CRITERIA III: RESEARCH, INNOVATION AND EXTENSION

3.2.1: NUMBER OF PAPERS PUBLISHED PER TEACHER IN THE JOURNALS NOTIFIED ON UGC WEBSITE

Data Collection Year for Assessment

2019-20

A materialistic perspective of consumer decision-making styles

Sartaj Chaudhary
Jagannath International Management School, New Delhi, India, and

Ajoy Kumar Dey
Birla Institute of Management Technology, Noida, India

Abstract

Purpose – Materialism has become a topic of increasing interest to researchers and policymakers because it can influence consumer behavior. However, a clear picture of how this phenomenon impacts consumers has proven to be elusive. Using an integrated framework, this paper aims to derive hypotheses from theoretical concepts of materialism and consumer decision-making styles and uses a survey of 1,216 respondents in India to test the hypotheses.

Design/methodology/approach – The present study is the first of its kind to test the impact of materialist values on consumer decision-making styles among a sample of 13-18 years old school children. The constructs are validated through a first- and second-order confirmatory factor analysis and an integrated second-order structural model is developed.

Findings – This study finds that materialism is a positive predictor of "recreation/ hedonistic",; "confused by over-choice"; "brand consciousness"; "perfectionistic high-quality consciousness"; and "habitual brand loyal" style of consumers. Further, materialism has a negative impact on "price value consciousness" of consumers. These findings have important implications for theory and practice.

Research limitations/implications – This study is restricted to school children in the National Capital Region and hence cannot be generalized to the whole young population in India.

Originality/value – Conceptualizing both materialism and consumer decision-making styles as second-order constructs, this is a maiden study that examines the impact of materialistic values on the consumer decision-making styles of young consumers.

Keywords India, Structural equation modelling, Materialism, Teenagers, Consumer decision-making styles

Paper type Research paper

Introduction

Materialism has been conceptualized as a consumer value which is central to the lives of consumers. The ownership and acquisition of possessions have been ascribed to achieving major life goals like happiness, satisfaction and success (Richins, 2004). Materialism has a negative impact on consumers' subjective well- being (Manolis and Roberts, 2012) and induces the feelings of insecurity (Kasser *et al.*, 2004). It is directly related to buying products that confer status (Goldsmith and Clark, 2012). Of late, researchers around the world have voiced their concern about protecting young consumers from an obsession with possessions (Chan, 2013). Marketers are keen to promote their products to this segment as they influence family purchase and take consumption decisions (Kaur and Medury, 2010). An intriguing fact is that teenagers

Journal of Indian Business
Research
Vol. 12 No. 2, 2020
pp. 231-248
© Emerald Publishing Limited
1755-4195
DOI 10.1108/JIBR-07-2017-0108

Evolution Of Concept Of Intelligence

Suchitra Srivastava

Associate Professor, Department of Management,
Jagannath International Management School, Vasant Kunj, New Delhi, India.

Abstract

This review paper focuses on the different theories of intelligence and their relevance. It analyses the conceptual framework underlying the theories of prominent researchers across the world. It critically examines the different factors that impact intelligence and define intelligent behavior. It sheds the light on emotional intelligence as major determinant of success and well- being and traces its evolution. The paper also highlights the definitions and concepts of intelligence in the twenty first century. It also gives a new direction to the assessment of intelligence.

Key words: intelligence, intelligent behavior, factors, abilities, emotional intelligence

1.Introduction

The word intelligence is derived from the Latin verb 'intellegere' which means understanding. According to the Webster's dictionary, intelligence is the ability to retain knowledge, use reasoning to solve problems or have above average brain power. The Oxford dictionary defines it as "the ability to acquire and apply knowledge and skills." In other words, it refers to the ability of thinking, reasoning, and acquiring and applying knowledge. [28] The present study is a modest attempt to trace the evolution of the concept of intelligence. The researcher strives to highlight the conceptual framework proposed by the different researchers in the past. The key purpose of this paper is to find out the relevance of these frameworks in the present times.

The prime focus is to analyze whether a person is born intelligent, whether intelligence can be acquired, whether it manifests itself in particular situations or whether it is a complex combination of two or more of the above.

2.Literature Review

The concept of intelligence has been a major topic of research and discourse since time immemorial. The ancient thinkers and philosophers had diverse views. Towards the twentieth century, it started gaining prominence as a measurable quantity. The developments in the conceptual framework are discussed in the present paper.

2.1 Ancient- Greek, Latin, Islamic and Indian Philosophy

Firstly, Plato suggests that "intelligence is that which distinguishes the different social classes, and is unevenly endowed by God. So fixed and innate to the individual was this endowment." Also, his student Aristotle had a similar view on the issue of "intelligence". However, he was more egalitarian compared to Plato. Aristotle claimed that people were very similar to each other in terms of their intellects. What made humans different from animals was this intellect . [15] Ancient Greeks described wisdom or intelligence is Phronesis. It dealt with practical action, implying both good judgement and excellence of character and habits, or practical virtue. Phronesis was a much discussed in the ancient Greek philosophy. Because of its practical connotation and orientation, it was often translated as "practical wisdom", and sometimes (more traditionally) as "prudence", from Latin prudentia. According to Thomas McEvilley, the best translation of the concept could be "mindfulness". [40] It is beneficial to examine the concept of "intellect" also from Islamic tradition. Avicenna being one of the most famous Islamic philosophers, theorized about "intellect" and brought the theory of "active intellect" in philosophy.

Ahmed (2011) gives Avicenna's definition of mind in his book. Avicenna postulates that "The mind is a capacity of the soul that prepares (it) for acquiring knowledge". [1] This definition of "mind"

IMPACT OF HR PRACTICES ON EMPLOYEES IN AVIATION INDUSTRY OF INDIA

By

DR. VAISHALI SHARMA, Dean, Faculty of Management Studies, Jagannath University, Jaipur, vaishali.sharma@jagannathuniversity.org

And

MS. PRABHJOT KAUR, Assistant Professor, Jagannath International Management School, Vasant Kunj, New Delhi, affiliated to Guru Gobind Singh Indraprastha University, New Delhi, prabhjot.kaur@jagannath.org, 9818517245

ABSTRACT

As we all know that Indian Airlines Sector is effectively instrumental in the development of India's economic, therefore, developing Human Resources for this industry has become all the more essential to achieve the national objectives. Aviation Industry is like engines of growth of any economy, which can make the economy, go up in the sky. So, with this notion in mind and many other major incidences which the researcher has come across, the researcher has chosen this topic with a great zeal and enthusiasm to see the final results of the research work which has been done neutrally i.e., without any prior intentions or a bend towards particular direction. Let us see what this particular industry holds for us in terms of various HR Practices being adopted by different Indian Airlines and what impact they put on the employees working in these organizations.

Keywords: HR Practices, Work Engagement, Indian Aviation Industry

Introduction

There are various theories and concepts in HR literature to provide a structure to enhance employee work engagement. However, the HR literature has not appropriately addressed how the satisfaction level of employees with HR practices of the organization influence their work engagement level. HR practices are like the medium by which employees' perceptions, attitudes, and behaviors towards the organization are shaped. Moreover, investing in development of employees is believed to facilitate greater obligation from the employee's

A Study on HR Practices in Indian Aviation Sector

Dr. Vaishali Sharma[1] and Prabhjot Kaur[2]

[1]Jagannath University, Jaipur, India
[2]Jagannath International Management School, Vasant Kunj, affiliated to GGSIPU, Delhi, India.

Abstract. The airline service industry be it commercial or non-commercial is highly competitive, safety-driven and technologically advanced. People, employees and customers, not products and machines, must be the arena of an organization's core competence. The implications of these HR Practices are huge and all-pervasive, affecting the organization's design, competence, climate and several daily routine activities. In such a safety-driven, customer-oriented, service-centric environment, the traditional product-centric approach industrial structure would be inappropriate. Human Resource Management specialization is needed now, than ever before, to lead corporate's internal marketing strategies in order to gain customer loyalty. The core area of corporate's strategy focus should be the manner in which the HR department aligns its activities, practices and functions with the organizational strategic development imperatives. Now, in a highly competitive aviation market, where service-innovations are copied so easily, a key strategic variable that no one can replicate is an airline's unique human resources. The aviation sector is probably one of the most volatile sectors among all the industries in the world.

Keywords: HR Practices, Aviation Industry, Work Engagement.

