2018-19

IIRW'S

INTERNATIONAL CENTRE OF EXCELLENCE IN ENGINEERING AND MANAGEMENT (ICEEM

NAAC ACCREDITED

FACULTY APPRAISAL FORM

Faculty Appraisal Form for the Academic Year:		
Faculty Name: Jagtap D. D	Department: C. S. E	Position: Asst. professor
Mobile No.	Email ID: dinesh.jagtap@gmail.com	Date Conducted:

A. Academics (150 points)	Max	Secured
Section 1: Teaching (75)		
1. Teaching Effectiveness - Calculated based on adherences to academic calendar and student's performance and incorporation of feedback)	25	20
2. Innovations in Teaching & Learning– Implementation of active learning pedagogies to enhance students' learning	15	12
3. Strategies adopted to support slow and advanced learners.	15	12
4. Improvement in teaching practices based on mid-semester feedback collected from the students	10	06
5. Student feedback collected at the end of the semester	10	06
Section 2: Mentoring (25)		
6. Student mentoring: Effectiveness of mentoring students to monitor their progress and help them to succeed in the program	25	20
Section 3: Professional Development (50)		
7. Participation in teaching workshop/seminar to improve teaching through the 'Center for Engineering Education Development'.	10	08

Excellent (5)	Good (4)	Average (3)	Bad (2)	Poor (1)
	✓			

Excellent (5)	Good (4)	Average (3)	Bad (2)	Poor (1)
	✓			

Excellent (5)	Good (4)	Average (3)	Bad (2)	Poor (1)
	✓			

Excellent (5)	Good (4)	Average (3)	Bad (2)	Poor (1)
		✓		

Excellent (5)	Good (4)	Average (3)	Bad (2)	Poor (1)
		✓		

Excellent (5)	Good (4)	Average (3)	Bad (2)	Poor (1)
	✓			

Excellent (5)	Good (4)	Average (3)	Bad (2)	Poor (1)
	✓			

CAMPUS DIRECTOR
International Centre of
Excellence In Engg. & MGMT.
Aurangabad

8. Participation in external Faculty Development Programs (FDPs) Excellent (5) | Good (4) ✓ | Average (3) | Bad (2) | Poor (1)	05	04	2
9. Participation in Short Term Training Programs (STTPS) Excellent (5) | Good (4) ✓ | Average (3) | Bad (2) | Poor (1)	10	08	2
10. Registration and completion of online courses/MOOCs such as NPTEL, SWAYAM etc. Excellent (5) | Good (4) | Average (3) ✓ | Bad (2) | Poor (1)	10	06	—
11. Participation in state/national/international conferences Excellent (5) | Good (4) | Average (3) | Bad (2) | Poor (1)	10	00	
12. Holds membership in professional organizations Excellent (5) | Good (4) | Average (3) | Bad (2) | Poor (1)	05	00	
Total A	**150**	102	

Minimum Eligible Criteria: 60 % score

Academic Score = Total A = 102

CAMPUS DIRECTOR
International Centre of
Excellence In Engg. & MGMT.
Aurangabad

B. Research (150 points)	Max	Secured	Evidence

Section I: Academic Research

- Primary, Secondary and Tertiary authorships are considered.
- If all the other authors in a publication are students of the faculty then any position is considered for authorship.
- In some of the research areas more number of authors (more than 3) can be considered based upon the recommendation by the committee formed by the Research Council

1. Total Publications in refereed scholarly article in a national or international Journal papers indexed in **Web of Science or Scopus. Or UGC Care**

Excellent (5)	Good (4)	Average (3)	Bad (2)	Poor (1)
	✓			

Max: 25 | Secured: 20 | Evidence: ✓

2. Publications in refereed scholarly article in a national or international Conference papers **indexed in Web of Science or Scopus**

Excellent (5)	Good (4)	Average (3)	Bad (2)	Poor (1)

Max: 15 | Secured: 00 | Evidence:

3. Publications in refereed scholarly article in a national/international Journal and Conference papers **approved by UGC**

Excellent (5)	Good (4)	Average (3)	Bad (2)	Poor (1)

Max: 15 | Secured: 00 | Evidence:

4. Applied and secured research funding/Industrial consultancy/Patent

Excellent (5)	Good (4)	Average (3)	Bad (2)	Poor (1)

Max: 30 | Secured: 00 | Evidence:

Sections II: Other research related activities

5. Develops patents / Industrial consultancy through scholarly research dustrial relation/Internship/OJT to students)/Frequent visit.

Excellent (5)	Good (4)	Average (3)	Bad (2)	Poor (1)

Max: 40 | Secured: 00 | Evidence:

6. Guides UG students for research and projects (should result filing of patents or paper publication/product/semi finished product)

Excellent (5)	Good (4)	Average (3)	Bad (2)	Poor (1)
	✓			

Max: 25 | Secured: 20 | Evidence:

	Total B	150	40	

Note: Minimum Eligible Criteria: Minimum score for this category will be based upon the number of years of experience, qualification, cadre and responsibilities assigned.

Research Score = Total B = 40

CEED-Center for Engineering Education department
CRID- Center for Research & Innovation department
CIST- Center for Innovation in social Transformation
IRAC-Internal Quality Assessment Center

CAMPUS DIRECTOR
International Centre of Excellence In Engg. & MGMT. Aurangabad

C. Service to the Institution (200 points)	Max	Secured	Evidence
Institute Level			
1. Support provided as department level coordinators for effective functioning of the various departments like CRID, Training and Placements, IQAC, Accreditation, CEED, , Department of Student Affairs, Institute Innovation Council, Entrepreneurship Development Cell, Communications & CIST etc Excellent (5) / Good (4) ✓ / Average (3) / Bad (2) / Poor (1)	15	12	✓
2. Coordination of accreditation related activities Excellent (5) / Good (4) ✓ / Average (3) / Bad (2) / Poor (1)	20	16	✓
3. Coordination of admissions related tasks Excellent (5) / Good (4) ✓ / Average (3) / Bad (2) / Poor (1)	30	24	✓
4. Support in organizing institute level programs such as conferences, students activity, Startup Fest, Innovation Summit, Teaching and Learning Symposium, Service-Learning Symposium, Emerging Technology Summit etc. Excellent (5) / Good (4) ✓ / Average (3) / Bad (2) / Poor (1)	15	12	✓
5. Serves as in-charge for student club/organization (technical, cultural, and social impact clubs), head or members of committee's such as Student Grievance Cell, Anti-ragging cell, Women's cell, NSS Cell, Disciplinary committee, Sports Committee, Library Committee etc.) Excellent (5) / Good (4) ✓ / Average (3) / Bad (2) / Poor (1)	20	16	✓
6. Serves as members of Academic Council and Board of Studies of other institutions, external reviewer to journals or professional organizations committee's such as IEEE, CSE, SAE, IETE etc/university level assessment Excellent (5) / Good (4) / Average (3) ✓ / Bad (2) / Poor (1)	15	00	
Department Level			
7. Support provided in effective functioning of the department. Excellent (5) ✓ / Good (4) / Average (3) / Bad (2) / Poor (1)	10	10	✓

CAMPUS DIRECTOR
International Centre of
Excellence In Engg. & MGMT.
Aurangabad

Serves as part of Department Advisory Board, Project Review Committee, Program Assessment Committee, Finance Committee, and Department Development Committee Excellent (5) | Good (4) | Average (3) ✓ | Bad (2) | Poor (1)	15	09	✓
9. Organization of workshop/guest lecture/training program and Authors departmental reports or documents Excellent (5) | Good (4) ✓ | Average (3) | Bad (2) | Poor (1)	15	12	✓
10. Mentors junior faculty to succeed in their role. Excellent (5) | Good (4) ✓ | Average (3) | Bad (2) | Poor (1)	05	04	✓
11. Adherence towards examination related duties. Excellent (5) | Good (4) ✓ | Average (3) | Bad (2) | Poor (1)	10	08	✓
12. Performance evaluation by head of department Excellent (5) | Good (4) ✓ | Average (3) | Bad (2) | Poor (1)	10	08	✓
13. Performance evaluation by other reporting in-charge (if applicable). Excellent (5) | Good (4) ✓ | Average (3) | Bad (2) | Poor (1)	10	08	✓
14. Feedback.	10	08	
Choose as many areas as applicable			
Total C	**200**	147	

Minimum Eligible Criteria: Minimum score for this category will be based upon the number of years of experience, qualification, cadre and responsibilities assigned.

