IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 1, January 2017

Cloud-Based Remote Venue Recommendation Framework

Ms. K. C. Sonawane[1], Ms. S. S. Ponde[2]

Student, Department of Computer Science, DIEMS, Aurangabad, India [1]

Assistant Professor, Department of Computer Science, DIEMS, Aurangabad, India [2]

Abstract: Now a days, recommendation systems have seen outstanding evolution in the field of knowledge engineering. Most of the existing recommendation systems based their models on collaborative filtering approaches that make them simple to implement. Still, performance of most of the existing collaborative filtering. Based recommendation system suffers due to the challenges, like: (a) cold start, (b) data sparseness, and (c) scalability. But, recommendation problem is often characterized by the presence of many conflicting objectives or decision variables, such as users' preferences and venue closeness. In this work, introduced Cloud-based Bi-Objective Recommendation Framework (BORF) for mobile social networks. These utilizes multi-objective optimization approaches to generate personalized recommendations. To address the issues pertaining to cold start and data sparseness, the BORF performs data pre-processing by using the Hub-Average (HA) inference model. Moreover, the Weighted Sum Approach (WSA) is implemented for Pyramid Maintenance Algorithm (PMA) is applied for vector optimization to provide optimal suggestions to the users about a venue. The results of comprehensive experiments on a large area real dataset confirm the accuracy of the proposed recommendation framework. In the future, we would like to increase our work by incorporating more contextual data in the form of objective functions, such as the check-in time, users' profiles, and interests in our proposed framework. Moreover, we intend to integrate other concepts, such as machine learning, text mining, and artificial neural networks to refine our existing framework.

Keywords: Context-Aware Web Services, MOPNAR, Multi-objective Optimization, Collaborative Filtering.

I. INTRODUCTION

Recommender systems goal to help users with data access and retrieval applications when large collections of items are involved. In general, they work by means of suggesting those items that should be the most attractive ones to the users based on their past personal preferences. In recent years, recommendation systems have seen significant evolution in the field of knowledge engineering. Mostly the existing recommendation systems based their frameworks on collaborative filtering approaches that make them simple to implement.

However, performance of most of the existing collaborative filtering- based recommendation system suffers caused by the challenges, such as: (a) cold start, (b) data sparseness, and (c) scalability. Moreover, recommendation problem is often characterized by the presence of many conflicting objectives or decision variables, such as users' preferences and venue closeness.
Recommender system is a software application agent that presents the choices, liking and preferences of individual persons/ users and makes recommendation accordingly. During online search they provide simple technique for users to make decisions based on their recommendations. Majority of existing approaches of recommender system focus on factors like individual persons choice of item, group community opinions and do not consider importance such as contextual data like Location of user from big data. This was possible using the technique named as collaborative filtering which is based on past group community opinions for user and item and correlates them to provide results to the user questions/queries. Most of the community opinions were a group of tuple i.e. (user, ratings, and item) of user and numeric rating for item. At present, large number of applications makes use of location based ratings to provide user and item locations. For example, social networks based on location like Foursquare, Facebook and Twitter allows users to find spatial destinations (e.g. constructions, Shopping malls etc.) And provide ratings on their visit and hence capable of associating both user and item. Content-based Recommendation It recommends items to users that are similar to those they preferred previously. The analysis of similarity is based on the items attributes. Collaborative Recommendation It recommends items to users according to the item ratings of other people who have characteristics similar to their own. The analysis of similarity is based on the user's tastes and preferences. Hybrid Recommendation It is a combination of content-based and collaborative recommendations.

In this paper, introduced Cloud-based Bi-Objective Recommendation Framework (BORF) for mobile social networks. The Context utilizes multi-objective optimization techniques to generate personalized recommendations. To address the issues pertaining to cold start and data sparseness, the BORF performs data pre-

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 1, January 2017

processing. Moreover, the Weighted Sum Approach (WSA) is implemented for Pyramid Maintenance Algorithm and MOPNAR Algorithm is applied for vector optimization to provide optimal suggestions to the users about a venue. The results of comprehensive experiments on a large -scale real dataset confirm the accuracy of the proposed recommendation framework.

A good example of recommendation systems for e-commerce applications is Amazon.com, Flipcart.com where customers receive personalized recommendations on a variety of products. In the past few years, several social networking applications, such as Foursquare, Gowalla, and Google Latitude were developed for mobile Devices. These applications allow users to perform a "check-in" at venues that users visit to share experiences in the form of a feedback or tip. Moreover, these services collect and hold huge volumes of users' geospatial check-in data. Based on the data extracted by the mobile social Networking Applications, several location-based recommendation systems were developed in the recent years which recommend venues to users closely related to their preferences. A major research challenge for such systems is to generate real-time venue recommendations for a given individual from a massively diverse dataset of users' historical check-ins. To generate an optimal recommendation for an individual, the system must simultaneously consider the following factors: **(a)** personal preferences, **(b)** past check-ins, **(c)** current context, such as time and location, and **(d)** collaborative social opinions (other individuals' preferences). The proliferation of wireless and cellular networks over the last few years has led to a remarkable rise in the number of users who are using a variety of modern mobile Internet-enabled devices such as iPhones, iPads, and Android-based smart phones to consume online services. Mobile users are increasingly requiring services tailored to their context as they are on the move. Therefore, enterprise services should be context-aware to deal with the changing environment of the user. Several definitions of the notion of context have been provided in the literature. According to Author, "Context is any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves."

According to this definition, the amount of information that can be categorized as context information is extremely wide. Location, time, temperature, humidity, pressure, and mobile user activity are the most widely used context indicators by applications. Specialized services, that we call context services, capture, store, analyse and aggregate data to provide high-level context information to consumer application services as needed. Context services and context consumers are often physically distributed. Besides, it is likely that these context sources provide the same context information but with different QoC. The QoC concept is explained. Context-awareness raises challenges like aggregation of context information in a structured format, discovery, and selection of appropriate context services for context delivery to context consumers.