1   Introduction

Indian Aviation Sector is a sector whose way of doing business is changing constantly largely due to outside forces. This industry is not only subject to regular bankruptcies, mergers and acquisitions; it is also subject to uncontrollable factors such as the political and economic situation of the society and its customer base as well. Thus, considering all these aspects, the HR manager of an airline sector has the challenge of staffing for these ever-changing needs. As the airline's needs are in a constant state of fluctuation, the first step the HR manager must take is to establish a system that allows for a need analysis and then staff the human resources, based on these analyzed needs. Within this sector, there are several different levels of positions, from ground staff to pilots, from stewards to maintenance workforce. Hence, the HR manager has to recruit staff successfully only if each of these different requirements is met by creating direct lines of communication with each of these diverse departments in order to meet their specific recruitment requirements and then make the recruitment plan from this gathered information. HR policies and practices are replicable and as they are the main key to unlock the knowledge, skills and abilities of employees which is the main element through which the firm achieves a competitive edge, so there is a clear motive for organizations to cover the content of HR Practices for the management of their employees. In the Aviation sector, there is not only a purpose for management to cover HR Practices, but there is also a pressure from the external environment on internal management of the firm to conform to the best HR practices.

The Indian Aviation Sector, unfortunately, is not in a good position. Tactical and executive operations, particularly in customer oriented and service centric organizations, are based on three significant critical success factors: employee participation, management commitment, and customer orientation. But, these norms are not followed properly in the aviation sector, primarily because of communication problems. Prosperous and well-established airlines are often cited in the aviation literature for strongly shaping their

A Comparative Study on Job Search Behavior of Job Applicant towards Social Recruitment

Seema Wadhawan[1] and Nidhi Gupta[2]
1 Gitarattan International Business School (GIBS)
2 Jagannath International Management School (JIMS)

People and corporate around the globe are exploring social media, recruitment being one such avenue. The growing use of social networking websites facilitates the individuals to explore the broad range of new job opportunities. The main purpose of this paper is to identify the importance of social networking websites in job seeking behavior of applicants. The study analyzed job information and organizational attributes available on these websites to examine the preference of formal, informal social networking website and employer review website. Social Networking websites undertaken for the study were LinkedIn, Glassdoor and Facebook. Questionnaire was designed and primary data was collected from students, fresher's and those who are at entry level position in ITES sector of Delhi NCR. Convenience sampling was applied and data was collected from 360 respondents. The results of the study show that LinkedIn is regarded as prime channel of recruitment for job search by applicants in Delhi NCR today.

Keywords: *Job Seeking Behavior, LinkedIn, Recruitment, Social Networking Websites.*

INTRODUCTION

The growing population and increasing unemployment in India is a biggest challenge for the job seeking section of society today. With technological advancement and new innovations, adoption of social networking website for recruitment to meet the need of talented resource is increasing. The focus of organization has moved to social networking websites from other sources since they do not facilitate to attract semi passive and passive potential candidates (Dutta, 2014; Singh & Sharma, 2017). Also, accessing the information of job seekers through these new sites is time and cost effective. On the other hand individuals make connections on these social networking websites for different purposes. They join Facebook, to connect personally with their friends, family, groups and known for enjoyment. They join LinkedIn to widespread their scope of professional opportunities by making professional network. Job search is the main reason for individuals to make their account on LinkedIn (Stopfer & Gosling, 2013). People Join Glassdoor, an employer review website and LinkedIn to gather more detailed and realistic and trustworthy information about organizations (Kaur, T 2016; Wadhawan, 2018).

Wide use of internet and Social networking websites SNW in job search and recruitment process is increasing today. Different articles in newspapers, reports show the increasing use of these platforms by individuals. (Global trend report, 2015), gave the

AMITY
BUSINESS S(

DETERMINANTS OF CHILD LABOUR IN INDIA

[1]Mala Parashar, [2]Ashok Sharma, [3]Pramod Kumar Pandey

ABSTRACT: *Child labour is a dynamic problem of worldwide concern. The eradication of child labour is major social development goal needs to be achieved by India. This study is an attempt to discuss various factors which affects child labour, such as poverty rate, population, male literacy, female literacy, school drop-out rate, fertility rate, urbanization rate, and per capita state domestic product. In the present study the Census data for the year 2001 & 2011, along with the RBI data on per capita state domestic product for the same years have been used. The number of child workers in a state has shown a decline from 2001 to 2011. Huge state level variation in the number can be seen in study.*

Key Words: *Child labour, per capita state domestic product, poverty rate, population, male literacy, female literacy, school drop-out rate, fertility rate and urbanization rate*

I.   INTRODUCTION

Child labour is a curse to many developing and new industrial countries. Children are that group of population who need a lot of social care because of their vulnerability and dependence. They can be used, mishandle and directed into unwanted channels by unscrupulous section in the community as per National Institute of Public Cooperation and Child Development (NIPCCD), 1997. The child labour

System, which poses tremendous physical, psychological and moral risks for children themselves, often obscures the entire outlook for the poorest countries of the world. As per the estimates of International labour organization (ILO), about 218 million kids between the 5 and 17 years of age are in labour force. Out of those 218 million, 152 million are target of child labour, almost 73 million works under unsafe working conditions. The highest number of working children can be found in countries like India, Pakistan and Bangladesh. The position of child labour across countries reflects the country' economic as well as social status, the poorer the nation the higher will be the child labours.

According to Census, child labour in India observed a fluctuating trend over the years. Child labour increased from 1.08 crore in the year 1971 to 1.36 crore in the year 1981. In the past 50 years child labour was highest in the year 1981 and from then it is showing a declining trend in the numbers. The 5% of the total children aged 5-14 years were working in 2001, which accounts for the total of 1.27 crore children, whereas in 2011, the count of child labour was 1.01 crore which shows that 4% of the total children aged 5-14 were working in 2011.

---

[1] *Research Scholar, GGSIPU, New Delhi*
[2] *Associate Professor, JIMS, Vasant Kunj, New Delhi*
[3] *Research Scholar, Jagannath University, Jaipur*

Digitalized Transformation, Social Networking and its Effect on Talent Acquisition

Seema Wadhawan, Nidhi Gupta

Abstract: Advancement in technology has led to adoption of digitalized platforms for recruitment. Hiring in the current age is candidate driven. Competitive success of organizations has made it imperative for the recruiters to draw right talent. A vital appropriate digitalized professional social networking platform facilitates the recruiters to connect personally and professionally. The study attempts to analyze effectiveness of LinkedIn as a digitalized SNS platform by analyzing information content and website usability. Research was conducted through a structured questionnaire in Delhi, NCR with a sample size of 125 recruiters. Factor analysis was applied to identify varied attributes of the LinkedIn for its adoption by talent acquisition teams. Correlation and Regression enabled to study the influence of information and website quality on Intention to use LinkedIn. Findings of the study showed that perceived usefulness and information relevance influence the intention to use LinkedIn by recruiters the most.

Keywords: Information relevance, LinkedIn, Perceived Usefulness, Recruitment, Social Networking Site and Talent Acquisition, TAM (Technology Acceptance Model)

I. INTRODUCTION

In this growing global economy, to have sustainable development organizations and HR professionals are enforced to digitalize the HR function. Digital Transformation in HR practices, especially talent acquisition has overcome the challenges of attracting, hiring and retaining competent resource (Schlechter et al., 2014). The traditional channels of recruitment are important but not sufficient. Talent seekers are deprived of the semi active or passive job seekers who can be captured through digitalized tools. Recruitment in digital era is candidate driven, so it is all the more important for the organization to adopt new SNS platforms like Facebook, LinkedIn and employee blogs (Nikolaou, 2014).

Digitalized HR practices have facilitated organizations to create their employer brand. Social media platform like Facebook, LinkedIn, Glassdoor enable the companies to share business and job information with the public at large, bridging the gap between potential candidate and recruiter. In current digital age, majority of job seekers have social presence. Social Networking websites have the feature of open access, thus recruiters can gain information about the candidate; their knowledge, skills, abilities and personal information which is vital for screening process. 68% of recruiters in western countries hire potential candidates through social networking sites (Swallov, 2011).

LinkedIn is considered as most popular SNS for recruitment (Kluemper, 2013). LinkedIn is a preferred source for job search by job seekers. It enables the job seeker to build their professional network. The information content and system usage facilitates the user to increase its usage (Wadhawan, et.al., 2020). With increased use of social media for business and varied functional domain, present study focuses to determine the factors influence the recruiter's perception to use LinkedIn as SNS in talent acquisition as a preferred channel.