Minimum Eligible Criteria: 60 % score

Service Score = Total C = 147

CAMPUS DIRECTOR
International Centre of
Excellence In Engg. & MGMT.
Aurangabad.

Total Score obtained = (Total A + Total B + Total C) = [Score Obtained] = 289

Total Score in % = [(Score obtained) /500] * 100 % = 57.80 %.

Faculty Member's Signature: _____ Date: _____

Appraiser's Signature: _____ Date: _____

Note:

More than 80% - Excellent
More than 70% - Very good
More than 60% - good
} **Recommend for increment 7.5, 10, 15, 20%**

More than 50% - Average - One year chance
More than 40% - Poor - six month chance
Less than 40% - Bad- Termination

CAMPUS DIRECTOR
International Centre of
Excellence In Engg. & MGMT.
Aurangabad

Big Data using Hadoop

Dinesh D. Jagtap[1]
CSE Department,
Everest Educational Society's Group of Institutions
Aurangabad,Maharashtra, India
dinesh.jagtap@gmail.com

Prof.B.K.Patil[2]
CSE Department,
Everest Educational Society's Group of Institutions,
Aurangabad, Maharashtra,India.
cseroyal7@gmail.com

Abstract—: "Big Data" is data that becomes large enough that it cannot be processed using conventional methods. The term Big Data concerns with the huge volume, complex and rapidly growing data sets with multiple, independent sources .Due to fast development of networking ,data storage and data collection capacity the concept of big data is now rapidly expanding in all science and engineering domains including biological, physical and biomedical sciences. Social networking sites, mobile phones ,banking and stock exchange sectors, sensors and science contribute to production of peta bytes of data daily. That's why Big Data analysis now drives almost every aspect like mobile services, retail, financial services, manufacturing and life sciences. We all have heard a lot about "big data," but "big" is actually a red herring. Telecommunications companies, Oil companies, and other data-relevant industries have had vast datasets for a long time. And as storage capacity continues to enlarge, today's "big" is certainly tomorrow's "medium" and "small." In next week. The best meaningful definition of "big data" is when the size of the data itself becomes part of the problem.

Keywords— Big Data, data mining, heterogeneity, autonomous sources, complex and evolving associations

I. INTRODUCTION

In the last few years, there has been tremendous increased in the amount of data that's available. Whether we're talking about tweet streams, web server logs, records of online transactions, government data, or some other source data. The problem is not only finding data, it's figuring out what to do with the available data. And it's not just companies using their own data, or the data contributed by users of that company. Data mining allows users to examine the data from many dissimilar magnitudes or angles, sort it, and summarize the associations identified. Strictly, data mining is the process of finding correlations or patterns among dozens of fields in big relational databases. Another fundamental characteristics of the Big Data is large volume of data is represented by heterogeneous and diverse dimensionalities. This is because of different information collector prefers their own schemata for recording the data and also the nature of application also results in diverse representation of data. When the size of data increases obviously the complexity and relationships underneath the data.Hadoop is an open source software project that enables processing of large data sets distributed across the clusters of product servers. We're discussing data problems that are ranging from gigabytes to petabytes size of the data. At particular point, conventional techniques for working with data run out of the stream.

Information platforms are somewhat like as traditional data warehouses, but different. They describe rich APIs, and are designed for exploring and considering the data rather than for traditional analysis and reporting. They allow all data formats, with the most messy, and their schemas grow.

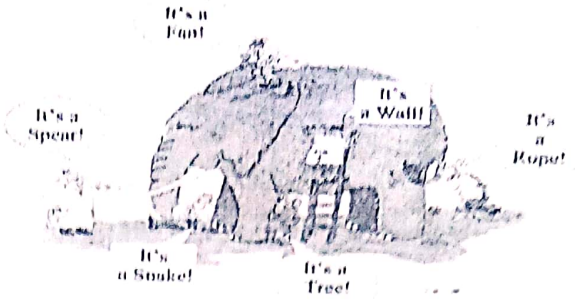

Figure 1 The blind men and the giant elephant: the localized (limited) view of each blind man leads to a biased conclusion.

II.REQUIREMENT

Most of the organizations that have built data platforms have establish it necessary to go further than the relational database model. Conventional relational database systems stop being valuable at this balance. Managing shading and replication across a mass of database servers is difficult and slow. The need to define a schema in advance conflict with reality of numerous, formless data sources, in which you may not know what's important until after you've analyzed the data. Relational databases are premeditated for uniformity, to support complex transactions that can easily be rolled back if any one of a composite set of operations fails.

To store vast datasets efficiently, we've seen a new type of databases appear. These are normally called NoSQL databases or Non-Relational databases, while neither term is very practical. Many of these databases are the logical offspring of Google's Big Table and Amazon's Dynamo, and are intended to be distributed across many nodes, to provide "ultimate uniformity" but not absolute consistency, and to have very flexible schema. Whereas there are two dozen or so products available (about all of them open source), a few leaders have recognized themselves:

III.OBJECTIVES

Data is only useful if you can do something with it, and massive datasets introduces computational issues. Google popularized the MapReduce approach, which is mainly a divide-and-conquer policy for distributing an tremendously large problem across an very large computing cluster. In the "map" stage, a programming task is divided into a number of equal subtasks, which are after that distributed across many processors; the halfway results are then combined by a single shrink task. In perception, MapReduce seems like an clear solution to Google's major trouble, creating large searches. It's so easy to allocate a search among number of processors, and after that merge the results into a single set of answers. What's less understandable is that MapReduce has proven that to be broadly valid to many large data troubles, ranging from searching to machine learning. Architecturally, the cause you're able to deal with lots of data is because Hadoop spreads it out. And the reason you're able to ask complicated computational question is only because you've got all of these processors, working in parallel, harness mutually.

IV.THEME

Using data effectively requires something different from traditional statistics, where actuaries in business suits perform arcane but fairly well-defined kinds of analysis. What differentiates data science from statistics is that data science is a holistic approach. We're increasingly finding data in the wild, and data scientists are involved with gathering data, massaging it into a tractable form,making it tell its story, and presenting that story to others.To meet the challenge of processing such large data sets, Google created Map-Reduce. Google's work and Yahoo's creation of the Hadoop MapReduce implementation has spawned an ecosystem of big data processing tools.

A. Literature Survey

Data is everywhere: your administration, your web server, your business partners and even your body, we are finding that almost everything can be instrumented. At O'Reilly, we normally merge publishing industry data from Nielsen BookScan

with our own sales data, openly available Amazon data, and even job data to see what's happening in the publishing industry. Sites like Infochimps and Factual gives access to numerous large datasets, including weather data, MySpace activity.

Storage and Reduce Big data" is data that becomes huge enough that it cannot be processed using straight methods. Social networks, mobile phones, Banking sector and government agencies contribute to peta bytes of data created daily.

- To face the number of challenge of processing such kind of huge data sets, Google invented Map Reduce. Google's work and Yahoo's creation of the Hadoop MapReduce implementation has spawned an environment of big data processing tools.
- As MapReduce has grown-up in reputation, a stack for big data systems has invented, comprising layers of Storage, MapReduce and Query (SMAQ).
    - SMAQ systems are normally open source, distributed, and run on commodity hardware.

