Real Time Shop Recommendation Via Text Messaging

Ashish S. Dhumal¹, Swapnil V. Dhadse², Vishal N. Gavhane³, Sandeep S. Kokate⁴ & Prof. Reshma Patil⁵

¹,²,³,⁴,⁵ KJCOEMR, Department Of Computer Engineering, Pune, India.

Abstract—Now days it’s very difficult to find compatible shop along with requirements, for find shop nearby location marketing, cellphone calling and internet is necessary. In this paper we are discussing about the SMS services which help the user to gather necessary information along with user requirements using SMS services. With the help of data crawler meaning technique and Hybrid recommendation it is very easy to get details for user along with requirement. Most of the system on SMS based is having some performance issue that can load lead to lower accuracy of the technique so this paper process a novel idea of SMS. Communication speed is very important in SMS technology, with the Crawler and SMS data gathering and frequent communication is very easy.

Keywords—Distributed computing, Data Crowler, Hybrid Recomentaion, SMS service

I. INTRODUCTION

In this digital world because of the Internet communication between source to destination is very easy. Social networking is widely used for sharing information. For accessing internet services mobile must have to compatible with internet services and able to access social services. In another way, SMS services is just in time and secure communication media for communication. For sharing information using Mobile technology SMS service is real time and secure.

Term frequency reward words having high accuracy in a document. Invers document frequency penalized the word appearing the many times in a document collection. To generate words should not have high weighted eg. ‘or’, ‘hot’, ‘is’, ‘the’ etc. Hence word rate in a document collection but frequency in a particular document get high weighted. After combining TF Algorithm and IDF Algorithm together assign high weight to discriminative words in a document foe a text document representation TF-IDF is well established technique.

In Social networking sites and online shopping sites for example Facebook, Twitter, Amazon, Flipchart user review is important for company as well an new customer. For processing customer review the Fuzzy Logic algorithm is widely used, Fuzzy Logic result is depend on Sentiment of the Sentence, the sentiment should be positive, negative, or neutral. Based on the sentiment analysis review rating is decided for a particular user review.

Hybrid Recommendation system is combination of content based filtering and collaborative filtering, could take advantaged form both the sides and similarities among users. Although there are several ways in which to combine the two techniques a distinction can be made between two basis approaches. In hybrid recommendation approach, it combines the more than one types of information it is also possible to use the recommendations of multiple filtering techniques independently.

It’s possible to have same product in one location at multiple shops. With the help of similarity based search algorithm it’s very easy to find multiple shops in one location (eg. Find Detol soap in Kondhwa.), in this example user want to find shops in Kondhwa location for keyword Soap.

II Literature Survey

[1]. A New Approach to Target Dependent Sentiment analysis with onto fuzzy logic.

In this paper [1], Authors analyze the most powerful social Networking site twitter to analyze the mood of peoples. Mostly peoples expressed their thoughts using commenting on trending topics (using Hash Tag). These comments are in the form of positive, negative or neutral. These comments encourage companies to provide best services to their customers. In this paper they uses Naïve string searching algorithm, Loopy Belief Algorithm, Enhanced Boosting Classification, Fuzzy logic etc.

[2]. Social Network Comment Classification using Fuzzy based classifier technique.

The authors of paper [2], knows today Internet become the most essential part of society. In this content the odd user can misbehave or can be involved in unsocial activity over the pure
environment. By analyzing the data of peoples on social networking site they categories the peoples by their behavior. They identify the different techniques which currently using in many applications such as Data mining, machine learning, Bayesian classifier and k-nearest neighbour algorithm, fuzzy classification technique, etc.


In this paper [3], authors identify that this is era of digital age. In this paper they identify that many e-commerce website and online market is increase so for one product there are many options are available like automobile field having many options today. So authors of this paper bring a platform where hybrid recommendation is available for consumers so they can choose the best for them. This will grows the decision making power of consumers. They take the feedback of consumers as internal feedback and external feedback using both collaborative filtering and content based filtering.

[4]. Hamming Distance based Approximate Similarity text search algorithm.

Today data is creating in tremendous amount which is in form of zettabyte (ZP). For that many techniques are invented for handling of this large amount of data like machine learning, data mining, so on. They propose a Hamming distance based approximate similarity text search (BASTS) algorithm to increase the quality of queries in huge text data. The HASTS algorithm first constructs an index table with the substrings extracted randomly from the feature fingerprints generated by the SimHash algorithm.

In this paper, section 2 is dedicated for literature review of past works. Section 3 describes the proposed methodology and Section 4 discusses the results and evaluation of the proposed technique. Finally Section 5 concludes this paper with future extension possibilities.

III. PROPOSED SYSTEM

The proposed methodology for Shop recommendation system through messaging can be efficiently explain with the below mentioned steps.

---

**Figure 1: System Overview Diagram**

**Step 1:** This is the primitive step of our proposed system.

Where the end user sends a text message for the area and the item to be purchased in that area to the number of the SMS server.

**Step 2:** SMS server is connected with the GSM mobile modem that is powered with the AT commands. As soon as the SMS arrives at the GSM modem it will be read through the AT commands which are written using the java communication API in our model. This receiving process is running constantly with the live processing thread at SMS Server. On receiving SMS String SMS server sends this string to the processing server to get the desired recommendation for the user.