II. LITERATURE STUDY AND HYPOTHESES

In the current age, HRM role is widespread covering not only the operational function but also the strategic function of contributing to business. Thereby, making it imperative for the organizations to recruit most appropriate personnel's, and not just fill the vacant position. Recruitment is not an isolated organizational function. Rather, identifying, attracting and recruiting the right talent is a key success factor of any talent management strategy thus digitalized platform and applicant of SNS is essential (Gallardo-Gallardo and Thunnissen, 2016).

A. Recruitment and E-Recruitment

Recruitment is the process of finding and engaging the people the organization needs (Armstrong, 2009). Over the years the recruitment function of HR has evolved with the development of technology. Adoption of digitalized platform of SNS has been the need of the hour. In India 20th to 21st century led to adoption of internet supported web based sites such as naukri.com, monster.com, jobsahead.com.

E-Recruitment is a concept where internet is used as a tool to search the potential job applicant. Marketing discipline contributed for focusing on branding effect on job applicant (Dineen and Noe, 2009) and Information technology contributed by focusing on information system, and acceptance to use the system (Grobler et al., 2014). Over the decade the company website and job boards became the most dependable sources to increase the applicant pool with the organization. A dramatic change was witnessed in mid 2000 in job hunt activities which were recorded at 46% in 2003 (Singh, 2017). The 2017 CIPD survey of resourcing and talent planning found most effective methods for recruiting was own corporate websites (74%), followed by professional networking sites (60%), commercial job boards (58%) and recruitment consultants (52%). Recruitment majorly depend upon internet as the most vital source for recruiting till 2012. Employers have adopted Social Networking Website for recruitment and screening purposes; and among these 97% of recruiter uses LinkedIn for hiring needs (Hebberd, 2017).

An Observable Network Route Support on Interpretation of Cloud Computing

Pramod Kumar Pandey, Sakshi Chhabra, Ashok Sharma

Abstract: The Commercial cloud computing is appropriate conventional and funding agencies beyond prototyping, and initiated fund Production exercise. An important feature of any technical computing Program is moving production data, inward and outward. By means of the virtual machine performance and cost relatively glowing assumed, Network performance and cost is not. This article provides an authentication in the regions of Amazon Web Services, Microsoft Azure network and between Google clouds platforms, cloud both resources and major DTNs In research platform in the Pacific, including the Federation of OSG data cache Network backbone, cloud inside their own. This article contains both qualitative results of the analysis, as well as latency and throughput measuring. It also includes analysis of the cost of contribution Cloud Based on the network.
Keywords: Cloud Computing, AWS, Microsoft Azure, Computer Network, GCP

I. INTRODUCTION TO CLOUD COMPUTING

The commercial cloud computing gaining popularity in the field of scientific computing. Due it is very flexible nature and large total capacity, which is a great resource and prototype. It makes urgent computing needs a good platform. The fund institution Began to pay attention, explicitly mentioned in some recent grants commercial cloud Use. An important aspect of any major scientific computing project is the data migration. The Scholars and support teams need to understand the basic features Connect to the resources they use, including network latency and throughput also Costs related to the proper use of resources planning decisions under these circumstances Commercial cloud resources, performance and cost calculation illustrations Relatively well documented and understood, the same cannot be said for the network link and Move large data. In the direction of address this deficiency, we had a network characteristics Autumn campaign early in 2019, collecting information on throughput and latency, Amazon Web Services (AWS), Microsoft Azure and Google each area of cloud computing Platform (GCP), were studied between resources and major DTNs Pacific cloud Platform (PRP / TNRP), including the Open Science Grid (OSG) data federation cache Internet to backbone, and between different regions within clouds they themselves. Furthermost of the benchmark results were contrasting by the three commercial operating objects store Cloud providers. In place of storage endpoints because it provides an easy Access scalable endpoint, data migration for real life is usually designed for those the service refer to the data movement inside a cloud performance area.

The situation delivers information about the network performance between the cloud regions using the Infrastructure-as-a-Service (IaaS) cloud service provider. Moving data between the performance of cloud resources and operating hardware.

II. THE NETWORKING INSIDE CLOUD REGIONS

In the direction of measure the performance of the network in the area, we must choose a scalable Endpoint. In line for to the large pool of distributed storage of all business operations of cloud service providers object storage interface, we thought of as our Lord aims. In accumulation, many scientific computing workloads may access cloud-native objects Storage, also contrasting is very important in its own right performance such measurements to achieve. Test setup is the same service provider for all commercial cloud. We created a set of File size of 1 GB each, and upload them to the cloud object storage in each region for test the commercial cloud service providers, that is, one in AWS, Azure and a one in GCP. At that time a calculation example provisioning, and large quantities of concurrent Download on their way, the timing of the log collected to measure implemented performance. Used to perform the actual download tool is aria2 and workload Management system is HT Condor. We participated numerous tests at each test cloud zone, only in a few cases at the beginning of then progressively increase instance counts. All three test scalability Commercial cloud providers are excellent, shown in Figure 1, we further stop After scaling we exceed the total bandwidth of about 1 Tbps.
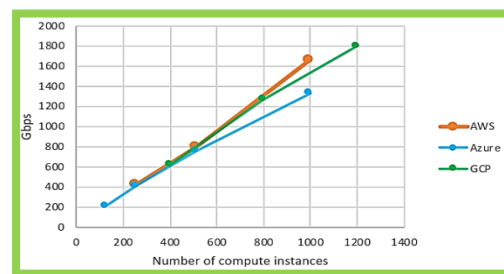
Figure-1: Showing the Highest throughput observed in a Cloud region while downloading from a local object storage instance

III. NETWORKING AMONG THE CLOUD PROVIDERS INSIDE THE REGION

An Observing to high-throughput applications distributed High Throughput Computing (DHTC) paradigm, like OSG those currently running resources, you can easily use multiple cloud area, in order to Maximum scalability

ORIGINAL ARTICLE

Artificial bee colony algorithm in data flow testing for optimal test suite generation

Snehlata Sheoran[1] · Neetu Mittal[1] · Alexander Gelbukh[2]

Abstract Meta-heuristic Artificial Bee Colony Algorithm finds its applications in the optimization of numerical problems. The intelligent searching behaviour of honey bees forms the base of this algorithm. The Artificial Bee Colony Algorithm is responsible for performing a global search along with a local search. One of the major usage areas of Artificial Bee Colony Algorithm is software testing, such as in structural testing and test suite optimization. The implementation of Artificial Bee Colony Algorithm in the field of data flow testing is still unexplored. In data flow testing, the definition-use paths which are not definition-clear paths are the potential trouble spots. The main aim of this paper is to present a simple and novel algorithm by making use of artificial bee colony algorithm in the field of data flow testing to find out and prioritize the definition-use paths which are not definition-clear paths.

Keywords Swarm intelligence · Data flow testing · Artificial intelligence · Test suite optimization · Artificial Bee Colony (ABC)

✉ Snehlata Sheoran
   snehlata2@student.amity.edu

   Neetu Mittal
   nmittal1@amity.edu

   Alexander Gelbukh
   gelbukh@cic.ipn.mx

[1] Amity University Uttar Pradesh, Noida, Uttar Pradesh, India

[2] Instituto Politécnico Nacional [IPN], Mexico City, Mexico

1 Introduction

In software development lifecycle, one of the most requisite part is software testing. This is based on the methodologies of the organization developing software, software testing and various other factors (Sommerville 2007). The software testing is performed with the intention of finding errors (Aggarwal and Yogesh 2005). For high quality software and satisfaction of users, testing needs to be done efficiently, to reduce the maintenance cost by producing more reliable and accurate results.

In literature, various optimization techniques such as differential evolution, Particle Swarm Optimization, Ant Colony Optimization and genetic algorithms find their usage in various fields. The unit test suite generation has been presented by Campos et al. (2018) and they did an empirical estimation by evolutionary algorithms. Varshney and Mehrotra (2016) considered generation of test data for data-flow coverage by making use of differential evolution strategy. Arcuri (2017) presented a novel independent objective algorithm by considering specific properties of test suite generation. Shamshiri et al. (2015) performed an empirical study for object oriented test suite generation by comparing random and genetic algorithms. Liu et al. (2000) proposed the use of data flow testing for web applications. Berndt et al. (2003) used genetic algorithm in software testing phase for breeding the software test cases. Yoo and Harman (2010) considered multi-objective test suite minimization and presented the concept of pareto efficiency by using genetic algorithm. Srivastava (2009) proposed the application of genetic algorithm for optimization of software testing. Chen et al. (2009) proposed the application of ant colony optimization for building prioritized pairwise interaction test suite. Srivastava and

Baby (2010) presented the use of ant colony optimization for all state transitions in a system by considering full and automated coverage and compared the results with genetic algorithm. Nayak and Mohapatra (2010) proposed the use of particle swarm optimization for generating test cases for data flow testing and it has been observed that the proposed approach was better than genetic algorithm by giving 100% def-use coverage. Haider et al. (2012) proposed an expert system using fuzzy logic for optimization of test suite. Srivatsava et al. (2013) presented a technique using the metaheuristic firefly algorithm for generating optimal test paths by using applicable objective function and guidance matrix introduction for traversing the graph. Mao et al. (2015) reformed the elementary ant colony optimization into discrete version for structural testing test data generation.

