysis of text. They measured the public opinion derived from polls with sentiment measured with text from many popular micro blogging sites. Now, many researchers or scholars start determining the temporal sentiment analysis on social media, they effectively use twitter information for their work. Mike et al. [6] determined the sentiment in twitter events. They described three approaches for sentiment analysis in twitter: Full-text machine learning, Lexicon-based methods and Linguistic analysis. Jannik et al. [7] analyzed the sentiment or text on different domains. They used TIMEX2 attributes for normalized the four type of temporal expressions: DATES, TIME, DURATION, SETS. They had created colloquial and scientific corpus and compare them with news and narrative. Tadahiko et al. [8] proposed a system for visualizing twitter users based on temporal changes in impressions. They proposed a web application system for visualizing twitter.

Andre et al. [9] described methods for spatial and temporal sentiment analysis and sentiment polarity. They implemented two approaches for sentiment classification of tweets and compare their results: 1) SVM 2) Naïve Bayes. Gael et al. [10] proposed method to build a temporal ontology which contributes to time related applications. They had used Synsets and categorized them into present, past and future. Yuanyuan et al. [11] worked on mapping of dense Geo-tweets and web pages.They had described many mapping functions such as acquisition of tweets, filtering, acquisition of web pages and clustering of tweets.

Luciano et al. [12] discussed about the happiness level of people in a city. They measured the sentiment expressions in tweet that are posted from popular areas in city. Syed et al. [13] done work on Localized twitter opinion mining using sentiment analysis and had discussed a methodology which allows utilization and interpretation of twitter data to determine public opinions. They had analyzed the gender of user. Now, some researchers start working on Temponyms tagging. Edral et al. [14] describes the integration of wide range temponyms. The authors had detected temporal expressions and make their sentiments accessible by normalizing them into standard formats. For tagging they used three data sources to create large collection of explicit temponyms with their temporal values: 1) Yago 2) Aida 3) Temponyms pattern directory.

III. TEMPORAL SENTIMENT ANALYSIS

Sentiment analysis is the part of Natural Language Processing, in which the opinion of user is detected and classification of attitudes in texts are involved [22]. For the classification of sentiment, different approaches or methods can be used such as supervised learning, unsupervised learning and semi-supervised learning. Nowadays, time is one of the key feature that conclude a document reliability besides importance, correctness, impartiality and exposure. Temporal expressions are usually anchored by events and attributes of the temporal expressions are useful in distinguish between temporal relation of the events and ordering of the events. NLP and Information retrieval indicates diûerent title of analysis for temporality. NLP aims to understand time at a ûne grained level. From the IR viewpoint, temporality has been studied at a coarse-grained level.

TempoWordNet is an essential source for time related applications both in NLP and IR. Temporality means having some relationship with time. It is traditionally the linear sequence of past, present and future. In last few years, temporality has accepted increased attention in natural language processing and information retrieval. In temporal information retrieval, text document contains the systematic and algorithmic protocol which involves steps for the creation, extraction and normalization. In text documents two standards are