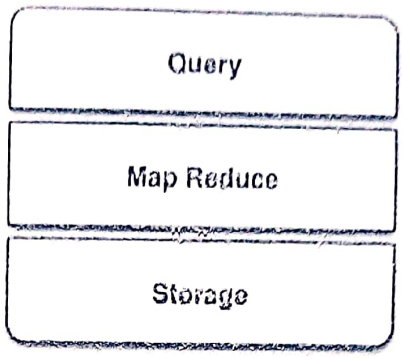

Figure.2. SMAQ systems

- Created at Google in response to the difficulty of creating web search indexes, the MapReduce framework is the thrust behind most of today's big data processing.
- The key improvement of MapReduce is the capability to take a query over a data set, divide it, and run it in parallel over many nodes.

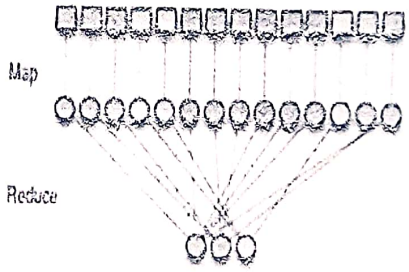

Figure. 3. Map Reduce Technique

- **Loading the data**—This operation is more properly called Extract, Transform, Load (ETL) in data warehousing language. Data should be extracted from its source, prepared to make it ready for further processing.
- **MapReduce**—This segment will retrieve data from storage, process it, and transfer its results to the storage.
- **Extracting the result**—Once processing is completed, for the result to be useful to humans, it must be retrieved from the storage and presented.
- Many SMAQ systems have characteristics designed to solve the operation of each of these stages.
- **Storage-MapReduce** requires storage from which to retrive data and in which to store the obtained results of the computation. The data predicted by MapReduce is not the relational data as generally used by conventional database system. Instead, data is consumed in chunks, which are then divided among nodes and fed to the map phase as key value pairs. This data does not need a schema, and may be formless.

CAMPUS DIRECTOR
International Centre of Excellence In Engg. & MGMT.
Aurangabad

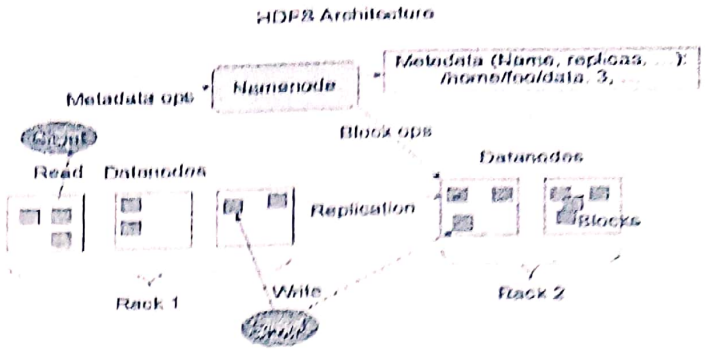

Figure.4 HDFS Architecture

- **Hadoop** is leading open source map reduce implementation created by yahoo emerged in 2006 creator is Doug cutting.

To communicate between node in 2nd generation uses replication factor Hadoop and HDFS utilize a master- slave architecture. HDFS is written in Java, with an HDFS cluster consisting of a primary name node a master server that manages the file system namespace and also regulates right of entry to data by clients. An elective secondary Name Node for fail over purposes also may be configured. Consecutively. HDFS has many goals. Here are some of the most prominent:

- Fault tolerance is easy to find by detecting faults and applying rapid and automatic recovery.
- Data access via MapReduce streaming.
- Processing logic is close to the data instead of the data close to the processing Logic.

V.BIG DATA CHARACTERISTICS: HACE THEOREM

HACE Theorem. Big Data starts with large-volume, heterogeneous, autonomous sources with distributed and decentralized control, and seeks to explore complex and evolving relationships among data.These above mentioned characteristics make it an extreme challenge for taking out meaningful facts from the Big Data. In a naïve sense, we can imagine that there are number of blind men are trying to size up a giant elephant (see Fig. 1), which will be the Big Data in this context. The main aim of each blind man is to construct a picture of the elephant on the basis of the part of information collects during the process. Because each person's view is restricted to his local region, so the blind men will each conclude alone that the elephant "feels" like a rope, a hose, or a wall, depending on the region each of them is limited to. To make this scenario more complex let us imagine that the elephant is growing rapidly and its pose changes constantly, and) each blind man may have his own information sources that tell him about biased knowledge about the elephant (e.g., so one blind man can share his feeling about the present pose of elephant with another blind man, where the knowledge which is shared is inherently biased. Describing the Big Data in this scenario is corresponding to aggregating heterogeneous information from number of sources (blind men) to help draw a best possible picture of the elephant in a real-time position.[3]

5.1 Huge Data with Heterogeneous and Diverse Dimensionality

One of the fundamental characteristics of the Big Data is the enormous volume of data is represented by heterogeneous and varied dimensionalities. Because different information collectors prefer their own schemata or protocols for recording of the data, different applications and their nature also results in miscellaneous data representations. The simple example is, each human being in a biomedical world can be represented by using simple demographic information such as gender, age, family disease history, and so on. For X-ray and CT scan examination of each patient, images or videos are provides visual information used for doctors to carry detailed examination that's why images and videos are useful entity. Under such situation, the heterogeneous features refer to the representations for the same individuals in different types, and the diverse features refer to the variety of the features involved to represent each single observation. Imagine that different organizations or health practitioners can have their own kind of schemata to represent each individual, if we want to enable aggregation of data by combining data from all sources then the data heterogeneity and diverse dimensionality are become major challenges.[3]

5.2 Autonomous Sources with Distributed and Decentralized Control

Autonomous data sources with distributed and decentra-lized controls aare a main quality of Big Data applications. Being autonomous, there is no any centralized control on it so, each data source can produce and gather information without involving and relying on any centralized control. This is same as World Wide Web (WWW) setting where each web server is independent and provides a certain amount of information without necessarily depanding on other servers.The enormous volumes of the data can also make an application more vulnerable to malfunctions, if the whole system has to be depended only on single centralized control unit.Todays well known social sites such as Google, Facebook, and Walmart, has set of large number of server farms which are situated all over the world to ensure nonstop services and quick responses for local markets. More particularly, the local government regulations also impact on the wholesale management process and it result in reorganized data representations and data warehouses for local markets.[3]

5.3 Complex and Evolving Relationships

While the amount of the Big Data increases, so the complexity and the relationships under the data. In the early stage of data centralized information systems, the main aim to finding best feature values to represent each observation. This is same as using a number of data fields, such as gender,age, income, education background, to describe each individual. This type of representation of sample-feature inherently treats each individual as an independent entity without considering their societal relations, which is one of the most important factors of the human society. In real world our friend circles may be formed based on the frequent hobbies or people are connected by biological dealings. Such social connections also are very popular in cyberworlds. For example, major social networking applications, such as Facebook or Twitter, are mostly characterized by social functions such as friend-connections and followers (in Twitter). In the sample-feature representation, individuals persons are regarded alike if they are sharing similar feature values, whereas in the sample-feature-relationship representation, two persons can be linked together even though they might share nothing in common in the feature domains at all. In a dynamic world, the features used to represent the individuals and the social ties used to represent our connections may also evolve with respect to sequential, spatial, and other factors. Such a complication is becoming part of the reality for Big Data applications, where the key is to take the complex data relationships along with evolving changes into consideration, to find out valuable patterns from Big Data collections. [3]