The artificial bee colony algorithm has great potential and widespread applications in various domains with clear scope of structural modifications for overall improvement in its performance (Karaboga et al. 2014). Lam et al. (2012) used artificial bee colony algorithm for optimization of test suite and independent path automated generation. The artificial bee colony algorithm for software test suite optimization has been implemented by Mala et al. (2009) and compared the results with ant colony optimization. McCaffrey (2009) worked on generation of pair wise test sets by application of simulated bee colony algorithm and better result were obtained as compared to other deterministic algorithms. Mala et al. (2010) proposed a technique for automated software test optimization by using the parallel behaviour of the three bees and made a comparison with random testing, genetic algorithm and sequential ABC. Dahiya et al. (2010) worked for automatic generation of structural software tests by using ABC algorithm. Baykasoğlu et al. (2007) considered solving generalized assignment problem by application of an ABC algorithm. Pham et al. (2006) presented a novel tool based on the bee's algorithm for complex optimization problems. Kulkarni et al. (2011) proposed the use of artificial bee colony algorithm for software test case prioritization for software under test in time constraint environment. Akay and Karaboga (2012) introduced the modified versions of Artificial Bee Colony algorithm for solving the real-parameter optimization problem efficiently. Gao and Liu (2012) presented modified artificial bee colony algorithm for solving complex numerical problems. Singh (2009) considered the leaf-constrained minimum spanning tree problem and implemented an artificial bee colony. Banharnsakun et al. (2011) considered artificial bee colony algorithm and presented solution updation of the onlooker bee by a modified method.

The optimization algorithm GA requires large computation time and suffers convergence problem for local minima/maxima. Further, GA is unable to effectively solve constrained optimization problems (Binitha and Sathya 2012). The PSO is a heuristic approach designed on swarm intelligence, which in presence of certain objectives generates various solutions to certain optimization problems in search space and predict the social behavior of either school of fish or flock of bird in search of food (Bashir and El-Hawary 2009). However, this technique comes with drawbacks such as partial optimism causing incomplete information with respect to various parameters used for analysis, doesn't provide feasible solutions for scattering problems and fails in implementation on non-coordinate systems. ACO is an acknowledged swarm-based algorithm used for various optimization problems (Selvi and Umarani 2010). However, ACO suffers from the following major disadvantages.

(1) There are no independent sequences of random decisions. (2) The iteration affects the probability distribution. (3) The time to convergence is not certain, although a guaranteed convergence is there. The ACO technique is a probabilistic algorithm utilizing graphs to provide the best solution for various computational problems. However, this algorithm may yield a local maxima/minima rather than a global maxima/minima Lin et al. 2013). To avoid a local maxima/minima, with a quick convergence at global maxima/minima, ABC algorithm is used (Nasiraghdam and Jadid 2012).

In the field of software testing with primary focus on data flow testing, the applications of artificial bee colony algorithm need to be explored. In order to efficiently determine and prioritize the definition-use paths which are not definition-clear paths, artificial bee colony algorithm has been applied in this paper. The rest of the paper is divided into 4 segments. Segment 2 covers testing strategies. Segment 3 presents the proposed method using ABC algorithm, segment 4 briefly presents the analysis and the result, and the last segment covers the conclusion.

2 Testing strategies

The commonly used types of software testing strategies are- black box testing and white box testing (Sommerville 2007). In black box testing, the internal structure of the code is ignored and considers the observation of the output for certain input values (Aggarwal and Yogesh 2005). In White box testing, the internal program module is tested thoroughly, which in turn helps in finding the optimal number of testing cases to guarantee a given level of coverage. White box testing incorporates three popular approaches which are Mutation Testing, Control Flow Testing and Data Flow Testing. In Data Flow Testing, the definition and usage of the variables are considered. The

definition-use paths which are not definition clear are probable concern spots.
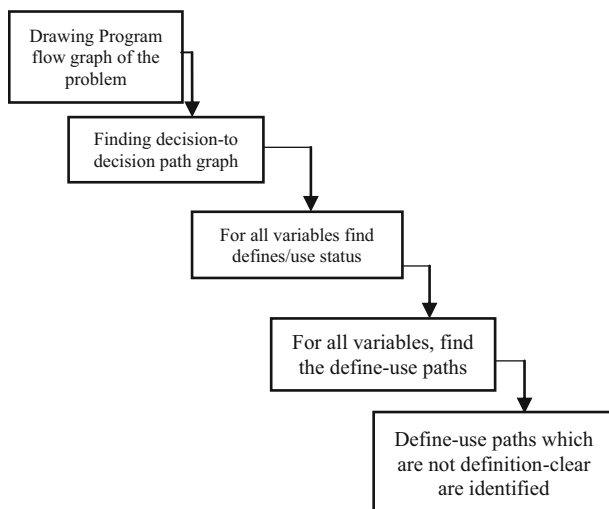
2.1 Data flow testing

Data flow testing gives the access of program's structure details to the tester and emphasize upon the variables used within a program and is a form of structural testing. In every program, first variables are defined and being used in the program. With respect to all variables in the program, the changing values are noted by the tester, through the concept of a program graph. The various definitions related to data flow testing are (Aggarwal and Yogesh 2005):

1. Defining Node: is a node where a variable is defined in the program.
2. Usage Node: is a node where the value of the variable is used in the program.
3. Definition use: It is a path where a variable is first defined and then put to use in a program.
4. Definition clear: It is a path where there is a define node and use node for a variable and no other defining node for the same variable exists in the path.

Few define/reference anomalies with respect to data flow testing are as follows (Aggarwal and Yogesh 2005):

1. A variable which has been defined but was never put to use
2. A variable which is being used, has not been defined.
3. A variable that is defined twice before it is used.

The define-use paths that are not definition clear are probable concern spots. Hence, the focus is on finding all DU-paths and then identifying those DU-paths which are not DC-paths. The steps followed in data flow testing are given in Fig. 1.

Fig. 1 Steps of data flow testing (Aggarwal and Yogesh 2005)

```
main()
1.    {
2.       int A, B, C;
3.       printf("Enter three values \n");
4.       scanf("%d%d%d", &A, &B, &C);
5.       printf("\n Largest value is :");
6.       if (A>B)
7.          {
8.             if (A>C)
9.                printf("%d\n, A);
10.           else
11.                printf("%d\n", C);
12.         }
13.      else
14.         {
15.            if(C>B)
16.                printf("%d\n", C)
17.            else
18.                printf("%d\n", B);
19.         }
20.   }
```
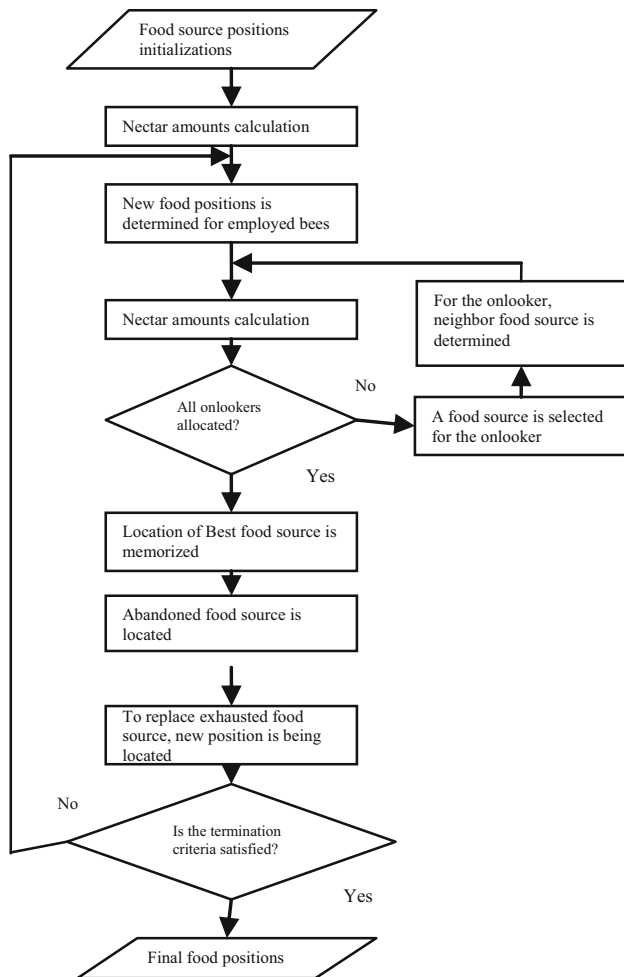
Fig. 2 Program code to find the largest number

Consider a simple program mentioned in Fig. 2. For a variable A, the define node is line number 4 and the usage node are 6 and 8. In data flow testing, there exists a define-use path for every variable. For variable A, the define-use paths are from node 4 to node 6 and from node 4 to node 8. Both the paths are definition clear as there is no redefinition node for the variable A.