Table I
Comparison between Related Works

Ref No.	*Paper Name*	*Work Done*	*Limitation*
1.	Sentiment in Twitter Events.	• Authors used three approaches for sentiment analysis: Full-text machine learning, Lexicon -based and linguistic analysis. • They classify the text as positive or negative.	1 month include only two special events the Oscars and the Olympics. The sentiment strength algorithm used is also an issue.
2.	TempoWordNet for sentence time tagging.	• The authors had used temporal synset having 21 sets and categorized the sentences into past, present future tense. • They used two classification method and one classification method for analyze the word as temporal and atemporal.	one-step process shows incapacity to solve the temporality issue.
3.	Dynamic Mapping of Dense Geo - Tweets and Web Pages based on Spatio-Temporal Analysis.	• In this tweets are related to its category name. They have used various mapping function such as acquisition of tweets, tweet filtering and acquisition of web pages. • Various methods for categorization of tweets: • K-NN method, Naive bayes method,Support vector machine. RMSE and average score value of tweets are calculated	Users cannot obtain the most recent information, while they browse web pages since these are not updated in real time.
4.	Temporal tagging on different domains: Challen ges, strategies and Gold standards.	• They had created colloquial and scientific corpus and compare them with news and narrative. • They analyzed the temporal tagging on different domains and addressing the different challenges.	
5.	Understanding Sentiment of People from News Articles: Temporal Sentiment Analysis of Social Events.	• Temporal sentiment analysis: The method takes texts with timestamps such as Weblog and news articles as input, and creates two types of graphs: a)topic graph b) sentiment graph	They did not proposed method for Weblog articles.
6.	Temponyms Tagging: Temporal Scopes for Textual Phrases.	• They had detected temporal expressions and make their semantics accessible by normalizing them into standard format. • For temponyms tagging, they combine three data sources to create a large collection of explicit temponyms with their temporal scopes: YAgo, Aida, Temponyms pattern directory.	They only addressed explicit temponyms.
7.	A spatial and temporal sentiment analysis approach applied to twitter microtexts.	• The author implemented the Classification of Sentiment Polarity by using two approaches and then comparing their results: 1.SVM 2.Naive Bayes • They detected the geographical location of tweets.	They concerned with the identification of the entity referred to by the opinion detected in tweets

contd. table I

8.	Domain specific sentiment analysis using contextual feature generation	• Classify a candidate word into positive, negative and neutral category using bootstrapping method. • They proposed a domain -specific sentiment analysis system utilizing context features in news texts.	System was not able to capture phrase-level clues. i.e the units of the clues need to be expend
9.	Visualizing temporal changes in impressions from tweets.	• They proposed a web application system for visualizing twitter users based on temporal changes in the impressions from the tweets posted by users.	This system is not suitable for the impressions for tweets that used the keywords.
10.	Localized twitter opinion mining using sentiment analysis	• The authors had discussed a methodology which allows utilization and interpretation of twitter data to determine public opinions. • They had analyzed the gender of user.	Quality of tweets was also very low. In gender classification there is still some errors.

used for annotating temporal information: TIDES TIMEX2 [19] and TimeML [20]. Both standards immediate guideline for how to verify the extents and how the value of temporal expressions is normalized.

The following attributes are used in annotation for normalization:

VAL: For temporal information it is a normalized form of the expressions in ISO standards

MOD: Modiûer of temporal expressions.

ANCHOR VAL: It is a normalized form of an anchoring date or time

ANCHOR DIR: the relation direction between VAL and ANCHOR VAL

SET: It is used to identify expressions denoting sets of times.

From these attributes the four types of temporal expressions dates, times, durations and sets can be normalized. The value attribute of date and time expressions directly pertain to a point in time.

For Example: “2016-03-15” for the expression “March 15, 2016”.

For durations and set expressions, it covers the length of the time interval, e.g., “P5D” for “Five days” and “every five days” [8]

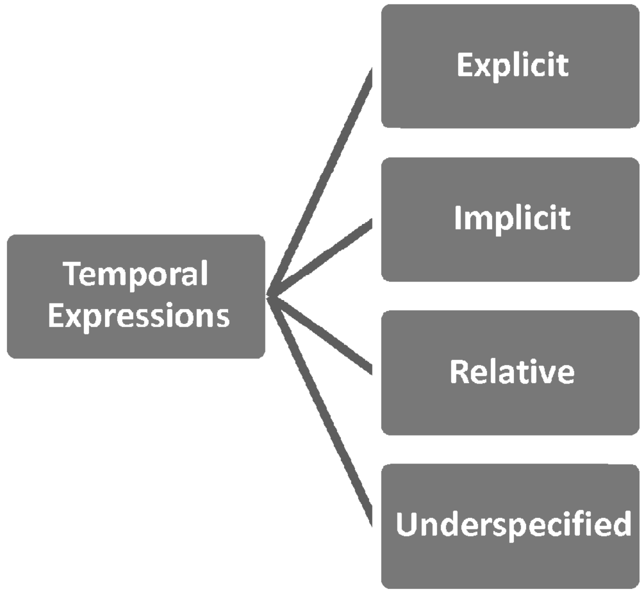
Thus, Temporal sentiment analysis analyzes temporal direction of sentiments and topics from a text documents that has timestamps. It is the relation of time with sentiment. It is useful in detecting the mood of user in different time frames. Temporal sentiment analysis is used in many research areas.

