



Speech Based Lexicon Analysis System Integrating Wrapping Attack Prevention Technique in Cloud Storage

A. Arun^a and M. Mayuranathan^a

^aDepartment of Computer Science and Engineering, Valliammai Engineering College, Chennai, India

E-mail: arunthomas.srm@gmail.com, manimayur@gmail.com.

Abstract: Sentimental analysis using speech to text conversion influences the consumer opinions given to each aspect over their overall opinions using sentimental analysis and also enhances content analysis by third party. The user driven reviews are stored in real cloud and third party checks the review and uploads in the application. To avoid data breach wrapping attack prevention techniques are implemented. The speech recognition server is also responsible for provoke the speech recognizer after it receives all the speech data and stores it to the server's hard disk. This is correlated by a system call. The speech recognition server and the speech recognition processes communicate with each other by reading and writing document in the server's local hard disks. In this user driven reviews about a product is taken into sentimental analysis to get positive, negative and neutral words. This would make the consumer come to a decision in a fraction of a section rather than going through n number of reviews, thus tremendously saving time.

Keywords: Sentimental analysis, speech recognition, wrapping, positive, negative, neutral.

1. INTRODUCTION

Cloud computing security, more simply, cloud security refers to a broad set of policies, technologies, and controls position to preserve data, applications, and the associated infrastructure of cloud computing. It is a root domain of computer security, network security, and, greater extent, information security.

Cloud computing and storage bane factor users with capabilities to store and process their data in relating to a person. Company use the cloud in a variety of different service models with descriptor such as SaaS, PaaS, and IaaS and deployment models private, public, hybrid, and community. Security concerns aggregation with cloud computing fall into two broad part: security issues faced by cloud provider's enterprise providing software, platform, or infrastructure-as-a-service via the cloud and public security faced by their customers companies or management who host applications or store data on the cloud.

The societal, weblog, meeting place, e-commerce web sites, etc. encourages people to share their opinion, concern and feelings publically. Today these internet sites are very popular and this resulted a huge increase of such opinions. People's opinions and augury are very valuable information in deciding process, but to get benefits from these opinion and familiarization, the accumulated content should be demodulate and analyzed properly.

This demodulate and analyzed sentiments or opinion are useful for consumer as well as invention as consumers can get a chance to valuate others opinion and experience related to some product or services before purchasing them. Likewise the contract can take these opinions as feedback from the consumers and they can perfect the worth of their product or services.

It also very useful for decision creator or policy makers of the politically entity. It is impossible for the policy creator to connect to the mass and collect their opinions efficaciously. It is true that via these media citizens can express their desires, problems, veneration and sensitiveness and the experts can make use of it by properly demodulate and analyzing it. But the demodulation and analysis of huge follow network content is beyond the human power and time. The content is mostly written in natural language.

This inclusion demands an automatic natural language processing tool that demodulate and analyzes the people sentiments from this unstructured text. Numerous probes are received in this direction. This research domain is called Opinion mining and sentiment analysis.

1.1. Speech Recognition

Speech recognition is the process of modifying the voice communication to a written text or some same form.

The major steps of a typical speech recognizer :

1. **Grammar design:** Analyzing grammars define the words that may be verbal by a user and the structure in which they may be voice communication. A grammar must be created and processed for a recognizer to know what it should listen for in incoming audio.
2. **Signal processing:** Analyzing the spectrum frequency characteristics of the incoming audio.
3. **Phoneme Analysis:** Compare the spectrum patterns to the patterns of the phonemes of the language being Analysis.
4. **Word Analysis:** Compare the combination of likely allophone against the words and patterns of words specified by the active grammars.
5. **Result Formation:** Corvine the application with information about the words the analysis has detected in the incoming audio. The result information is always provided once analysis of a single utterance frequently a sentence is complete, but may also be provided fetch the recognition process. The result always predicts the recognizer's best guess of what a user said, but may also predict alternative guesses.