1.1 Research Motivation

Recommender systemsaim to help users with information access and retrieval applications when large collections of items are involved. In general, they work by means of suggesting those items that should be the most appealing ones to the users based on their past personal preferences. As the most popular and successful recommendation method, Collaborative Filtering (CF) have been widely used in many popular commercial recommender systems such as YouTube, Reedit, and Last.fm and Amazon etc1. The core work of CF in such recommender systems is predicting the unknown rating values. Recommendation systems are increasingly emerging as an integral component of e - business applications [1]. For instance, the integrated recommendation system of Amazon provides customers with personalized recommendations for various items of interest. Recommendation systems utilize various knowledge discovery techniques on a user's historical data and current context to recommend products and services that best match the user's preferences. In recent years, emergence of numerous mobile social networking services, such as, Facebook and Google Latitudehas significantly gained the attraction of a large number of subscribers [1], [6]. A mobile social networking service allows a user to perform a "check-in" that is a small feedback about the place visited by the user [1], [2]. Large number of check-ins on daily bases results in the accumulation of massive volumes of data. Based on the data stored by such services, several Venue- based Recommendation Systems (VRS) were developed [1]–[3].Such systems are designed to perform recommendation of venues to users that most closely match with users' preferences. Despite having very promising features, the VRS suffer with numerous limitations and challenges. A major research challenge for such systems is to process.

1.2 Related Work

Cloud computing enables a service-provisioning model for computing services that relies on the Internet.

1.2.1 Cloud services

This model typically involves the provisioning of dynamically scalable and virtualized services. Applications or services offered by means of cloud computing are called cloud services. Typical examples of cloud services include office applications (word processing, spreadsheets, and presentations) that are traditionally found among desktop applications. Nearly, all large software corporations, such as Google, Microsoft, Amazon, IBM, and Oracle, are providing various kinds of cloud services. Besides, many small businesses have launched their own Web-based services, mainly to take advantage of the collaborative nature of cloud services. The user of a cloud service has access to the service through a Web interface or via an API. Once started, the

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 1, January 2017

cloud service application acts as if it is a normal desktop application. The difference is that working documents are on the cloud servers.

1.2.2 Cloud services models are:
1.2.2.1Infrastructure-as-a-Service (IaaS):
With IaaS, organizations rent computing resources and storage space and access them through a private network or across the Internet.

1.2.2.2 Platform-as-a-Service (PaaS):
With PaaS, organizations can develop their business applications in a cloud environment by using software tools supported by their cloud provider. Maintenance and management of the cloud infrastructure including severs and operating system is the responsibility of the cloud provider.

1.2.2.3 Software-as-a-Service (SaaS):
With SaaS, the cloud service application runs on the cloud provider servers and users access the service through a Web interface or via an API.

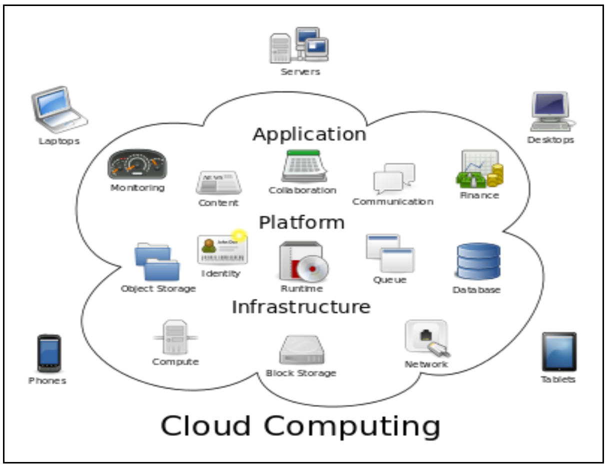

Fig 1: Cloud Computing

In the past, most work focused on trajectory-based approaches for venue recommendation systems. The trajectory based approaches record information about a user's visit pattern (in the form of GPS coordinates) to various locations, the routes taken, and dwell times. The authors in [3] applied data mining and machine learning on trajectory data to recommend most popular places. Although, trajectory-based approaches recommend locations to users based on their past trajectories, a major drawback of such approaches is that they are unable to simultaneously consider other influential factors apart from simple GPS trace that makes them produce less optimal recommendations. To address such deficiency, we utilized multi-objective optimization in our proposed framework. Another issue is that the trajectory-based approaches suffer from data sparseness problem as usually a person does not frequently visits many places, which results in sparse user-venue matrix. Moreover, the trajectory based approaches suffer from scalability issues as huge volumes of trajectory data needs to be processed causing considerable overhead. Some of the approaches, such as [3], [5] are based on the online ratings provided by the users to the visited places. The authors in [7] combine the available venue ratings with users' social ties to recommend venues that are high-ranked as well as most preferred by a user's friends. However, the authors did not compare their approach with any of the baseline approaches, and does not discuss complexity of their work. The aforementioned approaches perform different modelling to users' preferences, but they are not considering multiple objectives that we specifically considered in our study.

Moreover, they also suffer from data sparseness issues due to limited number of entries within the user-rating matrix. Apart from rating based approaches, few of the techniques have their models built on check-in based approaches where the users provide small feedbacks as check-ins about the places they visited [2]–[4], [7], [14]. For example, the authors in [6] applied random-walk-with-restart on a user-venue check-in matrix to generate personalized recommendations. Most of the above mentioned approaches have their designs built on memory-based CF that enables such approaches to provide recommendations to users on the basis of their past entries. However, such approaches suffer from common drawbacks of memory-based CF (e.g. cold start and data sparsity) which reduce their performance. Moreover, large number of similarity computations on user-to-venue matrix makes such approaches less scalable. There has been some limited work performed on applying multi-objective optimization on recommendation systems. One such contribution is where authors performed a weighted combination of numerous recommendation algorithms and applied optimization to find appropriate weights for the constituent algorithms. However, their approach is computation intensive and no time complexity was discussed. To address the issues cited above, we proposed a hybrid approach over a cloud architecture that combines the benefits of memory-based and model-based collaborative filtering along with multi-objective optimization to obtain an optimal list of venues to be recommended. Moreover, our proposed framework presents a solution for scalability, data sparseness, and cold start issues.