TEMPORAL EXPRESSIONS IN DOCUMENTS AND THEIR NOMALIZATION

TimeML, the standard mark-up language for temporal annotation, we have four types of temporal expressions: dates (Dec 3, 2016), times (7 p.m.), durations (one weeks), and sets (daily).

(a) Types of Temporal Expression

There are four kinds of temporal expressions: Explicit, Implicit, Relative Expression and Underspecified Expression as shown in figure. In temporal tagging each document is associated with a domain that is to be processed and pertain domain specific strategies for extraction and normalization of temporal expressions.



Explicit: In explicit expression, the temporality is captured through the metadata.

For example: *Explicit expression is referred as "25/November/2016" or "25-11-2016".*

Implicit: In implicit expressions, the temporality of the document is hidden in the document itself. It can be captured by applying some language rules over the text.

For example: *Implicit expression sometimes written as "Children's Day 2016".*

Implicit expressions are normalized by using knowledge about their meaning.

Relative: These expressions are those which form a specific relation with a temporal focus.

For example: *Relative expressions may written as "after three days" or "Yesterday".*

For the normalization of relative expression identification of reference time is required.

Underspecified: It is an expression where users neglect the elements of information, which then have to be improved from the context by using some language rules.

For example: *This expression sometimes written as "in June".*

To normalize the above expression, it is required to identify the temporal relation to the reference time. In the News and Narrative style document corpora, it is simple to identify the reference time. In these documents the reference time is mostly equals to DCT (Document creation time). [8]

IV. APPLICATION OF TEMPORAL SENTIMENT ANALYSIS

Recently many researchers find that online texts analysis can be helpful for trend or event prediction.

- *Forecasting Political Results*: Political orientations [16] are presage from the reviews and comments of users. Users post their ideas or opinions about each politician or party over the web sphere.
- *Mood Detection*: In sentiment analysis, time plays an important role in mood detection. Temporal sentiment analysis describes the mood swings of users. This is useful in predicting the behavior of user.

- *Popularity/ Infamy*: In [17] On the basis of sentiment analysis the achievement of box office is predicted. From the reviews or comments on movie we can rate the movie or drama.
- *Stock Market:* In [18] Stock markets movement is read from the news article sentiment or mood of twitter users.

V. COMPONENTS USED IN TEMPORAL SENTIMENT ANALYSIS

Temporal sentiment analysis is related to time with sentiment. There are various components which are used for analysis of sentiment. Some of the components are given as:

WordNet: WordNet sometimes also known as ontology. WordNet is any of the machine understandable databases of information about words/text for the English language. It assort words into sets of similar words or words having same meaning called synsets and narrative a number of relations among synonym sets.[24]

SentiWordNet: It is sentiment lexicon companion sentiment information to each WordNet synset. It is the combination of WordNet and sentiment information. SentiWordNet categorized each word in synset of WordNet into three categories: Positive Score, Negative Score and Objectivity Score.

TempoWordNet: It is a set of time-perspective synsets. TempoWordNet is a free verbal knowledge base for temporal analysis where each synset has its own intrinsic temporal value. TempoWordNet classify each synset of WordNet into four categories: Atemporal, past, present and future.[25]

Metadata: Metadata is data that depicts other data. Metadata is structured information that delineates, explain, detect, easier to extract and retrieve, use, manage an information. We use metadata for discovering useful and relevant information about each resource.

VI. PERFORMANCE MATRICS

There are the following metrics that are used to take measure of the system performance:[15]

To calculate Precision, Recall, F-Measure and Accuracy of the system following performance metrics need to be defined:[21]

True positives (TP): Number of positive examples classify as positive.

