A Framework for Opinion Mining- Root Cause Monitoring the Social System Which Keeps Tracks of Weak Points in Rural Areas

Suman A Lasrado[1], Dr. Vincent Mascarenhas[2]   , Pradeepkumar[3], Abdul Jamali[4]

Assistant Professor, Department of IT, SIG – Open Source Software, AIMIT, Beeri, Mangalore, India[1]

Professor, Department of IT, SIG – Open Source Software, AIMIT, Beeri, Mangalore, India[2]

Member, SIG – Open Source Software, AIMIT, Beeri, Mangalore, India[3]

Member, SIG – Open Source Software, AIMIT, Beeri, Mangalore, India[4]

ABSTRACT: The advent of Web 2.0 and social media content has stirred much excitement and created abundant opportunities for understanding the opinions of the general public and consumers toward social events, political movements, company strategies, marketing campaigns, and product preferences. Opinion mining, a sub discipline within data mining and computational linguistics, refers to the computational techniques for extracting, classifying, understanding, and assessing the opinions expressed in various online news sources, social media comments, and other user-generated content. Sentiment analysis is often used in opinion mining to identify sentiment, affect, subjectivity, and other emotional states in online text.

In this paper the idea is to keep track of all the weak points in a system. A constant monitoring system will keep track of all these failure points. As more and more weak points are gathered, then the criticality ration of the problem can be calculated and the management alerted of the impending problem. We all wait for disasters to happen and then respond to it. This can also be used by traffic management and other social welfare initiatives.

KEYWORDS: Opinion Mining, Sentiment Analysis, Subjectivity classification, semantic orientations
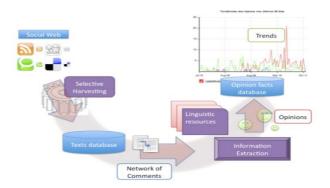
I. INTRODUCTION

Opinion mining is the area of research that attempts to make automatic systems to determine human opinion from text written in natural language. It aims to extract opinions from information sources such as reviews and present them to the user in a user friendly manner [1].What other people think has always been an important piece of information for most of us during the decision-making process. Opinion mining draws on computational linguistic, information retrieval, text mining, natural language processing, machine learning, statistics and predictive analysis. Textual information in the world can be broadly classified into two main categories, facts and opinions. Facts are objective statements about entities and events in the world. Opinions are subjective statements that reflect people's sentiments or perceptions about the entities and events. Much of the existing research on text information processing has been focused on mining and retrieval of factual information[2]. The research in the field started with sentiment and subjectivity classification, which treated the problem as a text classification problem. Sentiment classification classifies whether an opinionated document or sentence expresses a positive or negative opinion[2]. Subjectivity classification determines whether a sentence is subjective or objective[2].Many real-life applications, however, require more detailed analysis because users often want to know the subject ofopinions[4].

Research on opinion mining started with identifying opinion (or sentiment) bearing words, e.g., great, amazing, wonderful, bad, and poor.The next major development is sentiment classification of product reviews at the document level [10, 11, 16]. The objective of this task is to classify each review document as expressing a positive or a negative sentiment about an object.

Finding opinion sources and monitoring them on the Web is also a formidable task because a large number of diverse sources exist on the Web and each source also contains a huge volume of information. In many cases, opinions are hidden in long forum posts and blogs. It is very difficult for a human reader to find relevant sources, extract pertinent sentences, read them, summarize them and organize them into usable forms[2]. In this paper we are introducing aframework which will keep track of all these failure points, then the criticality ration of the problem can be calculated by gathering the weak points and the management alerted of the impending problem.

II. RELATED WORK

In this section, we present a review of the existing and related works on Opinion Mining (OM) and Sentiment Classification (SC) proposed in the literature. Most of the existing research like information retrieval, Web search, and other text mining and natural language processing tasks on text information processing has been focused on mining and retrieval of factual information. Only a little work has been done on the processing of opinions until recently. Though the number of research interest in this area is growing fast[1].

Users or consumers don't know what to do when there is a theft or burglary in their society and before they react the disaster already takes place .Opinion Mining suffers from several different challenges, such as determining which segment of text is opinionated, identifying the opinion holder, unable to reach out to the rural areas etc.Our methodology helps to overcome this problem by identifying the opinion holder through security code that is provided to the user once he registers into the system. When the user reports a problem, the system tracks the location of the users and alerts the responsible authorities to take quick action.

III. SIGNIFICANCE OF THE STUDY

As per the proposed system, the user will have to register on to the system. Once the user provides his/her details, a unique code will be generated and sent back to the user. The user will use the code sent back to him/her next time when he/she has to make use of the application i.e., the user can register beforehandss itself and can directly register a complaint regarding any issue rather than running around at the last minute in search of a particular phone number. A copy of this software will be available with all the concerned departments and they will be notified immediately and the concerned department will take the necessary action as soon as possible.

The proposed study educates the people to provide valuable information towards this application by conducting survey in rural areas. This application could also be used by getting survey results from certain areas. But this will require people to go around and collect that information. All data needs to be input by that team. This can also be done in a more efficient way. School can take an initiative to do surveys and keep the municipal agencies aware of certain problems becoming more frequent in certain areas. This can be distributed to all the schools and the students can send in SMS messages reporting these problems. Mobile phones tracking database can be used to track the mobile phones

and find out the location. If there is a theft happening people can inform the police authorities immediately. Easy to find the opinions of the people in the rural areas and take a quick action.