2. RELATED WORK

Decision making both on individual and organizational level is always ruled by the search of other's opinion on the same. With tremendous establishment of opinion rich resources like, reviews, for group meeting elaboration, blogs, micro-blogs, Twitter etc. provide a rich collection of sentiments. This user generated content can serve as a benefaction to market if the semantic positioning is measured. Opinion mining and sentiment analysis are the systematization for studying and construing opinions and sentiments. The digital system has itself sealed way for use of huge volume of opinionated data recorded. [1]

Attribute-based encryption (ABE) can provide fine-grained non-interactive access control and encryption functionalities simultaneously, it has become a promising technique for cloud computing. ABE scheme for semi-anonymity and access privilege in the standard model, based on which, they decentralized the central authority to limit the identity leakage and introduced the file privilege control to manage all the operations on the cloud data. The original access tree to several privilege trees, where each privilege tree described one operation on the cloud data. Due to the absence of a central authority, all the authorities should work jointly to create the master key for each authority, the private key for each user, the public parameter and the master key for the whole system. [2]

The performance of BOW sometimes remains limited due to some fundamental deficiencies in handling the opposition shift problem. We propose a model called dual sentiment analysis (DSA), this problem for sentiment categorization. First propose a fiction data extension technique by creating a sentiment converse canvas for each training and test review. On this basis, propose a dual training algorithm to make use of real and reversed training reviews in pairs for learning a sentiment classifier, and a duple prediction algorithm to classify the test reviews by considering two sides of one review. The Bag-of-words (BOW) model is typically used for text representation. In the BOW model, a review text is depicted by a vector of separate words. The statistical machine learning algorithms. [3]

Numerous consumer reviews of products are now available on the network. Consumer canvas contains rich and worthy Knowledge for both firms and users. However, the reviews are often disordered, leading to difficulties in information navigation and knowledge assumption. The development of a probabilistic aspect ranking algorithm is to infer the importance of aspects by at the same time relatives aspect incidence and the assurance of consumer opinions given to each sector over their overall opinions. The important product aspects are identified based on two measurements: the important aspects are usually commented on by a large number of feed and consumer opinions on the important characteristic greatly influence their overall opinions on the product. [4]

3. SYSTEM ANALYSIS

3.1. Problem Definition

1. **Sentimental analysis:** It uses voice to text modification and influence consumer review given to each characteristic over their overall opinions using sentimental analysis.
2. **Wrapping attack:** The server checks the valid user by the proper value (which is also duplicated) and integrity checking for the message is done. This demodulated and analyzed sentiments or opinion are useful for consumer as well as make as consumers can get a chance to evaluate others opinion and augury cognate to some product or services before purchasing. To assign weight age to identified positive, negative and neutral words implemented Sentiwordnet. SentiWordNet is a lexical resource for opinion mining. It assigns to each synset's of WordNet three sentiment scores: positivity, negativity, neutral. SentiWordNet (synsets) is associated with two numerical scores ranging from 0 to 1, each indicating the synset's positive and negative bias.

3.2. Proposed System Feature

The System develops a probabilistic sentimental analysis using speech to text conversion and influence consumer opinions given to each aspect over their overall opinions using sentimental analysis. Also, enhanced content analysis by third party. The user driven reviews are stored in real cloud storage and third party checks the review and uploads in the application. To avoid data breach to implemented wrapping attack prevention techniques. In this user driven reviews about a product is taken into sentimental analysis to get positive, negative and neutral words.

4. SYSTEM DESIGN AND IMPLEMENTATION

The architecture of proposed system Fig 1. Speech recognition servers is also responsible for invoking the speech recognizer after it receives all the speech data and stores it to the server's hard disk. This is correlated by a system call. The speech recognition server and the speech recognition processes communicate with each other by reading and writing files in the server's local hard disks. Java-based speech recognition system, which uses a small-vocabulary, isolated word recognition task, and in particular the digit recognition task. The core of the recognizer was based on continuous relative density HMMs.