VI.CONCLUSION

Real-world applications and key industrial stakeholders and initialized by national funding agencies, managing and mining Big Data have shown to be a challenging yet very compelling task.Term Big Data is accurately concerns about volume of dat, our HACE theorem suggests the key characteristics of the Big Data that are. First is, Large with heterogeneous and diverse data sources,Second is autonomous data sources with distributed and decentralized control, and third one is complex and evolving in data and knowledge associations. Such kind of mutual characteristics suggest that Big Data need a "big mind" to combine data for maximum values. To describe the concept of Big Data, we have review several challenges at the data, model, and system levels. To support Big Data mining, high-performance powerful computing platforms are required, which impose regular designs to unleash the full power of the Big Data. At the data level the autonomous information sources and the range of the data collection environments, often gives result in data with complex situation, such as missing values. In certain situations, solitude concerns, noise, and errors can be introduced into the data, to create tainted data copies. Developing of a secure and sound information sharing protocol is a major challenge. At the model level, the key challenge is to create universal models by combining nearby discovered patterns to form a unifying view. This requires carefully designed algorithms to analyze model correlations between scattered sites, and combine decisions from several sources to gain a best model out of the Big Data. At the system level, the important challenge is that a Big Data mining framework needs to consider complex relations between models, samples and data sources, along with their evolving changes with time and other likely factors. A system needs to be watchfully designed so that formless data can be linked through their compound associations to form helpful patterns, and the growth of data volumes and item relationships should help form rightful patterns to estimate the tendency and view. We stare Big Data as an capable style and the necessity for Big Data mining is arising in all science and engineering fields. With the use of Big Data technologies we will with any luck be able to give the best part of applicable and most precise social sensing feedback to improved realize our society at realtime. We can additionally stimulate the association of the public audiences in the data building loop for community and economical events. The period of Big Data has arrived.

REFERENCES

[1] Apache HBase http://hbase.apache.org[2] Apache Accumulo http://accumulo.apache.org[3] Xindong Wu, Fellow, IEEE, Xingquan Zhu, Senior Member, IEEE,Gong-Qing Wu, and Wei Ding, IEEE"Data mining with Big Data" Transaction on knowledge and data Engineering vol. 26 No.1 January 2014.[4] J. Kepner and S. Ahalt, "MatlabMPI," Journal of Parallel and Distributed Computing, vol. 64, issue 8, August, 2004.[5]B. Hindman, A. Konwinski, M. Zaharia, A. Ghodsi, A.D. Joseph, R.Katz, S. Shenker and I. Stoica, "Mesos: A Platform for Fine-Grained[6] N. Bliss, R. Bond, H. Kim, A. Reuther, and J. Kepner, "Interactive grid computing at Lincoln Laboratory," Lincoln Laboratory Journal, vol. 16,no. 1, 2006.[7] J. Kepner et al., "Dynamic distributed dimensional data model (D4M) database and computation system," 37th IEEE International1989.
[8] A. Jacobs, "The Pathologies of Big Data," Comm. ACM, vol. 52, no. 8, pp. 36-44, 2009.
[9] A. Jacobs, "The Pathologies of Big Data," Comm. ACM, vol. 52, no. 8, pp. 36-44,2009.
[10] R. Ahmed and G. Karypis, "Algorithms for Mining the Evolution of Conserved Relational States in Dynamic Networks," Knowledge and Information Systems, vol. 33, no. 3, pp. 603-630, Dec. 2012.
[11] M.H. Alam, J.W. Ha, and S.K. Lee, "Novel Approaches to Crawling Important Pages Early," Knowledge and Information Systems, vol. 33, no. 3, pp 707-734, Dec. 2012
[12] S. Aral and D. Walker, "Identifying Influential and Susceptible Members of Social Networks," Science, vol. 337, pp. 337-341, 2012.

CAMPUS DIRECTOR
International Centre of
Excellence In Engg. & MGMT.
Aurangabad

CAMPUS DIR...
Excel... In... & ...NT.
Aurangabad

|| Volume 2 || Issue 5 || November 2017||ISSN (Online) 2456-0774
INTERNATIONAL JOURNAL OF ADVANCE SCIENTIFIC RESEARCH
AND ENGINEERING TRENDS

Privacy Preserving Data Publishing Using Slicing

Dinesh D. Jagtap, Prof. R. A. Auti
Student, CSE Department, Everest Educational Society's Group Of Institutions Aurangabad ,Maharashtra, India[1]
Asst. Professor, CSE Department, Everest Educational Society's Group of Institutions Aurangabad Maharashtra, India[2]
dinesh.jagtapd@gmail.com

Abstract— **There are number of techniques are researched and proposed for anonymization of data for privacy preserving data mining Data anonymization techniques for privacy – preserving data publishing has got the more importance in recent year. In any large organization such as hospitals, government agency, private firms there is a huge micro data is generated on daily basis so, maintaining the privacy and identify of any entity such like as patient, bank customer is having the highest priority. Anonymity is the way which the can hide anyone's identity or make it concealed. Most common and well known generalization & bucketisation are the two techniques which are design for privacy preserving micro data publishing. The previous study of generalization shows that it loses subsequent amount of data while performing generalization process. Bucketization, on the other hand not applicable for preventing member ship disclosure. In this paper, we proposed an innovative technique called slicing, using this we can partitions the data both horizontally as well as vertically. In this paper we also shown that this technique preserves improved data utility than generalization technique and it also support membership disclosure protection.**

Keywords:- *Privacy Preserving, Membership Disclosure, Data Anonymization, Generalization, Bucketization,* Tuples, sensitive attributes (SAS)

I INTRODUCTION

Publishing of detail data by maintaining privacy has been studied widely in last couple of years. Data anonymization is a technique that converts a clear text into a non-human readable form. Most popular Privacy Preserving techniques are generalization & bucketization. In both methods attribute values are mainly classified into three categories.(i) Identifier by which we can uniquely identify any individual such as name or social security number.(ii) Some may be Quasi- identifier (QI) such as zip-code, age and gender whose values when combined taken together can possibly identify and individual.(iii) Some attributes are subtle or sensitive attributes, which are unknown to the opponent and considered as sensitive, like Disease and salary. In both methods first identifiers are truncated from the data and then tuples will be partitioned into buckets.

II EXISTING SYSTEM

From recent studies it is observe that first generalization losses substantial amount of data particularly for high dimensional data. Generalization based privacy preserving is not suitable for high-dimensional data. To make generalization more operative records belongs to the same bucket must be closed to each other so that performing generalization of records does not result into any kind of information loss. With comparison to bucketization, it provides efficient data utilization over generalization, but it is also having several limitations. The reason is bucketization issues the I values as it is in their actual forms so an challenger can easily discovers whether an individual is present in published data or not.

III PROPOSED SYSTEM

Here we are presenting a novel data anonymity technique called overlapping slicing. This technique divides the provided data set horizontally and vertically. In Vertical partitioning mechanism by attributes are grouped into column on the basis of attributes correlation among other attributes. And horizontally partitioning is performed by tuples grouping into bucket. And finally within each bucket each column values are randomly. Her column linking with different column is breaked but their association between each column is maintained. The algorithm of privacy preserving mainly contains the three steps attribute portioning column generalization and tuple partitioning

IV VARIOUS ANONYMIZATION TECHNIQUES

A. Generalization

Performing anonymization using generalization is one of the basic idea. In this approach the QI values are put in place of values that are less specific but semantically constant. So, all QI identifier values from one group are generalized to the whole cluster extent in the QID space. If minimum two relations in one group contains dissimilar values in a other column (i.e. one is having item and the other don't), in this case all information related to that item will get lost.. All possible items from the log are grouped into the QID. Quasi-identifier may be present in large number, And possible items may be in the order of thousands so making generalization of that would acquire tremendously high information loss, And makes the data unusable. To make the generalization more efficient, records from the same bucket must be close to each other so making generalization would not result into data loss. In the case of

multi- dimensional data, most data points are having the same distance between each other. To initiate data analysis process on the table after generalization, the data analyst should assume that every value in generalized set is equally probable. So he has to make uniform distribution of all the values.