2.2 Swarm intelligence

Swarm intelligence is an imperative concept in artificial intelligence where the goals are performance optimization and robustness. Collective behaviour emerged from social insects forms swarm intelligence in which social exchanges amongst the individual representative help in finding out the optimal solutions for NP-hard problems. Various types of swarm intelligence algorithms used for optimization are bat algorithm, particle swarm optimization, firefly algorithm, ant colony and bee colony.

The first studies related to Artificial bee colony algorithm included the performance evaluation on set of numerical benchmark test functions along with the comparison with PSO, DE, ACO and GA. Artificial bee colony algorithm is a single-objective optimizer, which motivated the researchers for the application of Artificial bee colony algorithm in other areas as well such as neural networks, software engineering, industrial engineering, mechanical and civil engineering and image processing. The emphasis stretched from solving unconstrained optimization problems to constrained problems also for ABC algorithm (Karaboga et al. 2014). There are 3 types of bees considered in artificial bee colony algorithm: the employed bee, onlooker bee and the scout bee. The random searches in the environment are carried out by scout bee. Employed bee is

Fig. 3 ABC algorithm flow chart (Karaboga 2009)

```
#inlcude<stdio.h>
#include<conio.h>
#include<math.h>
1       int main()
2       {
3       Int a, b, c, validinput=0, d;
4       Double D;
5       Printf("Enter the 'a' value:");
6       Scanf("%d",&a);
7       Printf("Enter the 'b' value:");
8       Scanf("%d",&b);
9       Printf("Enter the 'c' value:");
10      Scanf("%d",&c);
11      If((a>=0) && (a<=100) && (b>=0) && (b<=100) && (c>=0) &&
        (c<=100)) {
12          Validinput=1;
13              If(a==0){
14                  Validinput=-1;
15                  }
16      }
17      If(validinput==1){
18          d=b*b – 4*a*c;
19              if (d==0){
20              printf("The roots are equal and are r1,=r2=%f\n",-b/(2*(float)
        a));
21      }
22          Elseif(d>0) {
23              D=sqrt(d);
24          Printf("The roots are real and are r1=%f and r2=%f\n",(-b-
        D)/(2*a),(-b+D)/(2*a));
25              }
26          Else{
27              D=sqrt(-d)/(2*a);
28              Printf("The roots are imaginary and are r1= (%f, %f)
        and r2= (%f ,%f)\n", (2.0*a), D, -b/(2.0*a), -D);
29              }
30      }
31      Elseif(validinput==-1){
32              Printf("The values do not constitute a Quadratic Equation.");
33      }
34      Else{
35              Printf("The inputs belong to invalid range");
36      }
37      Getch();
38      Return 1;
39      }
```

Fig. 4 Program code of quadratic equation problem (Aggarwal and Yogesh 2005)

one who visits the food source which it had visited formerly, whereas the onlooker bee is the one waiting for decision making in the beehive. Steps involved in the algorithm are summarized below (Karaboga and Basturk 2007):

1. Initialization of population.
2. Repeat.
   a. Employed bees are being positioned on their respective food sources.
   b. Keeping in view the nectar amounts, the onlooker bees are positioned on their respective food sources.
   c. For locating the new food sources, the scout bees are sent to the search area.
   d. Position of best food source so far is memorized.
3. Until requirements are met.

The flow chart for Artificial Bee Colony algorithms is (Fig. 3).

Onlooker bee with the help of probabilistic value, selects a food source and this value of p_i is calculated as (Karaboga and Basturk 2007):

$$p_i = \frac{fit_i}{\sum_{n=1}^{SN} fit_n}$$

where the number of food sources if provided by SN and this count also represents the count of employed bees (BN). The value proportional to nectar amount of the food source is the fitness value *fit* for solution *i* and is being evaluated by its employed bee. For producing food position candidate with respect to old one, the algorithm uses the following expression:

$$v_{ij} = x_{ij} + \emptyset_{ij}(x_{ij} - x_{kj})$$

where *k* and *i* are different from each other and the randomly chosen indexes are represented by $k \in \{1, 2, \ldots BN\}$ and $j \in \{1, 2 \ldots D\}$. $\emptyset_{ij}$ is also a random number which lies between $[-1, 1]$ (Karaboga and Basturk 2007).
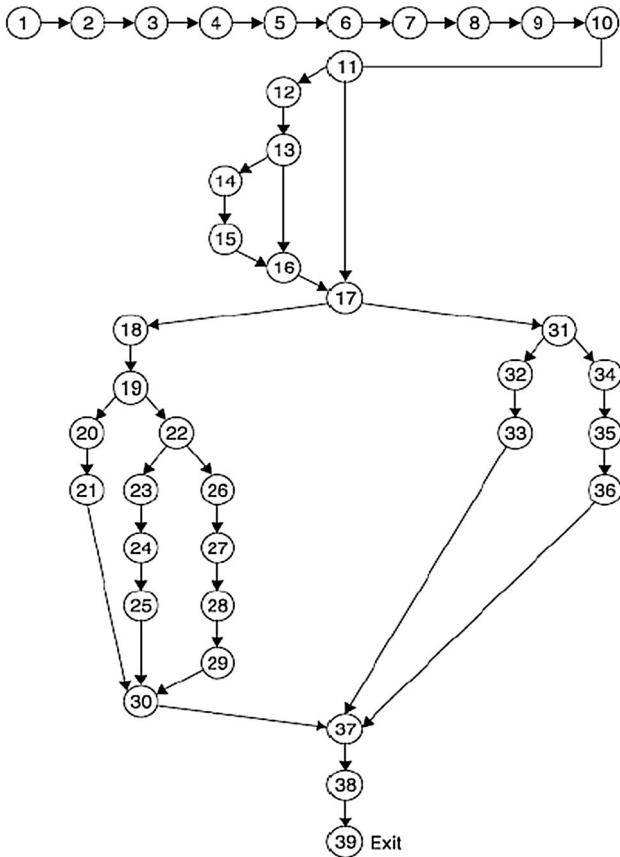
Fig. 5 CFG of quadratic equation problem (Aggarwal and Yogesh 2005)

Table 1 Define and use status of variables of problem (Aggarwal and Yogesh 2005)

Variable	Define node	Use node
a	6	11, 13, 18, 20, 24, 27, 28
b	8	11, 18, 20, 24, 28
c	10	11, 18
d	18	19, 22, 23, 27
D	23, 27	24, 28
Validinput	3, 12, 14	17, 31

3 Proposed methodology

In this work, the concepts of data flow testing have been implemented on eight benchmark problems using artificial bee colony algorithm.

The first step before applying this algorithm is to design a control flow graph for the problem considered and note down the define and use node for all variables mentioned in the program. Further, the out degree of all decision/predicate nodes is tabulated. The program code and CFG are given in Figs. 4 and 5.

Table 2 DU-paths with DC status (Aggarwal and Yogesh 2005)

Variable	Path	Definition clear
a	6–11	Yes
	6–13	Yes
	6–18	Yes
	6–20	Yes
	6–24	Yes
	6–27	Yes
	6–28	Yes
b	8–11	Yes
	8–18	Yes
	8–20	Yes
	8–24	Yes
	8–28	Yes
c	10–11	Yes
	10–18	Yes
d	18–19	Yes
	18–22	Yes
	18–23	Yes
	18–27	Yes
D	23–24	Yes
	23–28	Path not possible
	27–24	Path not possible
	27–28	Yes
Validinput	3–17	No
	3–31	No
	12–17	No
	12–31	No
	14–17	Yes
	14–31	Yes

The define node and use node for all the variables of Quadratic equation are represented in Table 1. The 'a' variable is defined at node 6 while it is put to use at node 11, 13, 18, 20, 24, 27 and 28 in the program. The 'b' variable is defined at node 8 and used at 11, 18, 20, 24 and 28. Similarly other variables are also defined and used at various nodes in the program.