1.3 Research Problem
Existing Recommender systems make use of community opinions to help users identify useful items from a considerably big data available. The methods implemented by many of these systems are collaborative filtering (CF), which analyses past community opinions to find correlations of similar users and items. In the last decade, the digital revolution has provided relatively inexpensive and accessible means of collecting and storing data. This unlimited growth of data has led to a situation in which the knowledge extraction process is more difficult and, in most cases, leads to problems of scalability and/or

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 1, January 2017

complexity [1]. In scientific literature, several works, such as [1]–[6], and [13] have applied Collaborative Filtering (CF) to the recommendation problem in VRS. The CF-based approaches in VRS tend to generate recommendations based on the similarity in actions and routines of users [1], [2], [5]. However, despite being less complicated, most CF-based recommendation techniques suffer from several limitations that make them less ideal choice in many real-life practical applications [13]. The following are the most common factors that affect the performance of many existing CF-based recommendation systems:

Cold start- The cold start problem occurs when a recommendation system has to suggest venues to the user that is newer to the system [2]. Insufficient check-ins for the new user results in zero similarity value that degradesthe performance of the recommendation system [13]. The only way for the system to provide recommendation in such scenario is to wait for sufficient check-ins by the user at different venues.

Data sparseness-Many existing recommendation systems suffer from data sparseness problem that occurs when users have visited only a limited number of venues [3]. This results into a sparsely filled user-to-venuecheck-in matrix. The sparseness of such matrix creates difficulty in finding sufficient reliable similarusers to generate good quality recommendation.

Scalability-Majority of traditional recommendation systems suffer from scalability issues. The fast and dynamic expansion of number of users causes recommender system to parse millions of check-in records to find the set of similar users.Some of the recommendation systems [3]employ data mining and machine learning techniques to reduce the dataset size. However, there is an inherent trade-off between reduced dataset size and recommendation quality [1].

1.4 Methods and Contributions

In this paper, we propose a hybrid cloud-based Bi-Objective Recommendation Framework (BORF) that overcomes the limitations exhibited by traditional CF-based approaches. The architecture to generate optimal recommendations for the current user. The memory-based CF model utilizes a user's historical data and user-to-venue closeness to predict venues for the current user. We adopt a bi-objective optimization approach that considers the two primary objectives: (a) venue preference and (b) location closeness. Venue preference determines how much the venue meets the criteria of user's interests, whereas venue closeness indicates how closely a desired venue is located relative to a user's location. To the best of our knowledge this is the first work to incorporate the bi-objective optimization techniques into VRS. The rest of the paper is organized as follows. Section 2 presents the system overview. In Section 3, we discuss the proposed BORF framework. Section 4 presents the complexity analysis of the proposed framework and the performance evaluation with simulation results. The related work is reviewed in Section 5, Section 6 concludes the paper and Section 6 present Conclusion.

II. SYSTEM OVERVIEW

Most of the existing recommendation systems (e.g., [2], [3], [5], and [7]) utilize centralized architectures that are not scalable enough to process large volume of geographically distributed data. The centralized architecture for venue recommendations must simultaneously consider users' preferences, check-in history, and social context to generate optimal venue recommendations. Therefore, to address the scalability issue, we introduce the decentralized cloud-based BORF approach. The following are some of the major components of the proposed framework.

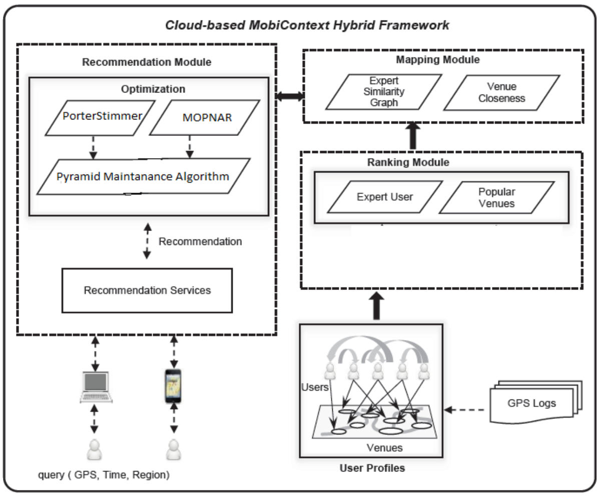

Fig. 1. Top level architecture of the Cloud-based BORF Framework

2.1 User Profiles

As reflected in Fig. 1, the framework maintains records of users' profiles for each geographical region. The arrows from users to venues at lower right of Fig. 1 indicate the number of check-ins performed by each user at various venues. A user's profile consists of the user's identification, venues visited by the user, and check-in time at a venue

2.2 Ranking Module

On top of users' profiles, the ranking module performs functionality during the pre-processing phase of data refinement. The pre-processing can be performed in the form of periodic batch jobs running at monthly or weekly basis as configured by system administrator. The ranking module applies model-based HA inference method on users' profiles to assign ranking to the set of users and venues based on mutual reinforcement relationships.

2.3 Mapping Module

The mapping module computes similarity graphs among expert users for a given region during pre-processing

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified

Vol. 6, Issue 1, January 2017

phase. The purpose of similarity graph computation is to generate a network of like-minded people who share the similar preferences for various venues they visit in a geographical region. The mapping module also computes venue closeness based on geographical distance between the current user and popular venues.

2.4 Recommendation Module

Fig. 1 depicts the online recommendation module that runs a service to receive recommendation queries from users. A user's request consists of: (a) current context (such as, GPS location of user, time, and region), and (b) a bounded region surrounding the user from where the topN venues will be selected for the current user (N is number of venues).

The recommendation service passes the user's query to optimization module that utilizes scalar and vector optimization techniques to generate an optimal set of venues. In our proposed framework, the scalar optimization technique utilizes the CF-based approach to generate user preferred recommendations.

III. ALGORITHM

1. PORTER STEMMER

Porter stemmer was developed by Martin Porter in 1980 at the University of Cambridge [4].Porter's algorithm is applied in many fields as a pre-processing step for the indexing task; itsmain use is as part of a term normalization process that is usually done when setting up anInformation retrieval system. The Porter stemmer is actually the most commonly used of all the stemmers. It showed improvements in retrieval performance and in other fields such asclassification, clustering, spam filtering. It becomes the most popular and the standard approachof stemming.The stemmer is based on the idea that the suffixes in the English language are mostly built of acombination of smaller and simpler suffixes; for instance, the suffix "fullness" is composed of two suffixes "full" and "ness". Thus, Porter stemmer is a linear step stemmer; it appliesmorphological rules sequentially allowing removing affixes in stages.The performance of information retrieval systems can be improved by matching key terms to any morphological variant. A stemming algorithm is a technique for automatically conflating morphologically related terms together. Several stemming algorithms exist with different techniques. The most common algorithm for English is Porter, Porter (1980). It has been widely adopted for information retrieval applications in a wide range of languages. However, it still has several drawbacks, since its simple rules cannot fully describe English morphology.