B. Bucketization

In this technique tuple T is partitioned into bucket and then non – sensitive values are get separated from sensitive attribute values within each bucket. Then the resultant data consists of the permuted sensitive values. In simple way let us set the partitions of bucketization more formally. First of tuples are get partitioned into buckets (partition the T horizontally), and inside every bucket random variation are applied to the column which contain s-values. After this generated set of buckets B is published. It is also having some limitations .Membership disclosure is not prevented in bucketization. Because of here QI values are published in their original forms so it is easy for adversary to find out whether published data contains record for certain user or not. Second it requires clear segregation between sensitive values and quassi identifiers.

C. Slicing

Slicing is a newly invented data anonymization technique. This technique divides the provided data set horizontally and vertically. In Vertical partitioning mechanism by attributes are grouped into column on the basis of attributes correlation among other attributes. And horizontally partitioning is performed by tuples grouping into bucket. And finally within each bucket each column values are randomly. here column linking with different column is braked but their association between each column is maintained. The algorithm of privacy preserving mainly contains the three steps attribute portioning column generalization and tuple partitioning. Slicing results into an efficient data utility because grouping of highly correlated data attributes is done and correlations between such attributes is maintained. So, slicing algorithm mainly classified into three steps.

1. Partitioning of attribute values.
2. Generalization of column.
3. Partitioned the tuples.

V SYSTEM ARCHITECTURE

A. Data Collection And Data Publishing

Figure 1 Describes the typical scenario for of collecting and publishing the data, the data holder obtains data from record owners i.e. Alice and Bob as shown in figure. Then in the data publishing phase, the collected data is forwarded to a data beneficiary, who will further perform data mining operation on the published data.

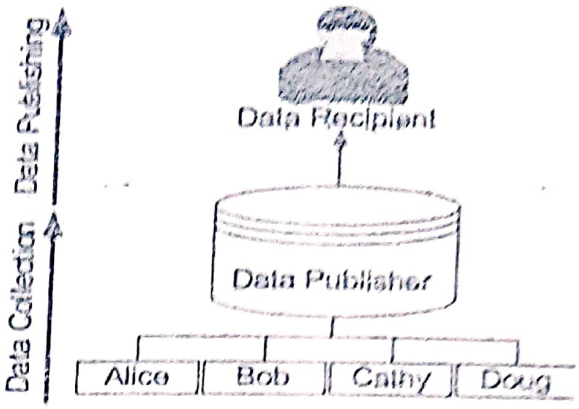

Figure 1 Collecting & Publishing data.

B Privacy-Preserving Data Publishing:

In this process of Privacy-Preserving Data Publishing (PPDP), the data can contains the set of attributes categorized into Clear Identifier, Quasi Identifier, Sensitive Attributes and non-Sensitive Attributes. Where Explicit Identifier is having attribute set, which individually identify any person such as name or social security number. When the values of quassi-identifier is taken into consideration combinely it, can potentially identify and individual. Some attributes are may be sensitive attributes (SAS), which are anonymous to the adversary and considered, sensitive, such as Disease and salary. And Non-Sensitive Attributes contains rest of the attributes which does not belongs to above mentioned three categories. These four set of attributes are considered as disjointly. Maximum studies adopt that every record in the table signifies a dissimilar record owner.

C Data Anonymization

In this phase clear text is converted into human non-readable form. Data anonymization technique for PPDP is gaining a lot much interest from last few years. Micro-data contains information about a household, a person, or an organization. Most popular Privacy Preserving techniques are generalization & bucketization. In both methods attributes are classified into three categories.(i) Identifier which can uniquely identify and individual such as name or social security number.(ii) Some may be Quasi- identifier (QI) such as zip code age and sex whose values when taken together can potentially identify and individual.(iii) Some attributes are sensitive attributes (SAS), which are unknown to the adversary and considered, sensitive, such as Disease and salary. In both methods first. Of all identifiers are removed from the data. The benefit of data anonymization is that using it we can transfer the information between two parties like two persons, departments, agencies, by overcoming threat factor of unintended disclosure,

and in confident environments that empowers estimation and analyses post-anonymization.

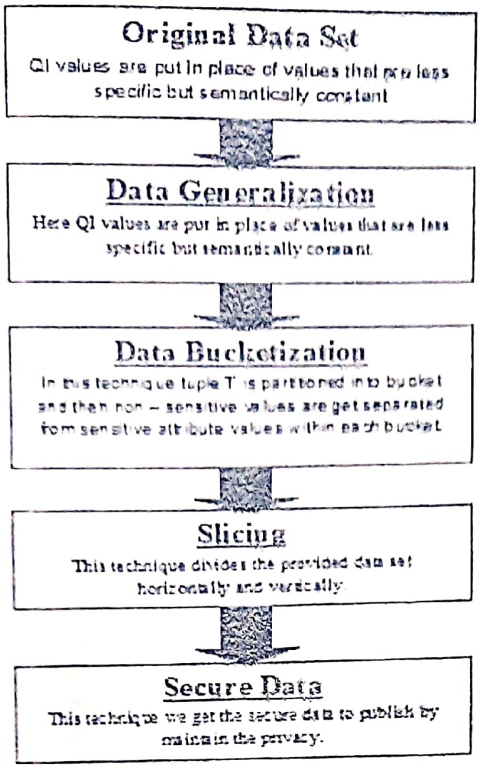

Figure 2: Proposed System Architecture

VI ALGORITHMS USED

Bucketization, generalization and such another algorithms having main focus on maintaining privacy, however they cannot prevent the attribute disclosure. So to solve this problem slicing algorithm is used. Slicing algorithm is mainly having three steps:

1. Attribute partitioning
2. Column generalization
3. Tuple partitioning

We will now describe these three phases.

A. Attribute Partitioning

As its name suggests, this algorithm makes partitions of attributes so, attributes which are highly correlated are placed into the similar column. To maintain utility and privacy this mechanism is somewhat good in nature. To achieve data utility, highly interrelated attributes are assembled and the associations among those attributes are maintained. In terms of privacy, the association of uncorrelated attributes is having higher risk of identification over the highly correlated attributes association. Because uncorrelated attribute associations is frequent and thus more specific to recognize.

B. Column Generalization

Generalization of column is not an essential phase, it can be helpful in some aspects. To maintain the membership disclosure protection column generalization is required. If the value of column value is unique in a column, then a tuple having this unique column value can only belongs to one matching bucket. In terms of maintaining privacy protection as in the case of generalization and bucketization where each tuple can only have single matching buckets. The main trouble is that the values which are unique in column can be identify. It would be beneficial to make column generalization mechanism to make sure that every columns value seems with at least some frequency. Second, to accomplish the equal privacy against attribute disclosure after performing column generalization, the size of bucket can be smaller. While column generalization may result in data loss, slighter bucket-sizes permit efficient utility of data. Therefore, the mechanism of tuple partitioning and column generalization is the alternative for each other.

C. Tuple Partitioning

Slicing algorithm is having two set of data structures: 1) a queue containing the group of buckets Q 2) and second set contains the sliced buckets SB. Initially, Q contains a buckets having all the tuples and SB is empty. For every repetition, the algorithm eliminates a bucket from the set Q and again divides the bucket into two more buckets. If the l-diversity is satisfied by sliced table, then the algorithm places the two buckets at the last of queue Q. Otherwise, bucket cannot be split anymore and the algorithm puts the bucket into SB. When group of bucket Q becomes empty, then we got the set of computed sliced table. SB is the group sliced buckets.