False positives (FP): Number of negative examples classify as positive.

True negatives (TN): Number of negative examples classify as negative.

False negatives (FN): Number of positive examples classify as negative.

1. Precision: The precision is the ratio of the number of relevant documents returned to the total numbers of documents for a given user query.

$$\text{Precision} = \frac{TP}{TP+FP}$$

2. Recall: The recall of a text can be defined as the ratio of the number of relevant documents returned to the total number of relevant documents for the user query in the set.

$$\text{Recall} = \frac{TP}{TP+FN}$$

3. F-Measure: F-Measure is defined as the measure that combines precision and recall is the melodic mean of precision and recall.

$$\text{F-Measure} = 2\left[\frac{precision.recall}{precision + recall}\right]$$

4. Accuracy: It is defined as the ratio of the number of correctly classified objects to the total number of objects.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

VII. RESEARCH GAP

Temporal Aspect is another important dimension of Semantic Analysis. Detecting attitude shifts is made possible identifying opinions and identifying when they were stated. Monitoring the impact of marketing campaigns or containing damage to brands and companies through quick response are few of the important applications of temporal dimension. Therefore, need of automatic assignment of high weightage to the recent reviews and low to the previously posted blog, review, etc is required Following are the fields where semantic analysis are used: [21]

1. *Geographical distribution of time*: The task of sentiment analysis based on the temporal aspect becomes more complex because of deviation in time worldwide. Hence, it is very monotonous to get a collaborative real-time sentiment of the whole world.
2. *Explicit/implicit text document*: Most text documents are given explicitly in the query or deduced implicitly. Extraction of topic implicitly is a best but complex option.
3. *Forecast analysis*: Researchers are working to improve forecast analysis by including the time in the process of sentiment analysis.
4. *Hinged weightage to reviews with respect to time*: As the time passes, the value of previous reviews is degrading. The present day review is more important. In Sentiment analysis to have time-oriented review the process should include joint weightage with respect to time.

VIII. CONCLUSION

Temporal sentiment analysis is an evolving area with a variety of real time applications. Although sentiment analysis is a challenging task, much attention is paid to it over the last decade. The rising need of real time applications gives temporality based sentiment analysis more valuable. Normalization makes the processing of web data more efficiently for sentiment analysis. Apart from the state-of-art temporal sentiment analysis, the research gaps described in this manuscript needs to be covered for increasing the performance of the analyzer.

REFRENCES

[1] Chowdhury GG. “Natural language processing.” Annual review of information science and technology, 37(1), 51-89, Jan 1 2003

[2] Preethi PG, Uma V. “Temporal sentiment analysis and causal rules extraction from tweets for event prediction.” Procedia Computer Science, 48, 84-89, Dec 31 2015.

[3] Dias GH, Hasanuzzaman M, Ferrari S, Mathet Y. “Tempowordnet for sentence time tagging.” In Proceedings of the 23rd International Conference on World Wide Web, Apr 7 2014, pp. 833-838.

[4] Fukuhara T, Nakagawa H, Nishida T. “Understanding Sentiment of People from News Articles: Temporal Sentiment Analysis of Social Events.” In ICWSM 2007.

[5] Choi Y, Kim Y, Myaeng SH. “Domain-specific sentiment analysis using contextual feature generation.” In Proceedings of the 1st international CIKM workshop on Topic-sentiment analysis for mass opinion, 2009 Nov 6, pp. 37-44.

[6] O'Connor B, Balasubramanyan R, Routledge BR, Smith NA. "From tweets to polls: Linking text sentiment to public opinion time series." ICWSM, 11, 2010 May 23, 122-129.

[7] Thelwall M, Buckley K, Paltoglou G. "Sentiment in Twitter events." Journal of the American Society for Information Science and Technology, 62(2), 406-418, Feb 1 2011.