Considering the define and use node of the variable, define-use paths for every variable are listed in Table 2 with definition clear status. Along any DU-path, if a variable is getting redefined again, then that path is considered as non-DC path.

Since the information related to CFG and define-use nodes has been gathered, the following assumptions are made for the global search algorithm:

1. 'N' is the total number of nodes in the CFG.
2. 'DN' is the total number of decision nodes in the CFG
3. 'OD[i]' is the outdegree of each decision node.

4. 'D[k]' (¥ all k = 1 to DN) represents an array of decision nodes.
5. 'P[i]' is the path saved after iteration.
6. 'All [P1, P2…..Pi]' is an array of all saved paths.
7. 'Vs' is the visited status of the nodes.

According to Algorithm 1, first the DU pair not yet traversed is selected and define and use nodes of the variable are set as start and end node. Then declare the number of decision nodes for that DU pair and also input the out degrees for each decision node. Initially, the visiting status of all the nodes is set to 0. The scout bee starts from the define node of the variable. It visits the first node and increments the visiting status by one. The scout bee then checks whether the visited node is an end node or decision node or exit node. If the visited node is an end node then that path is saved. If it is a decision node, then decision is being made where to proceed further by checking the next feasible node. If the visited node is an exit node, the path is discarded. Likewise, all the global paths are generated until the visiting status of all decision nodes becomes greater than their out degrees. The paths generated after global search are tabulated in Table 3 for the variable 'validinput' (Fig. 6).

Table 3 is for the variable 'validinput' as only for this variable, there are DU-paths which are not DC-paths. For this variable, the DU-pairs considered are 3–17, 3–31, 12–17 and 12–31 and not 14–17 and 14–31. For DU–pair 3–17, there are 3 unique paths: p1 is 3-4-5-6-7-8-9-10-11-17, p2 is 3-4-5-6-7-8-9-10-11-12-13-16-17 and p3-3-4-5-6-7-8-9-10-11-12-13-14-15-16-17. One path gets discarded as this path is the redundant path. Likewise, for every other DU-pair, unique paths are identified and redundant paths or the paths which does not end on use node are discarded.

Once the information about global path or unique path gets collected, then perform the local search. Assumptions made for local search Algorithm 2 are:

1. 'All [P1, P2,…Pi]' contains all the saved paths from Algorithm 1.
2. 'EB' represents the count of employed bees, which is also equal to the count of independent paths saved.
3. 'Keynode[k]' is the redefinition nodes for a variable.

Algorithm 2 conducts a local search with the help of employed and onlooker bees. Basically, in local search, the employed bees have a set of randomly generated solutions and the employed bees search within those solutions to find out the food source with the highest nectar quantity. On similar lines, Algorithm 2 has been proposed. Here the randomly generated solution is the saved path list generated by the scout bee with the help of Algorithm 1. The scout bee saves an independent path between the define node and the end node and shares this information with

Table 3 Paths generated after Algorithm 1

DU pair	Path number	Path selected	Path discarded
3–17	P1	3-4-5-6-7-8-9-10-11-17	3-4-5-6-7-8-9-10-11-17
	P2	3-4-5-6-7-8-9-10-11-12-13-16-17	
	P3	3-4-5-6-7-8-9-10-11-12-13-14-15-16-17	
3–31	P4	3-4-5-6-7-8-9-10-11-17-31	3-4-5-6-7-8-9-10-11-12-13-14-15-16-17-31
	P5	3-4-5-6-7-8-9-10-11-12-13-16-17-31	
	P6	3-4-5-6-7-8-9-10-11-12-13-14-15-16-17-31	3-4-5-6-7-8-9-10-11-12-13-14-15-16-17-31
			3-4-5-6-7-8-9-10-11-12-13-14-15-16-17-18-19-22-23-24-25-30-37-38-39
			3-4-5-6-7-8-9-10-11-12-13-14-15-16-17-18-19-20-21-30-37-38-39
			3-4-5-6-7-8-9-10-11-12-13-14-15-16-17-18-19-22-26-27-28-29-30-37-38-39
			3-4-5-6-7-8-9-10-11-12-13-14-15-16-17-18-19-20-21-30-37-38-39
			3-4-5-6-7-8-9-10-11-12-13-14-15-16-17-18-19-22-23-24-25-30-37-38-39
12–17	P7	12-13-16-17	
	P8	12-13-14-15-16-17	
12–31	P9	12-13-16-17-31	12-13-14-15-16-17-31
	P10	12-13-14-15-16-17-31	12-13-14-15-16-17-31
			12-13-14-15-16-17-18-19-22-23-25-30-37-38-39
			12-13-14-15-16-17-18-19-20-21-30-37-38-39
			12-13-14-15-16-17-18-19-22-26-27-28-29-30-37-38-39
			12-13-14-15-16-17-18-19-20-21-30-37-38-39
			12-13-14-15-16-17-18-19-22-23-24-25-30-37-38-39

1. If (Any DU pair not yet covered)
   Goto step 2
   Else (Goto Step 8)
2. Select DU pair not yet covered.
3. Initialization:
   Set Start = Define node
   Set End = Use node
   int N, DN, OD[i][j], D[i], P[i], All[P1,
   P2…Pi]
   Vs = 0 (¥ all k = 1 to N)
4. Move from Start to Next Node.
   V_s (Start) = V_s (Start) + 1
5. Do
   Vs (Next node) = Vs (Next Node) + 1
   If {(Next Node = End Node)
       {Save Path in P[i]
       Compare P[i] with All[P1, P2,…Pi], If
   redundant,
       Then discard path
       If (V_s (¥ all k = 1 to DN) ≥ OD[i])
           Goto Step 1.
       Else (Goto Step 4)}}
   Else {If {(Next Node = Decision Node)
           Select according to the following
   rule:
   A. If {(Feasible node = End node) && (V_s (End
       Node) < DN)}, then select that node and goto step
       6, else follow B.
   B. Select the node with smallest V_s and goto step 6,
       else follow C.
   C. If (Feasible node = Decision node), then select that
       node and goto step 6, else follow D.
   D. Select randomly and goto step 6. }
       Else {If (Next Node = Exit)
               Discard Path and goto step 6
               Else (Move to next adjacent
   node and goto step 5)}}
6. Repeat step 5 untill V_s(¥ all k = 1 to DN) > OD[i].
7. Goto step 1.
8. Stop.

Fig. 6 Proposed Algorithm 1 for global search

1. If {(Any DU pair not yet checked)
   Goto Step 2
   Else (Goto step 8)}
2. Select DU pair not yet checked.
3. Select all saved paths [All[P1, P2,…Pi]] along that
   DU pair.
4. int EB, keynode[k]
5. For each EB along every path
   Initialization:
       Set Start = Define node
       Set End = Use node
       Count = 0
6. Move from start node to next node.
7. If {(Next Node = Keynode)
       Count = Count +1
       Move to next adjacent node and follow step 7
   Else {If (Next node = End node)
       Return Count
       Goto step 1
       Else (Move to next adjacent node and follow 7)}}
8. Stop.

Fig. 7 Proposed Algorithm 2 for local search

Table 4 Paths generated after Algorithm 2 for quadratic equation

Variable	Defined at	DU Pair	Path number	Count
Validinput	3, 12, 14	3–17	P1	0
			P2	1
			P3	2
		3–31	P4	0
			P5	1
			P6	2
		12–17	P7	0
			P8	1
		12–31	P9	0
			P10	1

Table 5 Algorithm 1 results

Example	Total iterations	Path saved
Quadratic	25	10
Triangle	23	8
Next date	32	16
3 variables	34	28
Fibonacci	6	3
Swapping	33	19
Age problem	7	7
Largest side	10	10

other employed bees in the hive by performing the waggle dance (Fig. 7).