VII FUTURE SCOPE AND CONCLUSION

The limitations of generalization and bucketization are somewhat reduced by slicing approach. And it also maintains efficient data utility while protecting against privacy related issues. Attribute disclosure and membership disclosure is prevented by slicing mechanism. Slicing gives better result in terms of efficient data utility over generalization and is also more advanced and powerful than bucketization in workloads having the set of sensitive attribute. We assume slicing where every attribute is in just one column. An extension is the notion of overlapping slicing, which duplicates an attribute in more than one column. The random grouping is not the perfect solution to maintain privacy. We proposed to design more operative algorithm for tuple grouping. Additional way is to plan a data mining tasks using the anonymzed data generated by various Anonymizations methods. Privacy is preserve in slicing by breaking the association between attributes which are highly correlated. And reserve data utility by maintaining the highly correlated attribute association. Another important benefit of slicing is it can handle large and high-dimensional amount of data. The simple idea suggested by this work is first of all

consider the characteristic of data before performing annonymization on it, after analysis of characteristics performs anonymization strategy.

REFERENCES

[1]Tiancheng Li, Ninghui Li, Senior Member, IEEE, Jin Zhang, Member, IEEE, and Ian Molloy "Slicing: A New Approach for Privacy Preserving Data Publishing"Proc. IEEE Transactions On Knowledge And Data Engineering,Vol. 24, No. 3, March 2012.

[2]C.Aggarwal. On k-anonymity and the curse of dimensionality. In VLDB, pages 901–909, 2005.

[3]Asuncion and D. Newman. UCI machine learning repository, 2007.

[4]Blum, C. Dwork, F. McSherry, and K. Nissim. Practical privacy: the sulq framework. In PODS, pages 128–138, 2005.

[5]J.Brickell and V. Shmatikov. The cost of privacy: destruction of data-mining utility in anonymized data publishing. In KDD, pages 70–78, 2008.

[6]B.-C. Chen, R. Ramakrishnan, and K. LeFevre. Privacy skyline: Privacy with multidimensional adversarial knowledge. In VLDB, pages 770–781, 2007.

[7]H. Cramt'er. Mathematical Methods of Statistics. Princeton, 1948.

[8]Dinur and K. Nissim. Revealing information while preserving privacy. In PODS, pages 202–210,2003.

[9] A. Machanavajjhala, D. Kifer, J. Gehrke,and M. Venkitasubramaniam. "L-diversity: Privacy beyond k-anonymity", InICDE 2006.

CAMPUS DIRECTOR
International Centre of
Excellence In Engg. & MGMT.
Aurangabad

|| Volume 2 ||Issue 12 || JULY 2017 ||ISSN (Online) 2456-0774
INTERNATIONAL JOURNAL OF ADVANCE SCIENTIFIC RESEARCH
AND ENGINEERING TRENDS

Recommender System for Web Services Employing QoS Values and Physical Location

Dinesh D. Jagtap[1] , Prof. R. A. Auti[2]
CSE Department, Everest Educational Society's Group of Institutions, Aurangabad, Maharashtra, India[1]
Asst. Professor, CSE Department, Everest Educational Society's Group of Institutions, Aurangabad, Maharashtra, India[2]

Abstract— Recommendation of Web service is the popular area of research in the field of IT. Collaborative filtering (CF) is one of the popular method of web service recommendation which is based on a Quality of Service, (QoS) parameters of the service. Web services are nothing but a software component which are designed to perform machine to machine communication over the network. The QoS of web service is essential factor which is taken into consideration while selecting the appropriate web service. Previously a number of studies or methods were conducted for choosing web services and making their recommendation by performing collaborative filtering; here we are going to examine these methods with their advantages and limitations. And also based on this study we are proposing a new technique for web service recommendation which is based on past experience of user regarding QOS of web service and their location.

Keywords: *- Web Service, Service Computing, Collaborative filtering, QoS values, Web service recommendation, QoS prediction, collaborative filtering, privacy preservation.*

I INTRODUCTION

Web services are software components to support interoperable machine-to-machine interaction over a network. The increasing acceptance of web services in large organization demands efficient recommendation and selection techniques for optimum web service amongst the variety of available services on internet. Web services have been extensively employed by both individual developers and enterprises for building service-oriented applications. While considering the QoS properties of Web services, some features of web services are user independent and having equal values for different users (e. g Availability, price, popularity etc.)The values of the user independency of QoS properties are typically offered by service providers or third-party registers (for example, UDDI). In another case some QoS features for users are reliant and have dissimilar values for different users (for example, Invocation failure rate, response time, etc.).

Client-side Web service evaluation requires real web service calls and encounters the following drawbacks: First, real Web service invocations enforce costs for service users and utilize the resources of the service provider.

It can exist on many Web service candidate analyzed and some suitable web services in the assessment list may not be detected and observed by the service user.

Finally, in web service evaluation not all the users are expert.

However, without adequate client-side evaluation, accurate values of the user-specific QoS properties cannot be obtained. Hence optimal Web service selection and recommendation are not easy to accomplish.

II RECOMMENDER SYSTEM

Large organization and individual user requires a particular system which can understand the interests of user and recommend them the best utilizable services. So according to their functionality recommender systems can be classified as content based filtering, collaborative filtering, Hybrid models[2].Recommender systems can help users by selecting most valuable items by calculating the similarities among other users with the help of collaborative filtering algorithms.

2.1 Collaborative Filtering Methods

Collaborative filtering is nothing but an identification of similar users, related web services and recommend them. Collaborative filtering is generally employed in commercial recommender systems like as Amazon.com and Netflix. The Web services suggestion for the user is based on the prior knowledge of web service history. Therefore, and accurate Web service QoS prediction is required for service user providers. On the basis of predicted QoS values the preferred service selection can be completed. Collaborative Filtering [3] was firstly proposed by Rich and has been extensively used in service recommendation systems.

Collaborative Filtering algorithm uses two processes:

Prediction process[3][4] where a numerical value expressing the predicted probability of web services that cannot be upheld certain users. This predicted value is in the same scale as opinion by the same user supplied values.

Recommendation process [3] where a list of N items that the active users like the most is recommended. This recommended list has those users who do not already have access to Web services. This interface of collaborative filtering algorithm Top N recommendation [13] is called Collaborative filtering process and is as shown in the following figure 1.

It is not practical for each user to dynamically determine QoS values due to the exclusive overhead of invoking a huge number of services.

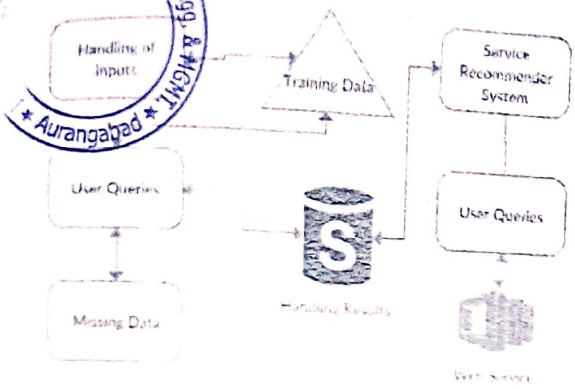

Figure 1: Web service recommendation process

To address this issue, collaborative QoS prediction has recently been proposed, and becomes a key step to QoS-based Web service recommendation [3], [4], [5]. Specifically, two types of CF approaches have been studied for QoS prediction of Web services [5] in recent literature. There are two types of collaborative filtering algorithms:

2.1.1 Model-Based Collaborative Filtering

On the basis of ratings of dataset a models is built in this technique. In other words, we take out some useful information from the dataset, and consider that model to make recommendations [5] without using the whole dataset every time. With the help of this approach ones can potentially offers advantage of both speed and scalability. With the help of model-based algorithms we can learn and understand the collection of QoS, a model which is then used for QoS predictions. Model-based CF algorithms consist of Bayesian models (probabilistic) and clustering models [6]. Model-based CF technique [6] deliver a predefined model by studying previous Qos parameters, and then the trained model can be used to predict the unknown QoS values. To address the QoS prediction problem matrix factorization [7] is one of the most popular model-based CF approaches that were first introduced. Matrix factorization model [7] treats the problem well sparsely and generally achieved better performance than neighborhood-based approaches. Typical examples include user-based approaches (e.g., UPCC [8]) that leverage the QoS information of similar users for prediction.

