

A Review on Summery Generation with Message Categories for Social Messaging

Kapade Ruchita R
Dept. of Computer Engg.
G.C.O.E.,Nagaon
Dhule, India

Sharma Vrajbala S
Dept. of Computer Engg.
G.C.O.E.,Nagaon
Dhule, India

Deore Megharani S
Dept. of Computer Engg.
G.C.O.E.,Nagaon
Dhule, India

Mr. S.P.Chavan
Asst. Prof. Dept. of Computer Engg.
G.C.O.E.,Nagaon
Dhule, India

Abstract: The development in smart phone technology and time reduce prices increase its use in modern day to day life. People started using smart phone for all sorts of activities as they found it easy to use and they can perform their task on the move. One such application which made peoples life was message like WhatsApp, hika. This was one of the biggest revolutions in social media and people now days have almost been addicted to use of WhatsApp. Our proposed system will make this use simpler by providing a summary view to user of their set of messages and also categorized the messages so that user can get more simple and effective view of their messages. System will use text mining and image processing algorithms to generate summary of messages and categories the messages. Our system will be a prototype model to WhatsApp application and will allow user with some basic functionality of WhatsApp application with main aim to show summary view and categorized view of messages.

Keywords: : Messaging, summary, TimeLine, Text

I. INTRODUCTION

Recently social network sites are using so much as we can say it is part of everyone life. Social network sites are one of the modes of communication for the people all over the world. Any number of people from different part of the world can communicate over the internet. There are so many social networking sites such as what's app, Twitter, Facebook etc. there was a survey represented by Twitter is one of the famous social networking site, it also have huge number of interactions every day in form of billions of comments, messages.

All social media sites are very easy to use and convenient for expressing views on different topics. So popularity of such sites is very high among the people. User can review, express their feelings on that or also can simply forward it further, even one can like or leave comment on it. As the popularity of such social networking sites is more so number of such messages is very high with high generation rate. When any user wants to refer any certain message of comment, he has to refer them all which is impossible every time and note feasible. It will take lots of time of user in search of particular important messages or review. because of this things we summarize & categories the messages.& this is a main motivation of our work to summarize the messages. & categories the messages for easily access the required & important messages.

Summarization represents a set of documents by a summary consisting of several sentences. Intuitively, a good summary should cover the main topics (or sub topics) and have diversity among the sentences to reduce redundancy. Summarization is extensively used in content presentation, especially when users surf the internet with their mobile devices which have much smaller screens than PCs.

II. LITERATURE SURVEY

In paper [1], "Summarization and Timeline Generation for Social Stream with Categorization." (Snehal Chavan¹, Sanchayati Thorat², Tejaswini Surade³, Priyanka Ghongade⁴) proposed a technique tweet cluster vector algorithm for maintaining statistical data and compact cluster information to maintain dynamically in memory during stream processing, store and organize cluster snapshots of different moments. Generation of online and historical summaries with arbitrary time durations, in which propose TCV rank summarization algorithm. & also proposed an evaluation method which generates timeline, categorization based on topic evaluation.

In paper "Summarization Method and Timeline Generation of the Tweet" (Pooja Patil, Nilamvhatte, Srushti Rajput, Ujjwala Panhalkar, K. V. Deshpande) worked on a three types of operations on tweets, named as clustering of tweets, summarization and topic evaluation over tweeter data. summarization framework called Sumblr. It is proposed a continuous tweet stream summarization framework, namely Sumblr, to generate summaries and timelines in the context of streams. & use a novel data structure called TCV for stream processing, and propose the TCV Rank algorithm for online and historical summarization. We are using a topic evolution detection algorithm which produces timelines by monitoring three kinds of variations. Extensive experiments on real Twitter data sets demonstrate the efficiency and effectiveness of our framework. Produce multi topic summarization in our proposed will try to cover this point.

II. PROBLEM DEFINITION

The user register on the system then it can received and send the messages and simply forward. But in some messages can be repeated and other messages can be not read properly and time also waste.

IV. PROPOSED SYSTEM

The proposed system is a cloud based architecture where all the clients will communicate with one server via internet/ wifi medium. The system will consist of relational MSQYL database which will be managed by the server. System will work in private cloud environment.