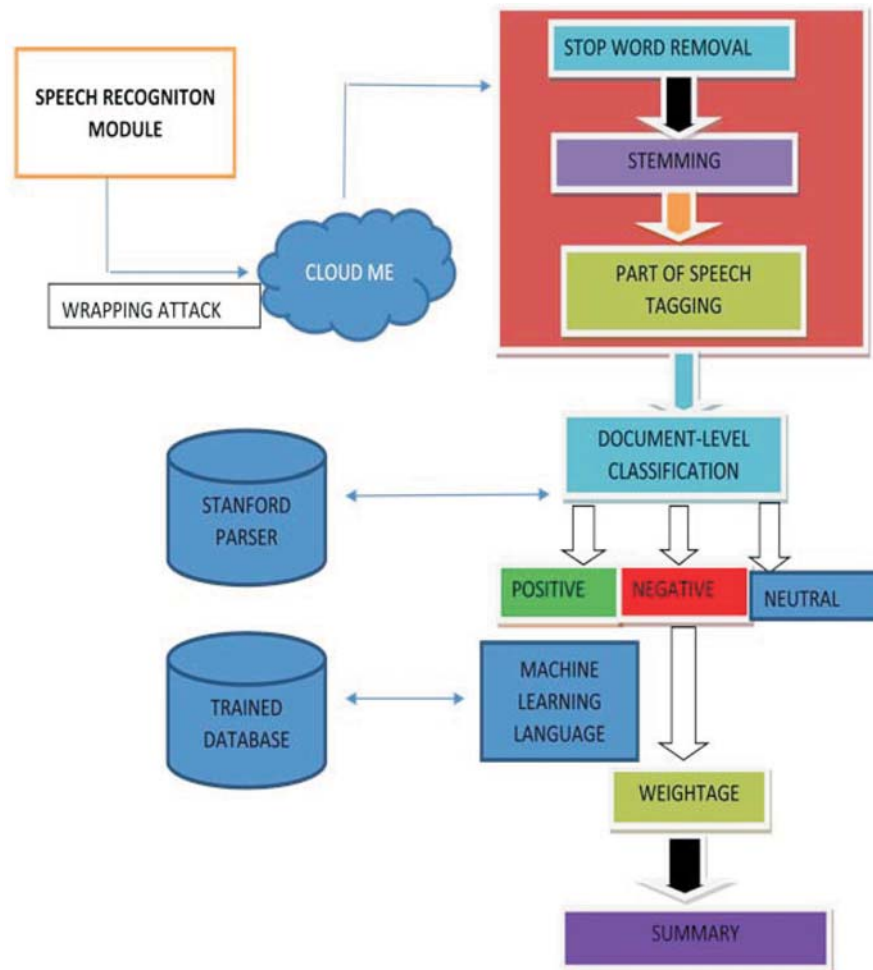


Figure 1: Architecture design

The pre-processing stage in speech recognition systems is used in order to increase the effectiveness of subsequent feature extraction and classification stages and therefore to increase the overall recognition performance. Commonly the pre-processing includes the sampling, a windowing and a de-noising.

4.1. Pre-Processing Methods

The pre-processing stage in speech recognition systems is used in order to increase the effective of subsequent feature extraction and classification stages and therefore to improve the overall recognition performance. Commonly the pre-processing includes the sampling step, a windowing and a de-noising step.

1. Stop Word Removal
2. Stemming
3. POS Tagging

1. **Stop Word Removal** : The working with text mining applications, too often hear of the term “stop line” or “stop word line” or even “stop count”. Stop words are basically a set of same used words in any language, not just English. Though stop words usually refer to the most common words in a language, there are no various lists of stop words used by likely tools, and indeed not all tools even use such a list.

2. **Stemming** : Stemming is the processes of reducing inflected or sometimes output words to their word stem, base or root form generally a written word.
3. **POS Tagging** : A Part-Of-Speech Tagger POS Tagger is a piece of software that reads text in some language and aerial parts of speech to each word such as noun, verb, adjective, etc., although generally estimation applications use more pulverized POS tags like ‘noun-plural’.

4.2. Document Level Classification

The document comprises an overview of the testing practice to be improvising when an upgrade or enhancement is made, or a module is added to an existing application. The emphasis is on testing sarcastic business computing, while minimizing the time required for testing while also mitigating risks.

1. Positive
 2. Negative
 3. Neutral
1. **Positive** : Words that give a very pleasant meaning are considered to be positive words, for example excellent, awesome, and very good.
 2. **Negative** : Words that give a very unpleasant meaning are considered to be negative words, for example bad, worst, not satisfactory
 3. **Neutral** : Words that give an overall meaning that is either good or bad, for example satisfactory, okay.