2.1.2 Memory Based Collaborative Filtering

In memory-based algorithms approach the collaborative filtering is perform by considering the complete database. As described by Breese et. al [9], It finds the users those who are similar to active user (i.e. the users we want to make predictions for) and it uses their preferences to forecast ratings for the current user. For making predictions memory-based algorithms uses the data (users, services and QoS data)stored in memory. They can be categorized in to nearest neighbor algorithms and top-N recommendation algorithms. This type of model for CF approaches use the experiential QoS data to compute the similarity values between users or services and use them further for QoS recommendation. Top-N recommendation is to recommend a number of N top Web services; this will be to a specific user of interest. Analyze Top N recommendation [10] techniques to correlate the user service matrix dissimilar users or services and use them to calculate the recommendations.

III RELATED WORK

3.1. QoS aware Web service recommendation

Due to the vast availability of web services on internet users firstly pays their attention to QoS instead of functionality than before. Web service QoS mostly contains non-functional attributes such as availability, security, response time and throughput, etc. It has been widely used in web service selection [11], [12] (Wang, Wang et al. 2013), service composition (Feng, Ngan et al. 2013), service recommendation (Cao, Wu et al. 2013; Jiang, Liu et al. 2011) and other popular topics in the field of Services Computing. In this section, we present the related work of efficient QoS-aware Web service [12] recommendation.

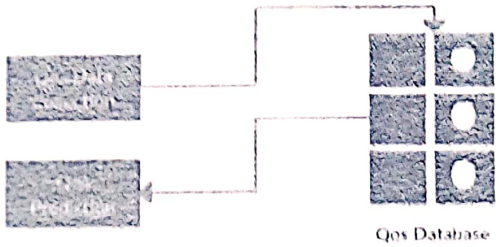

Figure 2: Web Service QoS Prediction

By studying user-based combination and item-based collaborative filtering method Web service QoS value prediction can be done. Their approach does not require Web service calls and help by analyzing QoS information of similar users Service users discover appropriate Web services. In its Web service [12] evaluations in paper reports, to reduce the effect of Web service calls to the real web services, they selected only one operation from a web service make for evaluations and use the power of this operation to the performance by presenting the Web service.

3.2. Web service recommendation based on location aware Qos:

Existing approaches fail because of QoS discrepancy according location of user to consider; and former recommender systems are all black boxes offers only partial information about the performance of the service candidates. Thus X. Chen, Z. Zheng, X. Liu, Z. Huang, H. and Sun [13], [13] proposed designed a novel collaborative filtering algorithm for large-scale Web service recommendation on location aware QoS. It firstly combines the memory-based and model-based and CF algorithms for Web service recommendation, So this

method clearly showing the recommendation accurgieness and time complexity improved as compared to previous service recommendation algorithms. Second, they create a visually appealing interface to browse the recommended web services, which allows a better understanding of the service performance. This algorithm uses the property of QoS of users in distinct regions clustering. Based on the feature region a refined nearest neighbor algorithm is proposed to generate QoS forecasts.

The concluding service recommendations are on a map by putting the underlying structure of QoS space to show and help users who accept recommendations. Similarly, they also change existing service similarity measurement of collaborative filtering which is used by service location information based on a hybrid collaborative filtering technology. Missing QoS values are find after finding similar users and services.

3.3. Web Service Recommendation Methods Based on Personalized Collaborative Filtering

There were number of techniques are available for selecting web services and recommendation on the basis of collaborative filtering, but not often do they take into account personal influence of users and services. Therefare Y. Jiang, J. Liu, Tang, X. Liu [14] proposed a technique of collaborative personalized recommendation effective filtering for Web service. A major piece of this method is the computation of the measure of the similarity of web services. Unlike the Pearson correlation coefficient (PCC) similarity measure, they consider the individual impact of services where among users and the individual impact of the measure of calculated likeness of services. On the basis of similarity measure of web services, they build up a custom hybrid efficient collaborative filtering technology (HICP) for integrating algorithm based on custom user and custom algorithm based item. Similarly, L. Shao, J. Zhang, Y. Wei, J. Znao, B. Xie and Mei H. being aware of different experiences of consumers QoS, they strike a collaborative approach to filtering based on mining matching decision and forecasting of consumer experiences.

IV FRAMEWORK OF QOS-AWARE WEB SERVICE RECOMMENDATION

The basic proposal of this scheme is that, the users those who are closely located with each other are more likely to have similar service experience than those who live far away from each other. Here we are employing the idea of user-collaboration in our web service recommender system.Our recommendation technique is designed as a two-phase process which is based on the collected QoS records. In the first step user are grouped according to their physical location and previous experience of web service QoS. In the second step, we search for similar users for the active user and make QoS prediction for the unused services. And finally service with best predicted QoS will be recommended to active user.

4.1. Location Information Representation, Acquisition and Processing

In this section we are going to discuss how to represent, acquire, and process location information of service users and web service which leaves an essential base for employing location-aware Web service recommendation method.

4.1.1. Location Representation:

User's location can be represented as a [IP Address], [Country], [IP No.], [AS], [Latitude], [Longitude]. Normally, a country has many ASs and an AS is within one country only. Ass is connected with each other and internet is composed of different number off Ass.

However, it is not always true that users located in the same AS are always geographically close, and vice versa. Therefore, it possible that even if two users are located in the same city, they may seem to be at different ASs. So this is the main reason behind choosing AS instead of other geographic positions instead of latitude and longitude for representing user's location.

4.1.2. Location Information:

Within the phase of acquisition we can fetch the location information of both Web services and service users. By using user's IP address it is possible to obtain his full location information, the things which we only needs are only to identify both the AS and the country in which he is located based on IP address. There are number of services and databases are presented for this purpose (e.g. the Who is lookup service2). In this work, we accomplished the IP to AS mapping and IP to country mapping using the GeoLite Autonompus System Number.

4.1.3. Similarity Computation and Similar Neighbor Selection

Here we have defined notations for the convenience of describing our method and algorithms. For computing similarity between both users and Web services we implemented weighted PCC algorithm, which takes personal QoS characteristics into consideration. Finally, author has discussed incorporating locations of both users and Web services into the similar neighbor selection.

4.1.4. Similar Neighbor Selection:

This section is a very important step of CF. In conventional type of user-based CF, the Top-N similar neighbor selection algorithm is used always . It selects the first N users as neighbors those who are most similar to the active user. Similarly, the Top-N similar neighbor selection algorithm can be useful to select top N Web services that are most similar to the web service for which we want to find recommendation. Old-fashioned Top-N algorithms ignore this problem and still

choose the top N most ones. Because of the resulting neighbors are not actually similar to the target user (service), doing this will degrade the prediction accuracy. Therefore, leaving those neighbors from the top N similar neighbor set is better if the similarity is not greater than zero. Secondly, as previously mentioned, Web service users may happen to notice similar QoS parameters on a few Web services.By taking location-relatedness of Web service QoS into consideration [5], authors have combined the user's location and Web services into similar neighbor selection.

4.2. User-Based QoS Value Prediction:

Authors offered a user-based location-aware CF method, named as ULACF. Traditional user-based CF methods usually adopted for finding value predictions. This equation, however, may be incorrect for Web service QoS value prediction. Web service QoS factors such as response time and throughput are the objective parameters and their values are not constant. Therefore, predicting QoS values based on the average QoS values perceived by the active user is not sufficient which gives faulty result. Intuitively, given two users that have the same estimated similarity degree to the target user, the user who is nearer to the target user should be placed more confidence in QoS prediction than the other.