SYSTEM ARCHITECTURE

System consists of following modules:

1. User android application
2. Web Server

User android application:

The mobile application will be android application which user can install on their android smart phones. The front end of this application will be developed using android XML layouts. The backend code of the application will be in android java. This application will have following features:

1. User Registration
2. User Login
3. User can post messages to other users
4. User will view message summary
5. User will view message in different categories

Web Server:

The web server will be developed in net beans editor. The web server will include java servlets for each and every operation that server will perform. The web server will access database using jdbc connection. Web server is responsible for all the request and response handling mechanism. Web server will be driven by apache tomcat application server. This server will be responsible for generating summary of user's messages and categorization of messages in different categories. Server will also send auto response message to the sender of message based on the category of message.

The following diagram depicts the architecture of the proposed system.

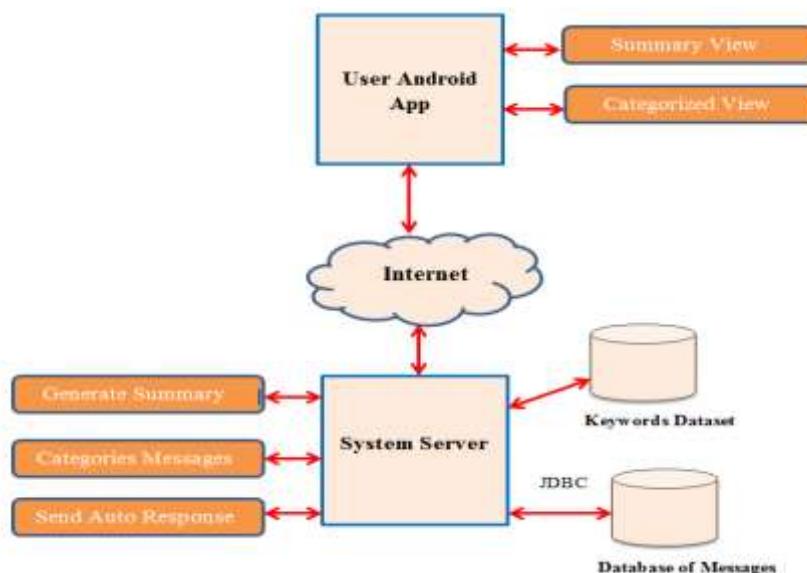


Fig 1: Architectural diagram

Algorithm-:

Text Mining Algorithms:

a) Stemming: The stemming process finds the root words of any input words. It eliminates the tense part from the word. The stemming algorithm increases the comparison efficiency. We have used the standard porter stemmer algorithm for the same. This algorithm works in five steps which are as follows.

Step 1: gets rid of plurals and -edoring

Step 2: turns terminal y to i when there is another vowel in the stem.

Step 3: maps double suffices to single ones. so -ization.

Step 4: deals with -ic-, -full, -ness etc.

Step 5: takes off -ant, -ence etc.

b) Term Frequency: By using term frequency we can find out the occurrence of the word in that particular text data. To find term frequency we have used the following equation:

$$\text{Term Frequency (tf)} = \frac{\text{occurrence of word}}{\text{total word count in the text data}}$$

c) Cosine Equations: The cosine equations are used to find difference between two numeric values. The following is the equation of cosine similarity: $\text{coeff_Cosine} = \frac{x*y}{\text{Sqrt}(x^2*y^2)}$

CONCLUSION AND FUTURE SCOPE

Messaging like whatsapp has become an integral part of any mobile user due to its simplicity and user friendliness. Major problems with current version is user gets man duplicate messages and images and user need to check each and every message which is time consuming approach. Our system which will provide summary view will eliminate all the duplicate messages and provide user with only unique messages and images along with categorization of messages. Thus our system will add further user friendliness.

REFERENCE

1. Zhenhua Wang, Lidan Shou, Ke Chen, Gang Chen, and Sharad Mehrotra "On Summarization and Timeline Generation"IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 27, NO. 5, MAY 2015 for Evolutionary Tweet Streams
2. Snehal Chavan¹, Sanchayati Thorat², Tejaswini Surade³, Priyanka Ghongade⁴ Department of Computer Engineering, SND College of Engineering & RC, Yeola, India Summarization and Timeline Generation for Social Stream with Categorization. International Journal of Emerging Technology and Advanced Engineering Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 6, Issue 9, September 2016)
3. Pooja Patil, Nilamvhatte, Srushti Rajput, Ujjwala Panhalkar, K. V. Deshpande Department of Computer Engineering, RSCOE, Tathawade, Savitribai Phule Pune University, Maharashtra, "Summarization Method and Timeline Generation of the Tweet"
 - a. © 2018 IJSRSET | Volume 4 | Issue 8 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099 Themed Section : Engineering and Technology
4. <http://it.wikipedia.org/wiki/Tf>.