As it is known that the count of employed bee in the hive is equal to the count of food sources, therefore in Algorithm 2 the number of employed bees used is equal to the number of independent paths found between the DU pair by the scout bee. The employed bees, based on their previous experience and memory, also have information about different rich food sources in the neighborhood of the hive. This memorization functionality of the employed bee is also used in Algorithm 2. Here the employed bee memorizes the keynodes i.e. the nodes where the variable has been defined again before the actual use of that variable. Now as the employed bee moves along the path provided by the scout bee, bee compares every node with the keynode saved in memory. If a match is found by the employed bee, than that employed bee recruits onlooker bees along that path as that path helps in collecting more nectar. Match found, represented by count, is directly proportional to the number of recruited onlooker bee. If no match for keynode is found along the saved path than the employed bee would become a scout bee after reaching the end node of that path. Therefore, the Algorithm 2 helps in identifying the paths where maximum number of times the
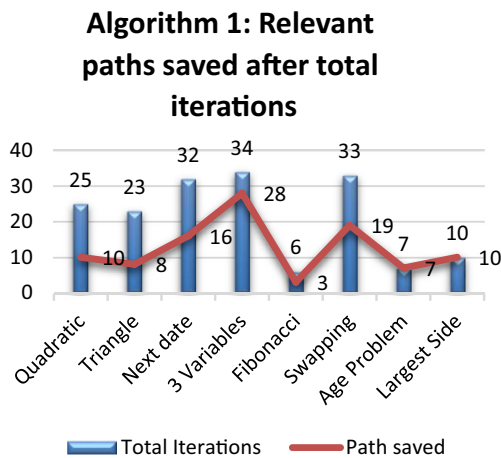
Fig. 8 Comparison of total iterations and path saved

Fig. 9 Comparison of outputs

variable has been redefined before the actual use of the variable. Hence the path where the employed bee recruits maximum number of onlooker bees is the path that needs to be tested with highest priority, as this path is not definition clear for the variable and has maximum number of redefinitions for the variable. The result for quadratic equation is tabulated in Table 4.

In Table 4, the variable considered is 'validinput' and the DU-pairs are 3–17, 3–31, 12–17 and 12–31. Along 3–17 DU-pair, there are three unique paths represented as p1, p2 and p3. Along p3, the variable is defined twice before it is put to use, whereas in p2, there is only one redefinition of the variable before usage of the same. In p1, there is no redefinition node. Therefore, it is suggested that in time-constraints environments, p3 should be tested first followed by p2 and last would be p1. Similarly, along other DU-pairs also, prioritization of paths is done.

4 Results and discussion

The proposed method has been established on eight standard problems. Table 5 provides the number of paths obtained from global search. For each example, the total
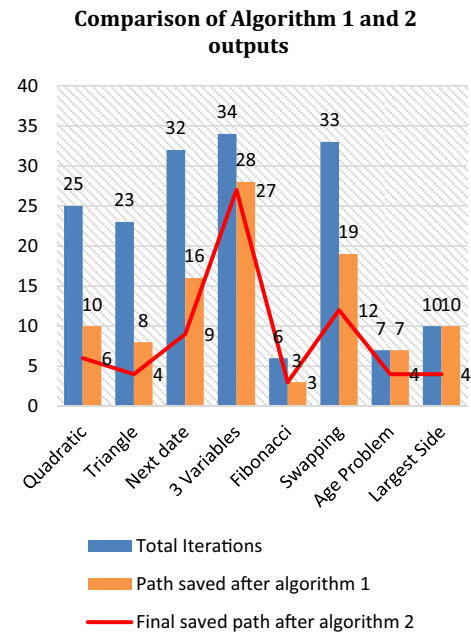
iterations represent the number of times the scout bee randomly searched for path reaching the end node. Saved path represents the successful paths where the scout bee was able to trace the end node. Path saved is less than the total iterations because the paths that did not end at the end node have been discarded. Saved paths also represents the number of independent paths along a DU pair for a particular variable from the define node to the end node (Fig. 8).

The comparison of results is shown in Table 6. Total iterations represent the random search conducted by the scout bee. Paths saved after Algorithm 1 are the independent paths along a DU pair. From Table 6, it has been analyzed that the total number of iterations for Quadratic problem are 25, which reflects the random paths searched by scout bee. Out of 25 paths, 10 paths are DU paths and at the time of testing the focus should be on 6 non-DC-paths for testing as these are the potential trouble spots. Similarly, for Triangle problem, 4 non-DC paths should be

S. no	Example	Total iterations	Path saved after Algorithm 1	Final saved path after Algorithm 2
1	Quadratic	25	10	6
2	Triangle	23	8	4
3	Next date	32	16	9
4	3 variables	34	28	27
5	Fibonacci	6	3	3
6	Swapping	33	19	12
7	Age problem	7	7	4
8	Largest side	10	10	4

Table 6 Comparative results of Algorithms 1 and 2

Table 7 Test suite savings

Examples	Total DU paths	DU not DC paths	Test suite saving in %
Quadratic	10	6	40
Triangle	8	4	50
Next date	16	9	43.75
3 variables X	12	12	0
3 variables Y	13	13	0
3 variables Z	3	2	33.34
Fibonacci	3	3	0
Swapping	10	6	40
Age	7	4	42.86
Largest side	7	4	42.86

tested firstly as compared to all 8 DU paths. In Next Date problem, there are 16 DU paths and 9 non-DC paths from a total of 32 paths. 27, 3, 12, 4 and 4 non-DC paths are there for 3 Variables, Fibonacci, Swapping, Age problem and Largest side, need to give preference at the time of testing as compared to all 28, 3, 19, 7 and 10 DU paths. Thus, after Algorithm 2, one can select all non-DC paths and avoid all DU-paths for testing. This in turn helps in generating optimal test suite in time constraint environment. Thus, when the tester is testing the software under development, more emphasis should be put on testing the paths saved after Algorithm 2 as the paths which are not definition clear paths are the potential concern spots (Fig. 9).

From Table 7, it has been observed that for Quadratic problem, 40% of effort has saved while forming the test suite for testing as only 6 paths needs to be tested in time constraint environment out of 10 DU paths. In Triangle problem, 50% saving in test suite has achieved. The test
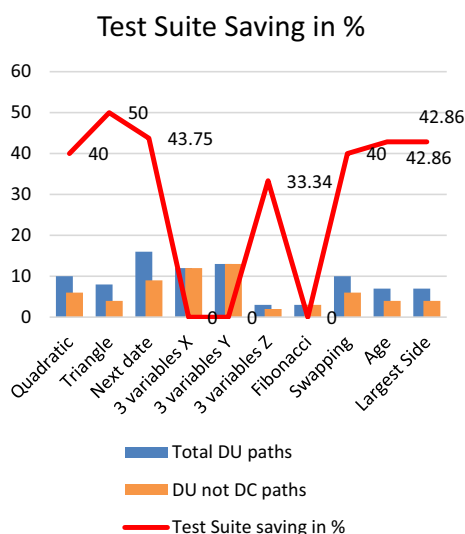
suite saving obtained for Next date, 3 variables Z, Swapping, Age and Largest side problem are 43.75%, 33.34%, 40%, 42.86% and 42.86% respectively. This reflects the efficiency of proposed algorithms (Fig. 10).

5 Conclusion

The artificial bee colony algorithm is simple and easy to implement. The local and global searches conducted by artificial bee colony algorithm for path extraction in data flow testing environment gives excellent results. The proposed algorithms have been applied on eight examples and the global search capability makes possible to trace all define-use paths by discarding redundant paths. The local search done by algorithm identifies the non-definition clear paths out of all define-use paths. Further, this has been suggested that these paths to be tested first to determine the errors in a time constraint environment. This prioritization of paths also helps in building the optimal test suite by suggesting the paths for which test cases to be generated and thus yields in saving of time, effort and resulting in generating test suite of optimal size. In this implementation of artificial bee colony algorithm, prior control flow graph design with details of define and use status of each variable are required. The proposed work is limited to non-loop programs only. However, in future, this work may be extended to programs with single and multiple loops.