IV. METHODOLOGY

This methodology helps us to come up with a data collection plan. Narrow down on the concentration of certain problem areas / Critical areas or types of problems (this could be new ones which is a spin-off from the current types of data that is gathered).Customer complaints about certain services. Traffic congestion stats and control. Pest control & Health - The authorities will reach the destination of the accident on time. Sanitary logistics – collection and disposal of harmful by products and disposition process.Burglary and thefts tracking and prevalence.It is easy to reachout to the rural areas by using this methodology. People residing in rural areas need not go searching for help during any disaster, this application provides instant help. This analysis acts as the mediator between two people, it communicates and tries to solve the problem.

V. SECURITY

Application security encompasses measures taken throughout the application's life-cycle to prevent exceptions in the security policy of an application or the underlying system vulnerabilities through flaws in the development and maintenance of the application. Applications only control the use of resources granted to them, and not which resources are granted to them. They, in turn, determine the use of these resources by users of the application through application security.

This application provides security to the customers by providing a security code while they register. Once the customer registers onto the system via a computer system, an email will be sent back. The email will contain a code which will be used by the user. This will prevent unauthorized users from accessing this system and no one can send spam email messages. In case of areas where computers are not available, the users can make use of cell phones to use this application. The procedure will be the same but instead of an email, a SMS will be sent to the cell phone number. This will help to track the location faster. This methodology tries to avoid burglary in the society by helping the cops to take immediate action on such issues. The system also provides Health protection.

VI. CONCLUSION

This paper presents a effective framework for opinion mining that can process data by building the Data storage and processing it. Opinion mining is the study of people's opinions and emotions toward entities, events and their attributes. In the past, when an organization wanted to find opinions of the general public about its products and services, it conducted surveys and focus groups. This application collects the opinion through messages, it sends the data regarding the issue to the main authority which will take the initiative immediately. The idea is to keep track of all the weak points in a system and to take action before the disaster happens. The research has been carried out towards finding out the opinion of the people in rural areas and react to the situation instantly. Thus the technique is highly effective and performs the methods significantly

REFERENCES

[1] Dipali V. Talele ME[CSE],GHRIEM, Chandrashekhar D. Badgujar Asst. Prof. in Dept. of CSE,   GHRIEM "The Art of Opinion Mining and Its Application Domains: -A Survey"at International   Conference in Recent Trends in Information Technology and Computer Science (ICRTITCS - 2012).
[2] Ding, X., Liu, B. and Yu, P. A Holistic Lexicon-Based Approach to Opinion Mining.Proceedings ofthe first ACM International Conference on Web search and Data Mining (WSDM'08), 2008.
[3]Ganapathibhotla, G. and Liu, B. Identifying Preferred Entities in Comparative Sentences. To appear in Proceedings of the 22nd International Conference on Computational Linguistics (COLING'08), 2008.
[4] Hatzivassiloglou, V. and McKeown, K. Predicting the Semantic Orientation of Adjectives. ACLEACL'97, 1997.
[5] Hu, M and Liu, B. Mining and Summarizing Customer Reviews. Proceedings of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'04), 2004.
[6] Jindal, N. and Liu, B. Mining Comparative Sentences and Relations. Proceedings of National Conference on Artificial Intelligence (AAAI'06),

2006.

[7] Kanayama, H. and Nasukawa, T. Fully Automatic Lexicon Expansion for Domain-Oriented Sentiment Analysis. Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing (EMNLP'06), 2006.

[8] Kim, S. and Hovy, E. Determining the Sentiment of Opinions. Proceedings of the 20th International Conference on Computational Linguistics (COLING'04), 2004.

[9] Liu, B., Hu, M. and Cheng, J. Opinion Observer: Analyzing and Comparing Opinions on the Web.   Proceedings of International World Wide Web Conference (WWW'05), 2005.

[10] Pang, B., Lee, L. and Vaithyanathan, S. Thumbs up? Sentiment Classification Using Machine LearningTechniques. Proceedings of the 2002 Conference on Empirical Methods in Natural   Language Processing (EMNLP'02), 2002.

[11] Popescu, A.-M. andEtzioni, O. Extracting Product Features and Opinions from Reviews. Proceedingsof the 2005 Conference on Empirical Methods in Natural Language Processing (EMNLP'05), 2005.

[12] Turney, P. Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews. ACL'02, 2002.

[13] Wiebe, J. and Riloff, E. Creating Subjective and Objective Sentence Classifiers from   Unannotated Texts. Proceedings of International Conference on Intelligent Text Processing and Computational Linguistics (CICLing'05), 2005.

[14] Wilson, T., Wiebe, J. and Hwa, R. Just How Mad Are You? Finding Strong and Weak Opinion Clauses. Proceedings of National Conference on Artificial Intelligence (AAAI'04), 2004.

[15] A. Abbasi et al., "Affect Analysis of Web Forums and Blogs using Correlation Ensembles," IEEE Trans. Knowledge and Data Eng., vol. 20, no. 9, 2008, pp. 1168–1180.

[16] Dave, D., Lawrence, A., and Pennock, D. Mining the Peanut Gallery: Opinion Extraction and Semantic Classification of Product Reviews. Proceedings of International World Wide Web Conference (WWW'03), 2003.