4.3. Graph Generation

The graphs are generated here based on the values obtained from the analysis of the reviews using the sentimental approach. The graphs are generated based on the positive, negative and neutral values obtained. To gather experimental evidence for conjectures, it is frequently required to verify that all graphs in a given class satisfy a desired property. The graphs are generated here based on the values obtained from the analysis of the reviews using the sentimental approach.

5. RESULT



Figure 2: Home page of lexicon Analysis system

The discuss about Figure 2 home page of lexicon Analysis system. It placed in Home, Registration and Login system.

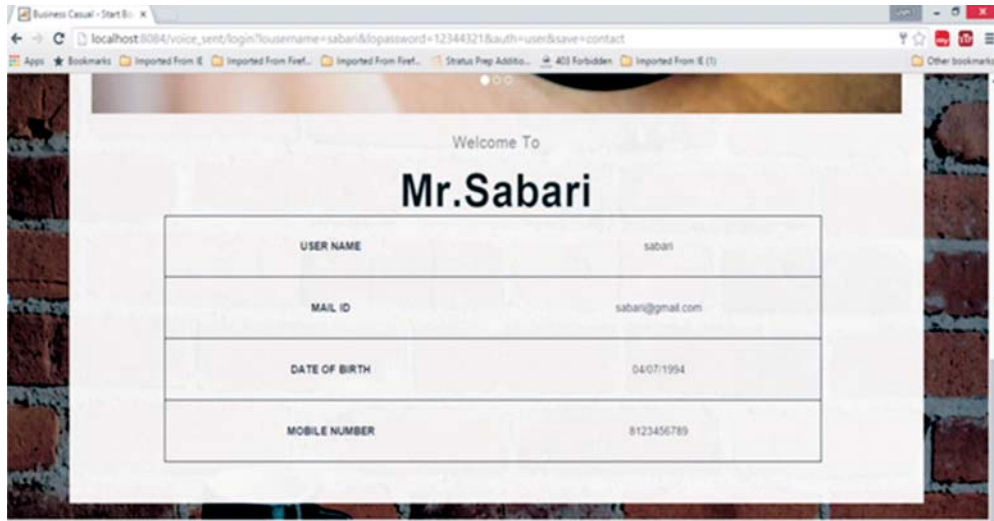


Figure 3: Customer profile page

The discuss about Figure 3 customer profile page is display particular customer details.