**Step 3:** As soon as processing server received the SMS String from SMS Server it first starts crawling all the super shops data from the web sites that are hosted in web server. For this purpose System designs an interactive web crawler that crawls all the data on the web pages of each super shop using the open URL Stream schemes of java network packages.

Once the URL’s are crawled their content will be fetch in the form of Strings that contains all the html tags for the respective URL, now these html tags need be shred off using html parser which eventually gives human readable data from the respective web page.

**Step 4:** In this step all the reviews data that are in the human readable text are need be preprocess to identify the required words like positive comments and negative comments. So to this comments data need to be preprocess which follows mainly three steps of preprocessing like

✓ Tokenization: This is the process of collecting all the words in a vector and assigning index tokens to them.
Stopword Removal: This process is nothing but removing of conjunctives from the given text. Conjunctives are like is, of, the, and etc.. On removing of these the meaning of the text doesn’t change at all.

Stemming: Stemming is the process of bringing the word to its base form, on doing so the meaning of the word doesn’t change. For example going word becomes go on stemming process, where we shred off “ing”.

This process of preprocessing can be depicted clearly with the below mentioned algorithm1.

**ALGORITHM 1: Preprocessing**

---

Step 0: Start
Step 1: Read string
Step 2: divide string into records on space and store in a vector V
Step 3: Remove Special Symbols
Step 4: Identify Stopwords
Step 5: Remove Stopwords
Step 6: Identify Stemming Substring
Step 7: Replace Substring to desire String
Step 8: Concatenate Strings
Step 9: stop

**Step 5:** To identify the importance the words from SMS string our proposed model is using the TF-IDF, which is nothing but term frequency and inverse document frequency for all the crawled data vectors. This can be depicted with the below algorithm 2

**ALGORITHM 2 : TF-IDF**

---

Step 0: Start
Step 1: Read the Preprocessed query string
Step 2: divide string into words on space and store in a vector V
Step 3: FOR i =0 to N (Where N is the length of V)
Step 4: W= V[i]
Step 5: Count W for the respective DB as TF
Step 6: Count W for the all other DB that is DF (DB Frequency)
Step 7: IDF= log ( DF)
Step 8: TF-IDF= TF* IDF
Step 9: END FOR
Step 10: Stop

**Step 6:** Once the TF-IDF is evaluated then the results of similarity indices are optimized to get the hybrid recommendation for providing the best shop names to the user. Finally This recommended string is send back to the user on his same mobile number though AT Commands which are handled by java communication API.

**IV. RESULTS AND DISCUSSIONS**

Proposed model is deployed using java technology using Netbeans 6.9.1 as core development IDE and MySQL 5.0 as SQL server for maintaining database. And the web part of the model is hosted using Apache tomcat server.

System is put under hammer to measure its accuracy using mean absolute errors MAE which can be represented using equation 2.

\[
    MAE = \frac{\sum_{i,j} | r_{i,j} - r'_{i,j} |}{N}
\]

Where

- \( r_{i,j} \) denotes the expected recommendation item \( j \) observed by service user \( i \),
- \( r'_{i,j} \) denotes the predicted recommendation value,
- \( N \) denotes the number of predicted values.

An experiment is conducted in between traditional model of collaborating filtering and our model of hybrid recommendation and the yielded results are tabulated in the below table 1.

<table>
<thead>
<tr>
<th>No of Runs</th>
<th>Hybrid Recommendation</th>
<th>Collaborative Recommendation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.2</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>1.8</td>
<td>7.5</td>
</tr>
<tr>
<td>3</td>
<td>1.6</td>
<td>3.4</td>
</tr>
<tr>
<td>4</td>
<td>1.74</td>
<td>8.4</td>
</tr>
<tr>
<td>5</td>
<td>2.2</td>
<td>14</td>
</tr>
</tbody>
</table>

**TABLE 1: RECOMMENDATION COMPARATIVE MAE**
The above plot indicates the efficiency of our system by achieving the less MAE than that of traditional collaborative filtering method. This indicates the superiority of the system due to using of hybrid recommendation system.

V. CONCLUSION AND FUTURE SCOPE

Now a days due to increasing number of mobile users there exist a problem of mobile network sustainability. This leads to a problem of poor availability of network at mobile nodes and there by poor connection of internet is always there. In that case it becomes problem for the mobile users to search the best shop to purchase their required items. So this paper proposed a methodology of providing the recommendation via text messaging using hybrid recommendation model in the same instance of requirement of the users.

This system can be enhance in the future by adding huge number of URL’s in big data database using interactive web crawlers that will work throughout the year to gather information.

REFERENCES

[1]. Shruti Joshi,sonal mehta,Purva mestry,Ashwini Save,” A New Approach to Target Dependent Sentiment Analysis withOnto-Fuzzy Logic”, 2nd IEEE International Conference on Engineering and Technology (ICETECH), 17th& 18th March 2016, Coimbatore, TN, India.


[4]. Haifeng Hu, Liang Zhang, and Jianshen Wu,” Hamming Distance based Approximate Similarity text search algorithm”, 7th International Conference on Advanced Computational Intelligence Mount Wuyi, Fujian, China; March 27-29,2015.