The present paper introduced an improved version of Porter stemming algorithm for English language. The resultant stemmer was evaluated using the error counting method. With this method, the performance of a stemmer is computed by counting the number of understemming and overstemming errors committed during the stemming process.

In this paper, we use Porter Stemmer algorithm. Porter Stemmer algorithm removes the stop words into user given review or feedback. After removing stopwords and after performing Porter Stemmeralgorithm, we pass review or feedback file to theMultiple Objective Positive Negative Association Rule (MOPNAR) algorithm.

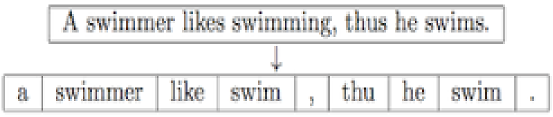

Fig 2. Example of Porter Stemmer Algorithm

Stemming is a technique to identify different Inflections and Derivations of the same word in order to transform them to one common root called Stem. A word's Stem is its most basic form which may or may not be a genuine word. Stems are obtained by stripping words of derivational and inflectional affixes to allow matching between the ones having the same semantic interpretation. In text mining applications using stemming, documents are represented by stems rather than by the original words. Thus, the index of a document containing the words "try", "tries" and "tried"will map all these words to one common root which is "try". This means that stemming algorithm can drastically reduce the dictionary size especially for highly inflected languages which leads to significant efficiency benefits in processing time and memory requirements.The most used is the Porter stemmer which contains support for English, French, Dutch and etc...

Specifically, the algorithm has five steps. Each step defines a set of rules. To stem a word, therules are tested sequentially; if one of these rules matched the current word, then the conditionsattached to that rule are tested. Once a rule is accepted; the suffix is removed and control movesto the next step. If the rule is not accepted then the next rule in the same step is tested, until eithera rule from that step is accepted or there are no more rules in that step and hence the control passes to the next step. This process continues for all the five steps; in the last step the resultantstem is returned by the stemmer. The whole algorithm can be resumed by the following activity diagram (Figure 3):

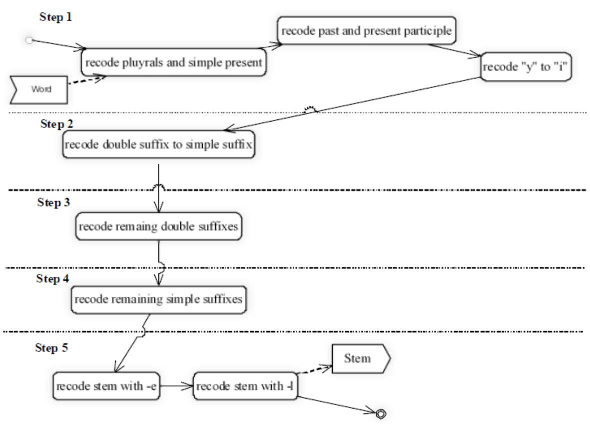

Fig 3. The steps of Porter Stemming Algorithm

Words' suffixes are removed step by step. The first step of the algorithm handles plurals, past participles, present participles, and transforms a terminal "y" to an "I". For example: "generalizations" is converted to "generalization", "agreed" to "agree", and "happy" to "happi". The second step deals with double suffixes to single ones.

For example: "generalization" is converted to "generalize", "oscillator" to "oscillate". The third step removes other double suffixes not handled in the previous step such as: "generalize" is changed into "general". The fourth step removes remaining suffixes such as: "general" is transformed into "gener", "oscillate" to "oscill". The fifth step treaties stems ending with –e, and treats words ending in double consonant. For example: "attribute" is recoded "attribut", "oscill" is converted to "oscil".

- Porter stemmer errors

Although Porter stemmer is known to be powerful, it still faces many problems. The major ones are Overstemming and Understemming errors. The first concept denotes the case where a word is cut too much which may lead to the point where completely different words are conjoined under the same stem. For instance, the words "general" and "generous" are stemmed under the same stem "gener". The latter describes the opposite case, where a word is not cut enough. This can lead to a situation where words derived from the same root do not have the same stem. This is due essentially to the fact that Porter stemmer ignores many cases and disregards many exceptions.There is a possibility to evaluate stemming by counting the numbers of two kinds of errors that occur during stemming, namely

- Under Stemming

This refers to words that should be grouped together by stemming, but aren't. This causes a single concept to be spread over various different stems, which will tend to decrease the Recall in an IR search.

- Over Stemming

This refers to words that shouldn't be grouped together by stemming, but are. This causes the meanings of the stems to be diluted, which will effect Precision of IR. Using a sample file of grouped words, these errors are then counted.

For example, Porter stemmer does not treat irregular verbs: "bought" remains "bought", and "buy" is stemmed "bii", whereas normally the two words have the same stem "buy". Irregular plural nouns are not handled by the stemmer: words ending with –men are the plural of words ending with –man. Porter stemmer makes other errors concerning the terminal –e. For instance, "do" is stemmed "do" and "does" is stemmed "doe". Many exceptions are not controlled: verb conjugation, possessive nouns, irregular comparative and superlative forms (e.g. good, better, best), etc. Moreover, more than 5000 suffixes are not handled by Porter such as –ativist, -ativistic, -ativism, atavistically, –ship, –ist, –atory, –ingly, got, gotton, ound, ank, unk, ook, ept, ew, own, etc. This would decrease the stemming quality, since related words are stemmed to different forms. In an information retrieval context, such cases reduce the performance since some useful documents will not be retrieved: a search for "ability", for example, will not return documents containing the word "able". This would decrease the efficiency of diverse systems applying Porter stemmer.