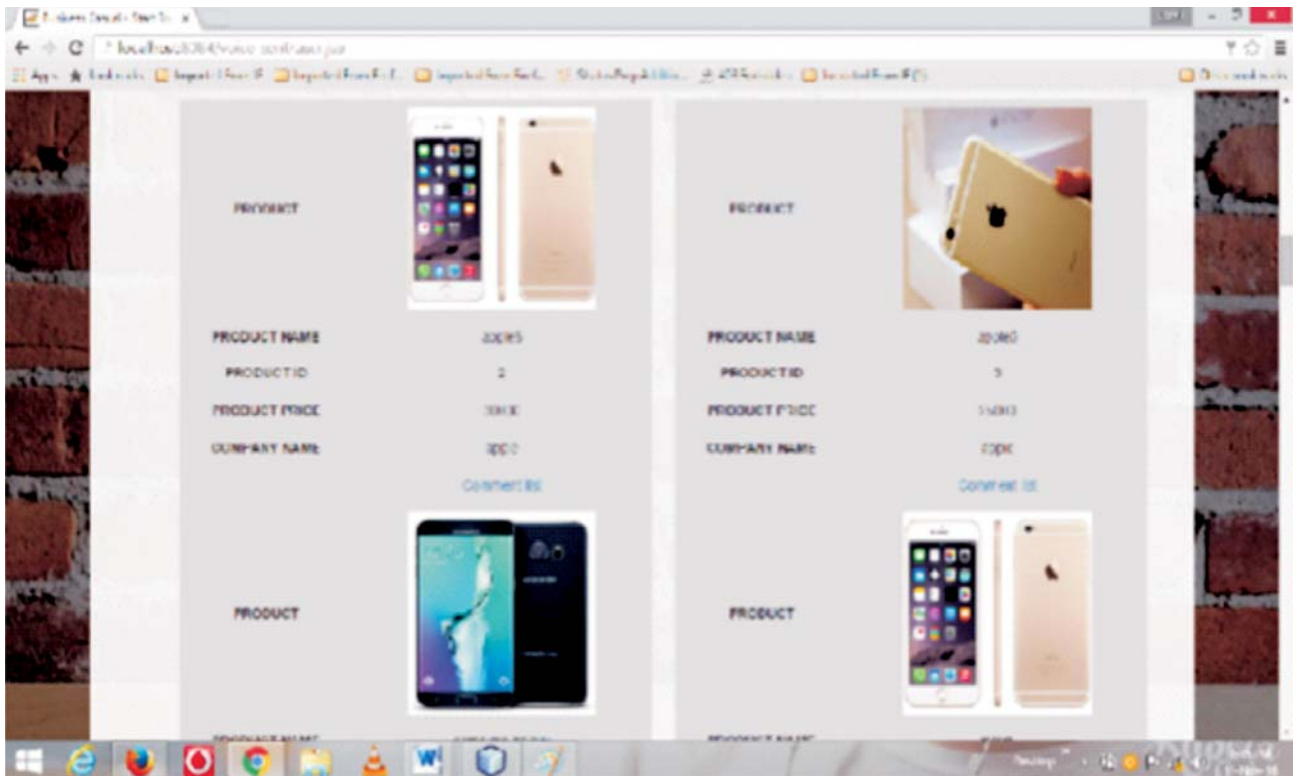


Figure 4: Customers to select the product with detailed reviews

In figure 4 product details and particular product reviews displays each and every product images and details.