V CONCLUSION

The association of the various QoS properties is important for the achievement of web service recommendation. Due to the increasing demand of Web services and the latency of dynamic service selection and integration, some service providers now provide parallel services. QoS is one of the modified factor to differentiate functionally similar Web services. The basic idea behind this work is to predict web service QoS values and recommend the best web service to active user best on past QoS records of web service. In this work we combine prediction results generated from user region and service region which gives better results than existing techniques. We also noticed that combination result is much better than the result from either one method of prediction from user region or the one generated from service region. Our future work includes the correlation between different QoS properties and detecting the users those who are contain inaccurate QoS information.

REFERENCES

[1] Jianxun Liu, Mingdong Tang, Member, IEEE, Zibin Zheng, Member, IEEE, Xiaoqing (Frank) Liu, Member, IEEE, Saixia Lyu, "Location-Aware and Personalized Collaborative Filtering for Web Service Recommendation", IEEE Transactions on services computing, manuscript Val X , No XX, 2015

[2] W. Zhang, H. Sun, X. Liu, and X. Guo, "Temporal QoS-aware web service recommendation via non-negative tensor factorization," in Proc. of the 23rd International World Wide Web Conference (WWW), 2014, pp. 585–596.

[3] X. Chen, Z. Zheng, Q. Yu, and M. R. Lyu, "Web service recommendation via exploiting location and QoS information," IEEE Trans. Parallel Distrib. Syst., vol. 25, no. 7, pp. 1913–1924, 2014.

[4] X. Chen, Z. Zheng, X. Liu, Z. Huang, and H. Sun, "Personalized QoS-Aware Web Service Recommendation and Visualization," IEEE Trans. Serv. Comput., vol. 6, no. 1, pp. 35-47,1st Quart., 2013.

[5] R. Salakhutdinov and A. Mnih, "Probabilistic matrix Factorization," in Proc. of the 21st Annual Conference on Neural Information Processing Systems (NIPS), 2007.

[6] M. Tang, Y. Jiang, J. Liu, X. F. Liu: Location-Aware Collaborative Filtering for QoS-Based Service Recommendation. in Proc. 10th International Conference on Web Services, Hawaii, USA, June 2012, pp.202-209

[7] S. Wang, Q. Sun, and F. Yang. "Towards Web Service Selection Based on QoS Estimation," Int'l J. Web Grid Serv., vol. 6, no. 4, pp. 424-443, Nov. 2012.

[8] G. Kang, J. Liu, M. Tang, X. F. Liu, K. K. Fletcher. "Web Service Selection for Resolving Conflicting Service Requests". Proceedings of International Conference on Web Services. IEEE Computer Society, pp. 387-394. 2011.

[9] Breese, J. S., Heckerman, D., Kadie, C. (1998). Empirical analysis of predictive algorithms for collaborative filtering. Proceedings of the 14 th Conference on Uncertainty in Artificial Intelligence (UAI 1998), Madison, Wisconsin, USA, July 24-26, 1998, 43-52.

[10] Y. Zhang, Z. Zheng, M. R. Lyu, "WSExpress: a QoS-aware search engine for Web services", in Proc. 8th IEEE International Conference on Web Services, Miami, FL, USA, July, 2010, pp.83-90

[11] M. Alrifai, and T. Risse, "Combining Global Optimization with Local Selection for Efficient QoS-aware Service Composition," in Proc. of the International World Wide Web Conference, Apr. 2009, pp. 881- 890.

[12] X. Su and T. M. Khoshgoftaar, "A survey of collaborative filtering techniques," Adv. Artificial Intelligence, vol.2009, 2009.

[13] Z. Zheng, H. Ma, M.R. Lyu, and I. King, "WSRec:A Collaborative Filtering Base Web Service Recommendation System," in Proc. 7th

[14] International. Conference on Web Services, Los Angeles, CA, USA, pp.437- 444, 2009.

NON-TEACHING STAFF PERFORMANCE APPRAISAL FORM

(Assessment Year 2018-19)

Instructions:
- *The details shall be provided for the academic year only*
- *Tick (☐) wherever applicable*

1. General Information:

Sr. No.	Particulars	Information			
1.	Full Name	Bhaskar Raghunath Akhand			
2.	Date of Birth & Age(in Years):	16/06/1969	Gender	Male ✓	Female
3.	Name of Department:	Store In-charge			
4.	Name of Constituent Institute:	ICEEM			
5.	Date of Joining the present Institute:	7/5/2011			
5.	Number of year/s of Service in ICEEM	13yrs.			
6.	Designation and Date of last Promotion, if any:	Store In-Charge			
7.	Current Designation:	Store In-Charge			
8.	Mobile No.	8788686819			
9.	E-mail Id:	a.bhaskarjiceem@gmail.com			

2. Additional Educational Qualification / Advanced skill learnt during the Assessment year:

Name of the program	Name of the Board / University	Date of Completion	Specialization / Major Subject
MSW.	BAMU Aurangabad.	2013.	Social work

The information provided in the above tables is true to the best of my knowledge.

Date Signature of Employee: 24/05/19

Date and Signature of DIRECTOR: 24/05/19

CAMPUS DIRECTOR
International Centre of
Excellence In Engg. & MGMT.
Aurangabad

Remarks of HOD/DIRECTOR

Sr. No.	Attributes	Poor	Fair	Good	Excellent
1	Discipline, Punctuality, Regularity, Sincerity			✓	
2	Moral Values - Honesty, Integrity, Behavior, Respectful etc.			✓	
3	Etiquettes and Manners			✓	
4	Communication skills			✓	
5	Inter-personal relationship with the fellow colleagues			✓	
6	Knowledge of System, Procedures, Rules & Regulations			✓	
7	Desire for work		✓	✓	
8	Ability to learn and grasp			✓	
9	Flexibility / Adaptability towards the work assigned			✓	
10	Amenable to change				
11	Accuracy and Quality of output			✓	
12	Timely accomplishment			✓	
13	Multi-tasking			✓	
14	Achieving the goal			✓	
15	Ability to take responsibilities			✓	
16	Potential to grow in present job			✓	

Specific Remarks, if any:

good

3. Recommendation (Yes / No)

For Increment	Yes
For additional Increment	No
For Incentive	No
For No Increment	—

Date and Signature of HOD: 24/10/19

CAMPUS DIRECTOR
International Centre of
Excellence In Engg. & MGMT.
Aurangabad

Date and Signature of DIRECTOR:

IIRW'S

INTERNATIONAL CENTRE OF EXCELLENCE IN ENGINEERING AND MANAGEMENT (ICEEM)

NAAC ACCREDITED

Summary Sheet

For

Non-Teaching staff Performance Appraisal Form(Assessment Year_____)

Sr. No.	Name of the Staff	Designation	Name of the Department / Section	Recommendation (Yes / No)				Specific Remarks, if any
				Increment	Additional Increment	Incentive	No Increment	
	A.R. Abhang	Store Incharge	Store	Yes	-	-	-	Nil

Date, Signature & Stamp of HOD/DIRECTOR.: 24/07/19

Note:

1. The summary sheet shall be filled by the Department / Section Head.
2. Please submit the hard copy of summary sheet in a sealed envelope to the Admin Office , **ICEEM; mentioning as 'Confidential' on the envelope.**

For Office use only

Remarks of HOD: 24/07/19

Remarks of Director:

CAMPUS DIRECTOR
International Centre of
Excellence In Engg. & MGMT.
Aurangabad

Add.: Gut No.4, Opp. Bajaj Auto Ltd. Main Gate, Aurangabad- Pune National Highway, Aurangabad - 431136 (MS) India.
Telephone : 0240 - 2558101 to 10 | Telefax 0240 - 2558111
Website : www.iceemabad.com | E-mail : director@iceemabad.com