Stemmers are used to conflate terms to improve retrieval effectiveness and /or to reduce the size of indexing file.Stemming will increase recall at the cost of decreased precision.Stemming can have marked effect on the size of indexing files,sometimes decreasing the size of file as much as 50 percent.Porter's algorithm is important for two reasons. First, it provides a simple approachto conflation that seems to work well in practice and that is applicable to a range oflanguages. Second, it has spurred interest in stemming as a topic for research in itsown right, rather than merely as a low-level component of an information retrievalsystem. The algorithm was first published in 1980; however, it and its descendantscontinue to be employed in a range of applications that stretch far beyond its originalintended use.

2. MOPNAR algorithm

Multiple Objective Positive-Negative Association Rule algorithm used for classify the negative and positive Review. MOPNAR stand for(Multiple Objective Positive Negative Association Rule), using this algorithm we can classify all negative and positivereview into individual file.This paper includes the approach and techniqueof Multi-Objective Positive Negative Association Rule(MOPNAR) based predictive Sentiment Analysis, which is Based on huge dataset of multiple opinion obtained. In this Study multiple opinions from customer, data analyst, writers, and composers has been used which are in the form of text for identification of predictive sentiments.

MOPNAR is an extension of multi-objective evolutionary algorithm (MOEA). It helps in mining with a low computational cost areduced set of positive and negative QARs that are easy to understand and have good trade-off between the number of rules, support,and coverage of the dataset. The main focus of the algorithm is to obtain a reduced set of PNQARs which are having good trade-offconsidering three objectives which are comprehensibility, interestingness and performance. In order to perform a learning of rules itextends the traditional MOEA model. It also introduces two new components namely EP and Restarting process.To decompose the MOEA it decomposes the multi-objective optimization problem into N scalar optimization. It uses EA to optimizethe sub problems gathered. In this system to store all the no dominated rules found, provoke diversity in the population, and improvethe coverage of

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 1, January 2017

the datasets the EP and the restart is introduced. Here EP will contain all the no dominated rules found and it will alsogenerate the updated offspring for each solution. Since the size of EP is not fixed we can store a large number of rules and can reducethe size of population. Whereas restarting process here deals with the local optima and provoke diversity in the population. Thisprocess is applied when number of new individuals of the population in one generation is less than % of the size of the current population.

From the last decade, the digital revolution has provided relatively inexpensive and accessible means of collecting and storing data. This unlimited growth of data has led to a situation in which the knowledge extraction process is more difficult and, in most cases, leads to problems of scalability and/or complexity [2]. Association discovery is one of the most common data mining techniques used to extract interesting knowledge from large datasets [3]. Association rules are used to identify and represent dependencies between items in a dataset [4]. Multi-objective sentimentanalysis and predictive mining is the field of study that analyses people's opinions, sentiments, evaluations, attitudes, and emotions from written language. It is one of the most active research areas in Natural Language Processing (NLP) and is also widely studied in data mining, Web mining, and text mining. Multi-Objective Sentiment analysis systems are being applied in almost every business and social domain because opinions are central to almost allhuman activities and are key influences of people behaviours. People beliefs and perceptions of reality, and the choices make by the user, arelargely conditioned on how others see and evaluate the world. For this reason, when person need to make a decision they often seek out the opinionsof others. This is true not only for individuals but also for organizations.

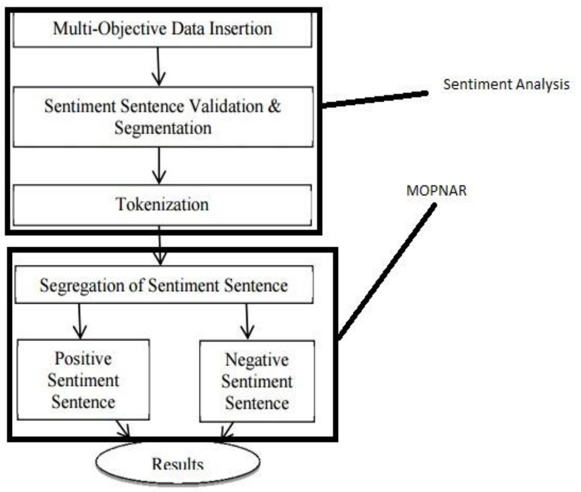

Fig 4. Diagram of Execution of mopnar

1. Multi-Objective Data Insertion
For Multi-Objective sentiment analysis programmer choose such kind of data which consists of all kind of hidden sentiments like positive, negative, neutral, exclamatory, questionnaire, etc. This kind of data can be inserted in three ways via manual, automatic file selection or via database storage. Pollution dataset is selected from web reviews.

2. Sentiment
Sentence Validation & Segmentation In this phase, validation of the data is done whether it is in specific desired format or not. Also if the huge data in paragraph form then segmentation of paragraph into sentence is performed. Once the entire paragraph is segmented then it will be ready for further analysis.

3. Tokenization
This is the phase where the specific words, punctuations, exclamation, etc. are captured & create token for each of them. In this step programmer also remove the stop words like punctuations which help us for further segregation process.

4. Segregation of Sentiment Sentence
Here the comparison of available sentences with evolutionary algorithm & grammatical rules for sentiment analysis is performed. In this way, segregation of the tokenized part to obtain multi-objective positive & negative sentiments hidden in the statements is accomplished.

5. Efficiency assessment with Association Rule and its result
Once they ready with positive and negative sentiments then apply the MOEA rule focusing on negative sentiments. With the help of association rules following results will comes out which conclude more efficiency compared to existing algorithm.

• Algorithm
N: number of sentences in the form of paragraph; x={x1, x2, x3… xN}.
Output: 1) Positive, negative and neutral sentences are Separated.
2) Positive, negative and neutral sentence's count.
3) Values for different parameters.
Steps:
1) Upload input data file containing 'n' number of Paragraphs.
2) Data will be validated using sentiment sentence validation.
3) Sentences will be segmented.
4) Tokens will be separated from all sentences.
5) Process of segregation is performed for finding the positive, negative and neutral words by matching the words with POS tagging words.
6) Calculate all parameters values using their formulas.

3. Pyramid Maintenance Algorithm
Pyramid maintenance algorithm used to selection of most popular hotel by taste, ambiance, Service, Cost for each.