[8] Strötgen J, Gertz M. "Temporal Tagging on Different Domains: Challenges, Strategies, and Gold Standards." In LREC, Vol. 12, pp. 3746-3753, 2012.

[9] Kumamoto T, Wada H, Suzuki T. "Visualizing Temporal Changes in Impressions from Tweets." In Proceedings of the 16th International Conference on Information Integration and Web-based Applications & Services, Dec 4 2014, pp. 116-125.

[10] Alves AL, de Souza Baptista C, Firmino AA, de Oliveira MG, de Paiva AC. "A Spatial and Temporal Sentiment Analysis Approach Applied to Twitter Microtexts." Journal of Information and Data Management, 6(2), 118, Jan 20 2016.

[11] Wang Y, Yasui G, Kawai Y, Akiyama T, Sumiya K, Ishikawa Y. "Dynamic mapping of dense geo-tweets and web pages based on spatio-temporal analysis." In Proceedings of the 31st Annual ACM Symposium on Applied Computing 2016 Apr 4 pp. 1170-1173.

[12] Gallegos L, Lerman K, Huang A, Garcia D. "Geography of Emotion: Where in a City are People Happier?." In Proceedings of the 25th International Conference Companion on World Wide Web 2016 Apr 11, pp. 569-574.

[13] Hridoy SA, Ekram MT, Islam MS, Ahmed F, Rahman RM. "Localized twitter opinion mining using sentiment analysis." Decision Analytics, 2(1), 1, Oct 22 2015.

[14] Kuzey E, Strötgen J, Setty V, Weikum G. "Temponym Tagging: Temporal Scopes for Textual Phrases." In Proceedings of the 25th International Conference Companion on World Wide Web 2016 Apr 11, pp. 841-842.

[15] Junker M, Hoch R, Dengel A. "On the evaluation of document analysis components by recall, precision, and accuracy." In Document Analysis and Recognition, 1999. ICDAR'99. Proceedings of the Fifth International Conference on 1999 Sep 20, pp. 713-716.

[16] Park, S., Ko, M., Kim, J., Liu, Y., Song, J. "The politics of comments: predicting political orientation of news stories with commenters' sentiment patterns." in: Proceedings of the ACM Conference on Computer Supported Cooperative Work. ACM, 2011, pp. 113–122.

[17] Schuller, B., Knaup, T. "Learning and knowledge-based sentiment analysis in movie review key excerpts." In: Toward Autonomous, Adaptive, and Context-Aware Multimodal Interfaces. Theoretical and Practical Issues. Springer, 2011, pp. 448–472.

[18] Li, X., Xie, H., Chen, L., Wang, J., Deng, X., 2014. " News impact on stock price return via sentiment analysis." Knowl.-Based Syst, 69, 14– 23, 2014.

[19] Lisa Ferro, Laurie Gerber, Inderjeet Mani, Beth Sundheim, andGeorgeWilson. "Standard for the Annotation of Temporal Expressions." TIDES, The MITRE Corporation, 2005.

[20] James Pustejovsky, Robert Knippen, Jessica Littman, and Roser Sauri. "Temporal and Event Information in Natural Language Text." Language Resources and Evaluation, 39(2-3), 123–164, 2005.

[21] Kaur, Sukhnandan, and Rajni Mohana. "A roadmap of sentiment analysis and its research directions." International Journal of Knowledge and Learning10.3, 296-323, 2015.

[22] Balahur, Alexandra, and Marco Turchi. "Comparative experiments using supervised learning and machine translation for multilingual sentiment analysis." Computer Speech & Language 28.1 (2014): 56-75.

[23] Fellbaum, Christine. "WordNet and wordnets." In: Brown, Keith et al. (eds), Encyclopedia of Languages and Linguistic, Second Edition, Oxford: Elsevier, 665-670, 2005.

[24] Baccianella, Stefano, Andrea Esuli, and Fabrizio Sebastiani. "SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining." *LREC*. Vol. 10. 2010.

[25] HULTECH, GREYC- CNRS 6072 Laboratory, Normandie University, Caen, France, January, 28, 2015.