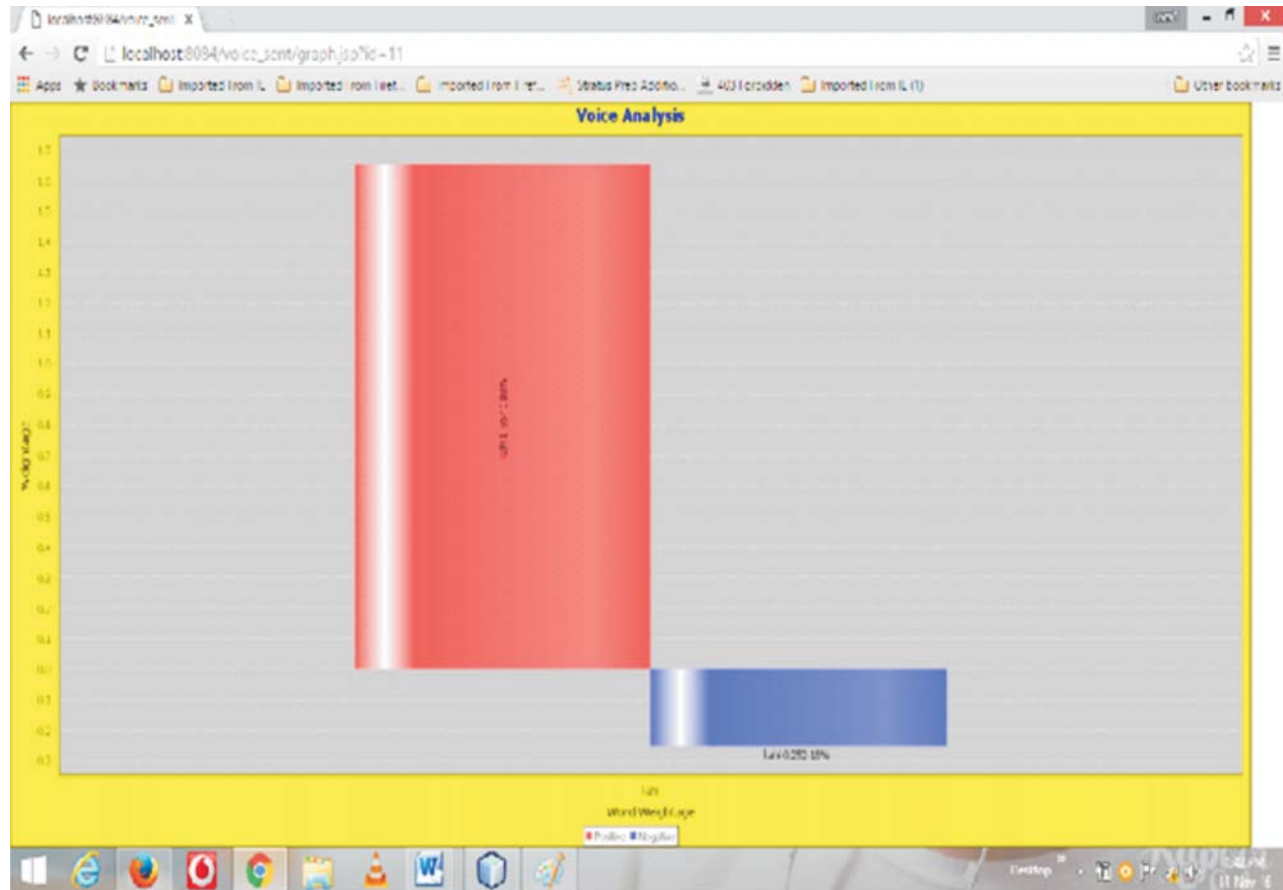


Figure 5: This is Bar chart for user review for positive, negative, neutral

In figure 5, Bar chart for user review for positive, negative and neutral process explains diagram model.

6. CONCLUSION AND FUTRURE WORK

6.1. Conclusion

This project is based on the observation that buyers often express opinions openly in free text feedback comments, propose common trust for trust evaluation by mining feedback comments. The user reviews (text) is stored in cloud for audit purpose. Once the audit is performed the reviews are posted in the respective applications. In this user driven reviews about a product is taken into sentimental analysis to get positive, negative and neutral words. This would make the consumer come to a decision in a fraction of a section rather than going through n number of reviews, thus tremendously saving time.

6.2. Future Work

Using this project can be implemented in social media such as flipchart, twitter, Face book, Amazon etc. And also, to use it in the social media and for customer use. Speech recognition is the process of converting spoken language to written text and influencing consumer opinions given to each aspect over their overall opinions using sentimental analysis.

REFERENCE

- [1] Pravesh Kumar Singh¹, Mohd Shahid Husain², “Methodological Study of Opinion Mining and Sentiment Analysis Techniques”, (IJSC) Vol. 5, No. 1, February (2014).
- [2] Hui Ma, Rui Zhang, and Wei Yuan Comments on “Control Cloud Data Access Privilege and Anonymity with Fully Anonymous Attribute- Based Encryption” (2014).
- [3] Rui Xia, Feng Xu, Chengqing Zong, Qianmu Li, Yong Qi, and Tao Li Dual Sentiment Analysis: Considerin Two Sides of One Review, (2015).
- [4] Zheng-Jun Zha, Jianxing Yu, Jinhui Tang, Meng Wang, Tat-Seng Chua 2014, Product Aspect Ranking and Its Applications, (2014).
- [5] Liu, B., “Sentiment Analysis and Subjectivity”. Appeared in Handbook of Natural Language Processing, Indurkha, N. & Damerau, and F.J. [Eds.]. p241-249 (2010).
- [6] Hu, M. & Lui, B. “Mining and Summarizing Customer Reviews”. In Proceedings of ACM SIGKDD Conference on Knowledge Discovery and Data Mining p. 168–177. (2004).
- [7] P. Turney and M. L. Littman, “Measuring praise and criticism: Inference of semantic orientation from association,” *ACM Transactions on Information Systems (TOIS)*, vol. 21, no. 4, pp. 315-346, (2003).
- [8] M. El-B`eze, A. Jackiewicz, and S. Hunston, “Opinions, sentiments et jugements d’evaluation,” *Traitement Automatique des Langues*, vol. 51, no. 3, (2010).
- [9] M. Mostafa, “More than words: “Social networks text mining for consumer brand sentiments,” *Expert Systems with Applications*, vol. 40, pp. 4241–4251, (2013).
- [10] . Pittermann, A. Pittermann, and W. Minker, “Emotion recognition an adaptation in spoken dialogue sy stems,” *International Journal of Speech Technology*, vol. 13, pp. 49– 60, (2010).