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

IJARCCE

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 1, January 2017

We use the pyramid maintenance algorithm for selection and searching purpose. Maintenance algorithms take input a pyramid data cell C and consists of level H. This section gives details about pyramid data structure, pyramid maintenance algorithm and travel penalty algorithm.

• PYRAMID DATA STRUCTURE
(a) The pyramid decomposes the space into H levels.
(b) The space is partitioned into 4h equal area grid Cells.
(c) Level 0, one grid cell represents the entire Geographic area
(d) Level 1 partitions space into four equi-area cells. Four equi-area cells, and so forth.

We represent each cell with a unique identifier cid. The pyramid data structure maintains three types of cells:(1) Recommendation Model Cell (α-Cell), (2) Statistics Cell (β-Cell), and (3) Empty Cell (γ-Cell)

In this algorithm 1, an α-Cell requires the highest storage and maintenance overhead because it maintains a CF model as well as the user or item ratings statistics. On the other hand, an α-Cell (as opposed to β-Cell and γ-Cell) is the only cell that can be leveraged to answer recommendation queries.

A pyramid structure that only contains α-Cells achieves the highest recommendation locality, and this is why an α-Cell is considered the highly ranked cell type in LARS*. A β-Cell is the secondly ranked cell type as it only maintains statistics about the user/item ratings. The storage and maintenance overhead incurred by a β-Cell is less expensive than an α-Cell. The statistics maintained at a β-Cell determines whether the children of that cell need to be maintained as α-Cells to serve more localized recommendation. Finally, a γ-Cell (lowest ranked cell type) has the least maintenance cost, as neither a CF model nor statistics are maintained for that cell.

Moreover, a γ-Cell is a leaf cell in the pyramid. RS* upgrades (downgrades) a cell to a higher (lower) cell rank, based on trade-offs between recommendation locality and system scalability. If recommendation locality is preferred over scalability, more α-Cells are maintained in the pyramid. On the other hand, if scalability is favoured over locality, more γ-Cells exist in the pyramid. β-Cells comes as an intermediary stage between α-Cells and γ-Cells to further increase the recommendation locality whereas the system scalability is not quite affected.

1) α-Cell: stores an item-based collaborative filtering model built using only the spatial ratings with user locations contained in the cell's spatial region. α-Cell and represents a "traditional" (i.e., non-spatial) item-based collaborative filtering model.

2) β-Cell: The user/item ratings that are located within the spatial range of the cell. RS prefers a β-Cell over an α-Cell to reduce the total system storage.

3) γ-Cell: γ -Cell is a cell that maintains neither the statistics nor the recommendation model for the ratings lying within its boundaries. γ-Cell cannot have any children. The figure below shows pyramid data structure with α-Cell, β-Cell and γ-Cell.

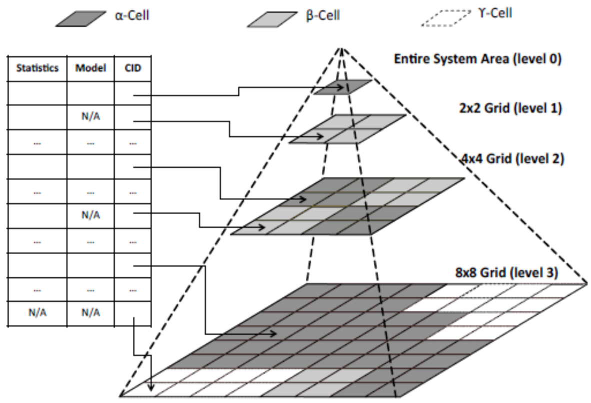

Fig 4. Pyramid data structure

• PYRAMID MAINTENANCE ALGORITHM

Algorithm Pyramid maintenance algorithm

```
/* Called after cell C receives N% new ratings */
Function PyramidMaintenance(Cell C, Level h)
/* Step I: Statistics Maintenance*/
Maintain cell C statistics
/*Step II: Model Rebuild */
if (Cell C is an α-Cell) then
    Rebuild item-based collaborative filtering model for cell C
end if
/*Step III: Cell Child Quadrant Maintenance */
if (C children quadrant q cells are α-Cells) then
    CheckDownGradeToSCells(q,C) /*
else if (C children quadrant q cells are γ-Cells) then
    CheckUpGradeToSCells(q,C)
else
    isSwitchedToMcells ← CheckUpGradeToMCells(q,C) /*

    if (isSwitchedToMcells is False) then
        CheckDownGradeToECells(q,C)
    end if
end if
return
```

IV. ANALYSIS

In this paper, one existing algorithm is used for comparing and analyzing the performance of newly proposed algorithm. The textual data file is used as an input data file, which is uploaded at runtime. The existing algorithm NSGA-II, only concentrated on randomly recommendation, but this proposed algorithm will work on systematic recommendation means the recommended hotels by users are arranged by proper manner. Previously developed algorithms only work on the discrete and binary data, but this new algorithm is work on the textual data which will contains the sentiment sentences. These sentiment sentences are the reviews from the public, which contains the sentiment or opinion of peoples.

We utilized the three standard performance evaluation matric to evaluate the proposed recommendation frameworks: (a) precision, (b) recall, and (c) F-measure.

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 1, January 2017

[26]. The precision presents a ratio of the accurate recommendations (true positive (tp)) to the total number of anticipated recommendations (tp+ false positive (fp)). An accurate recommendation is the recommendation that has been predicted correctly in the top-N recommended venues.

Precision is given as:

• Precision =

$$Precision = \frac{tp}{tp + fp}.$$

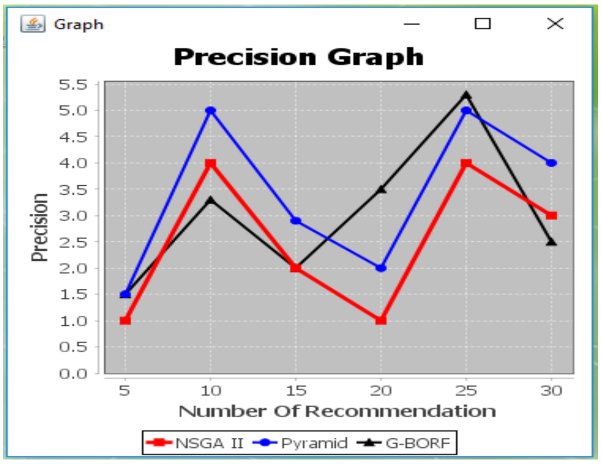

Fig. (a)

• Recall =

$$Recall = \frac{tp}{tp + fn}.$$

The recall measures the single user recommendation effectiveness by computing the average quality of the individual recommendations. Recall is defined as the ratio of correct recommendations (tp) to the total number of recommendations (tp + fn). The recall presents the proportion of all the accurate recommendations in the top-N recommended venues and can be represented as:

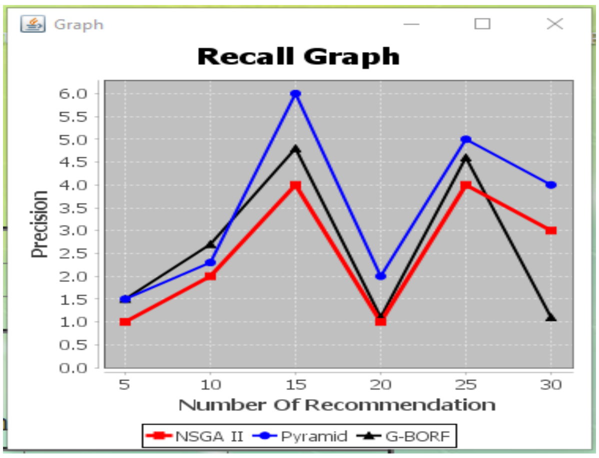

Fig. (b)

• F-measure =

$$F\text{-}measure = \frac{2 \times Precision \times Recall}{Precision + Recall}.$$

The F-measure is the harmonic mean of precision and recall and is denoted as follows:

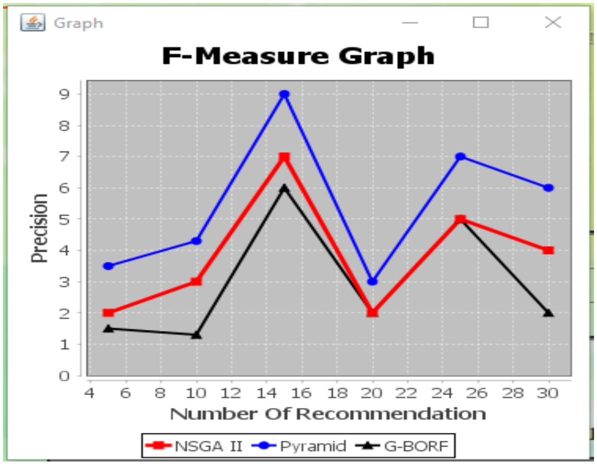

Fig. (c)

Fig.(a), (b), and (c) present the precision, recall, and f-measure results. These figures show better performance in terms of precision, recall, and f-Measure. Such improvement in results is due to the fact that the pre-processing phase reduces the negative effect of data sparseness over recommendation quality. Data sparseness results in zero similarity values in collaborative filtering, and with large number of zero entries in user-to-User.

As reflected in Fig. (a) ,(b), and (c),Pyramid Maintenance Algorithm demonstrates the better performance in terms of precision and recall as compared to the rest of the schemes (NSGA-II and G-BORF). In contrast, the NSGA-II and greedy-BORF approaches present slightly lower performance because of the aggregation method that maps the users' preferences and location closeness into single objective function. Such aggregation cannot provide accurate results especially when there is trade off between the user's preferences and location closeness. For instance, in the case of G-BORF, when there is no similarity between two users' preferred locations, the venue will be suggested to the current user on the bases of user-to-venue closeness. Such suggestion may not provide optimal recommendation and indicates lower performance in terms of precision and recall as presented in Fig. (a),(b) and (c).

Here values of different standard formulas are calculated and compared. Our proposed algorithm gives greater value for all parameters and comparison using graph is also shown figures. The proposed algorithm is gone through all transactions present in the input data files of dataset. This graph shows our algorithm is better as compared to existing algorithm.

V. APPLICATIONS

Hotel recommendation

To alleviate the problems faced by KBS developers due to the complex nature of some methodologies and also the lack of standards for the knowledge modelling, we have proposed steps to develop a knowledge model for a system recommending Mauritian hotels.

We have used UML for modelling the domain due to its common features between Object Oriented modelling and ontology modelling. We are presently working on the prototype of the application that will be used to validate the knowledge model.

Item recommendation

Recommender systems are a powerful new technology for extracting additional value for a business from its user data-bases. These systems help users and items they want to buy from a business. Recommender systems benefit users by enabling them to and items they like.

Conversely, they help the business by generating more sales. Recommender systems are rapidly becoming a crucial tool in E-commerce on the Web. Recommender systems are being stressed by the huge volume of user data in existing corporate databases, and will be stressed even more by the increasing volume of user data available on the Web.

Venue recommendation

Venue recommendation system, we collected behavioural, social, and spatial data from Gowalla and Foursquare (via., Twitter) for a range of the world's metropolises, and evaluated how a variety of learning algorithms from simple, non-personalized popular venue recommendation to predictions based on matrix factorization methods were able to rank and classify the new venues for each user.

Web-Page recommendation

Web-page recommendation plays an important role in intelligent Web systems. Useful knowledge discovery from Web usage data and satisfactory knowledge representation for effective Web-page recommendations are crucial and challenging. This paper proposes a novel method to efficiently provide better Web-page recommendation through semantic-enhancement by integrating the domain and Web usage knowledge of a website. Two new models are proposed to represent the domain knowledge.

Time-aware recommendation

Time-aware recommendations systems (TARS) are indeed receiving increasing attention. A wide range of approaches dealing with the time dimension in user modelling and recommendation strategies have been proposed. Recommender systems (RS) aim to help users with information access and retrieval applications when large collections of items are involved. In general, they work by means of suggesting those items that should be the most appealing ones to the users based on their past personal preferences. The methodologies used for TARS evaluations make use of several intermediate approaches that highly differ from one work to another, and existing differences may have a significant effect on the assessment of TARS performance.

Location-aware recommendation

LARS, a location-aware recommender system that uses location-based ratings to produce recommendations is proposed. Traditional recommender systems do not consider spatial properties of users nor items; LARS, on the other hand, supports a taxonomy of three novel classes of location-based ratings, namely, spatial ratings for non-spatial items, non-spatial ratings for spatial items, and spatial ratings for spatial items.

LARS exploits user rating locations through user partitioning, a technique that influences recommendations with ratings spatially close to querying users in a manner that maximizes system scalability while not sacrificing recommendation quality. LARS exploits item locations using travel penalty, a technique that favors recommendation candidates closer in travel distance to querying users in a way that avoids exhaustive access to all spatial items. LARS can apply these techniques separately, or together, depending on the type of location-based rating available. LARS is efficient, scalable, and capable of producing recommendations twice as accurate compared to existing recommendation approaches.

VI. CONCLUSION

This introduced a cloud-based framework that produces optimized recommendations by simultaneously considering the factors, like person's geographical location and location closeness. The significance of the proposed framework is the adaptation of collaborative filtering (CF) and bi-objective optimization approaches, like scalar and vector.

In this approach, data sparseness issue is determined by integrating the user-to-user similarity computation with confidence measure that calculates the amount of similar interest indicated by the two users. Moreover, the solution to cold start issue is discussed by introducing the Hub-Average inference model that assigns ranking to the users and has a precompiled number of popular not visited venues that can be recommended to the new user.

In this paper, the whole focus is on implementation of a new multi-objective evolutionary algorithm with positive negative Association rule. This proposed algorithm is working on textual data, which is stored in file. The proposed algorithm is compared with NSGA-II algorithm Based on the parameters. After comparing this algorithm, on two different datasets, it is found that the proposed algorithm performs better. Proposed algorithm will

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 1, January 2017

process all transactions present in input data file and the values calculated for all parameters are giving better values. This algorithm is working on textual data having three different kinds of sentences such as positive, negative and neutral, but previous algorithm is working on the binary and discrete values.

ACKNOWLEDGMENT

The author would like to offer special thanks to mentor **Ms S. S. Ponde** from our institute for directing the special concerns related with this topic and organizing this issue. I also owe great deal to my friends for contributing their efforts in enthusiastic manner.

REFERENCES

[1] A. Majid, L. Chen, G. Chen, H. Turab, I. Hussain, and J. Woodward, “A Context-aware Personalized Travel Recommendation System based on Geo-tagged Social Media Data Mining,” International Journal of Geographical Information Science, pp. 662-684, 2013.

[2] P. G. Campos, F. Díez, I. Cantador, “Time-aware Recommender Systems: A Comprehensive Survey and Analysis of Existing Evaluation Protocols,” User Modelling and User-Adapted Interaction, vol. 24, no.1-2, pp. 67-119, 2014.

[3] C. Chow, J. Bao, and M. Mokbel, “Towards Location-Based Social Networking Services,” In Proceedings of the 2nd ACM SIGSPATIAL International Workshop on Location Based Social Networks, ACM, pp. 31-38, 2010.

[4] M. Ye, P. Yin, and W. Lee, “Location recommendation for location-based social networks,” In Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems, ACM, pp. 458-461, 2010.

[5] A. Noulas, S. Scellato, N. Lathia, and C. Mascolo, “A Random Walk around the City: New Venue Recommendation in Location-Based Social Networks,” In Proceedings of International Conference on Social Computing (SocialCom), pp.144-153, 2012.

[6] Paired t test.Wiley Encyclopaedia of Clinical Trials, 2008.

[7] J. Abimbola, “A Non-linear Weights Selection in Weighted Sum for Convex Multi-objective Optimization,” Mathematics and Information, vol. 27, no. 3, 2012.

[8] H. Nasiri, M. Maghfoori, “Multiobjective Weighted Sum Approach Model reduction by Routh-Pade approximation using Harmony Search,” Turkish Journal of Electrical Engineering and Computer Science, vol. 21, no.2,pp. 2283-2293, 2013.

[9] M. Ribeiro, A. Lacerda, A. Veloso and N. Ziviani, “Pareto-Efficient Hybridization for Multi-objective Recommender Systems,” In Proceeding of 6th ACM Conference on Recommender Systems, pp. 19-26, 2012.

[10] K. Deb, A. Pratap, S. Agarwal and T. Meyarivan, “A Fast and Elitist Multi-objective Genetic Algorithm: NSGA-II,” IEEE Transaction on Evolutionary Computations, vol. 6, no. 2, pp. 182 197, 2002.

[11] Y. Wang, S. Wang, N. Stash, L. Aroyo and G. Schreiber, “Enhancing Content-Based Recommendation with the Task Model of Classification,” In Proceedings of the Knowledge and Management, pp. 431-440, 2010.

[12] J. Bobadilla, F. Ortega, A. Hernando, A. Gutiérrez, “Recommender Systems Survey,” Knowledge-Based Systems,vol.46, pp. 109-132, 2013.

[13] J. Bao, Y. Zheng, M.F. Mokbel, “Location-based and Preference Aware Recommendation using Sparse Geo-Social Networking Data,” In Proceeding of 20th International Conference on Advances in Geographic Information Systems, ACM New York, pp.199-208, 2012.

[14] B. Hidasi and D. Tick, “Initializing Matrix Factorization Methods on Implicit Feedback Database,” Journal of Universal Computer Science, vol. 19, no. 12, pp. 1835-1853, 2013.

BIOGRAPHIES

Ms. Khushabu C. Sonawane is currently studying in Masters of Computer Science and Engg. at DIEMS, Aurangabad. She has completed her B.E from SSBT COET, Jalgaon in 2014. Her research field includes Cloud computing, Data mining.

Ms. Smita S. Ponde has received the B.E. degree in Computer Science & Engineering from Government Engineering College, Aurangabad in 1999 and M.E. degree in Computer Science & Engineering from Government Engineering College, Aurangabad in 2009. Currently she is pursuing her Ph.D. in Computer Science & Engineering with the topic ‘Wireless sensor networks’ and is working as Assistant Professor in Deogiri Institute of Engineering and Management Studies. Her areas of specialization include Soft Computing, Wireless sensor networks.

.