Fig. 10 Display of test suite savings

References

Aggarwal KK, Yogesh S (2005) Software engineering, 2nd edn. New Age International Publishers, New Delhi

Akay B, Karaboga D (2012) A modified artificial bee colony algorithm for real-parameter optimization. Inf Sci 192:120–142

Arcuri A (2017) Many independent objective (MIO) algorithm for test suite generation. In: International symposium on search based software engineering (pp. 3–17). Springer, Cham

Banharnsakun A, Achalakul T, Sirinaovakul B (2011) The best-so-far selection in artificial bee colony algorithm. Appl Soft Comput 11(2):2888–2901

Bashir ZA, El-Hawary ME (2009) Applying wavelets to short-term load forecasting using PSO-based neural networks. IEEE Trans Power Syst 24(1):20–27

Baykasoğlu A, Özbakır L, Tapkan P (2007) Artificial bee colony algorithm and its application to generalized assignment problem. In: Swarm intelligence, focus on ant and particle swarm optimization. InTech

Berndt D, Fisher J, Johnson L, Pinglikar J, Watkins A (2003) Breeding software test cases with genetic algorithms. In: Proceedings of the 36th annual Hawaii international conference on system sciences (pp. 10). IEEE

Binitha S, Sathya SS (2012) A survey of bio inspired optimization algorithms. Int J Soft Comput Eng 2(2):137–151

Campos J, Ge Y, Albunian N, Fraser G, Eler M, Arcuri A (2018) An empirical evaluation of evolutionary algorithms for unit test suite generation. Inf Softw Technol 104:207–235

Chen X, Gu Q, Zhang X, Chen D (2009) Building prioritized pairwise interaction test suites with ant colony optimization. In: 2009 ninth international conference on quality software, pp 347–352. IEEE

Dahiya SS, Chhabra JK, Kumar S (2010) Application of artificial bee colony algorithm to software testing. In: 2010 21st Australian software engineering conference, pp 149–154. IEEE

Gao WF, Liu SY (2012) A modified artificial bee colony algorithm. Comput Oper Res 39(3):687–697

Haider AA, Rafiq S, Nadeem A (2012) Test suite optimization using fuzzy logic. In 2012 international conference on emerging technologies, pp 1–6. IEEE

Karaboga N (2009) A new design method based on artificial bee colony algorithm for digital IIR filters. J Frankl Inst 346(4):328–348

Karaboga D, Basturk B (2007) A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm. J Glob Optim 39(3):459–471

Karaboga D, Gorkemli B, Ozturk C, Karaboga N (2014) A comprehensive survey: artificial bee colony (ABC) algorithm and applications. Artif Intell Rev 42(1):21–57

Kulkarni NJ, Naveen KV, Singh P, Srivastava PR (2011) Test case optimization using artificial bee colony algorithm. In: International conference on advances in computing and communications, pp 570–579. Springer, Berlin

Lam SSB, Raju MHP, Ch S, Srivastav PR (2012) Automated generation of independent paths and test suite optimization using artificial bee colony. Procedia Eng 30:191–200

Lin Y-K, Yeh C-T, Huang P-S (2013) A hybrid ant-tabu algorithm for solving a multistate flow network reliability maximization problem. Appl Soft Comput 13:3529–3543

Liu CH, Kung DC, Hsia P (2000) Object-based data flow testing of web applications. In: Proceedings first Asia–Pacific conference on quality software, pp 7–16. IEEE

Mala DJ, Kamalapriya M, Shobana R, Mohan V (2009) A non-pheromone based intelligent swarm optimization technique in software test suite optimization. In: 2009 international conference on intelligent agent and multi-agent systems, pp 1–5. IEEE

Mala DJ, Mohan V, Kamalapriya M (2010) Automated software test optimisation framework—an artificial bee colony optimisation-based approach. IET Softw 4(5):334–348

Mao C, Xiao L, Yu X, Chen J (2015) Adapting ant colony optimization to generate test data for software structural testing. Swarm Evolut Comput 20:23–30

McCaffrey JD (2009) Generation of pairwise test sets using a genetic algorithm. In: 2009 33rd annual IEEE international computer software and applications conference, vol 1, pp 626–631. IEEE

Nasiraghdam H, Jadid S (2012) Optimal hybrid PV/WT/FC sizing and distribution system reconfiguration using multi-objective artificial bee colony (MOABC) algorithm. Sol Energy 86:3057–3071

Nayak N, Mohapatra DP (2010) Automatic test data generation for data flow testing using particle swarm optimization. In: International conference on contemporary computing, pp 1–12. Springer, Berlin

Pham DT, Ghanbarzadeh A, Koç E, Otri S, Rahim S, Zaidi M (2006) The bees algorithm—a novel tool for complex optimisation problems. In: Intelligent production machines and systems, pp 454–459. Elsevier Science Ltd, Amsterdam

Selvi V, Umarani R (2010) Comparative analysis of ant colony and particle swarm optimization techniques. Int J Comput Appl 5(4):1–6

Shamshiri S, Rojas JM, Fraser G, McMinn P (2015) Random or genetic algorithm search for object-oriented test suite generation? In: Proceedings of the 2015 annual conference on genetic and evolutionary computation, pp 1367–1374. ACM

Singh A (2009) An artificial bee colony algorithm for the leaf-constrained minimum spanning tree problem. Appl Soft Comput 9(2):625–631

Sommerville I (2007) Software engineering, Eight edn. Pearson Education Limited, Harlow

Srivastava PR (2009) Optimisation of software testing using genetic algorithm. Int J Artif Intell Soft Comput 1(2–4):363–375

Srivastava PR, Baby K (2010) Automated software testing using metahurestic technique based on an ant colony optimization. In: 2010 international symposium on electronic system design, pp 235–240. IEEE

Srivatsava PR, Mallikarjun B, Yang XS (2013) Optimal test sequence generation using firefly algorithm. Swarm Evolut Comput 8:44–53

Varshney S, Mehrotra M (2016) A differential evolution based approach to generate test data for data-flow coverage. In: 2016 international conference on computing, communication and automation (ICCCA), pp 796–801. IEEE

Yoo S, Harman M (2010) Using hybrid algorithm for pareto efficient multi-objective test suite minimisation. J Syst Softw 83(4):689–701

Analysis on application of swarm-based techniques in processing remote sensed data

Snehlata Sheoran[1] · Neetu Mittal[1] · Alexander Gelbukh[2]

Abstract

The remote sensed satellite images are big repository of information and provide the coverage of large areas. However, these images may not be able to describe the finer details of area being covered. Satellite Image optimization is the process of augmenting the components of an image for better and effective interpretations from satellite images. In order to obtain better visibility properties to fetch more information, various artificial intelligence techniques can be considered for the optimization process. Finding out the best technique for optimization is a challenging and time-consuming task [U1]. In this paper, applications of swarm-based artificial intelligence techniques such as ant colony optimization, particle swarm optimization, bat algorithm, artificial bee colony algorithm etc. are being analysed to process the remote sensed data. The detailed comparison with respect to classifier, utility, images considered, and observation are discussed. The comprehensive analysis revealed that particle swarm optimization is the most widely used technique. Further, various application areas such as land-use land-cover are discussed with possibilities of future research [U2].

Keywords Satellite images · Artificial intelligence (AI) · Swarm intelligence · Image processing · Remote sensing · Optimization · Change detection · Classification · Land-use land-cover · Segmentation · Particle swarm optimization (PSO) [U3]

Introduction

Image Processing is one of the major fields of signal processing. In order to augment the worth of an image or to draw worthwhile information from an image, the images need to be manipulated using various image processing techniques (Jain 1989). With the help of digitization, an image is converted into a suitable form to store on a computer device. The fundamental requirements are that, image must be sampled and quantized. Once the image has been stored, image processing operations may be executed on

the acquired image to get better and flawless information. Typical stages in image processing are acquisition and enhancement of an image. These techniques are used to carry out the restoration, segmentation (Zaitoun and Aqel 2015), object recognition, improvement in interpretability (Gajdhane and Deshpande 2014), better representation and description (Demirel et al. 2009) of an image [U4]. Image acquisition involves various pre-processing tasks such as histogram equalization, scaling, brightness control etc. These tasks are required for improvement of image data to suppress unwanted distortions for further processing. The image enhancement and restoration techniques primarily focus on refinements to control the appearance of an image (Demirel et al. 2009). The process of segregating an image into various sub-parts is known as image segmentation (Zaitoun and Aqel 2015). The image segmentation technique includes the partitioning of an image into sub-parts or objects and is a key step from image processing to image analysis. Recognition methods utilize the difference of grey values of an image (Lee et al. 1994). The image description provides the extraction of features to result in some quantifiable information of interest for segregating one class of objects from another class. The initial

---

✉ Snehlata Sheoran
  snehsheoran312@gmail.com

  Neetu Mittal
  nmittal1@amity.edu

  Alexander Gelbukh
  gelbukh@gelbukh.com

1 Amity University Uttar Pradesh, Noida, Uttar Pradesh, India

2 Instituto Politécnico Nacional (IPN), Mexico, Mexico

applications of image processing were majorly focused in the areas of news-paper and allied industries. With further advancements in technology, the application areas of image processing technology have increased to diversified fields such as medical imaging, Pattern recognition, robot vision and remote sensing imagery [U5].

The need for accurate, fast and cost-effective geospatial information provided by remote sensing technology is increasing day by day (Mohammadzadeh et al. 2009). The remote sensing technique contributes to take important decisions by providing the accurate information for a large area (Campbell and Wynne 2011). Many image processing and analysis techniques have been developed to aid the interpretation of remotely sensed images and to extract as much information as possible from the images. As the manual extraction and analysis techniques tend to be expensive with respect to efforts required, time consumed and quality. Therefore, there is a requirement for automatic image feature extraction technique to expedite the process, and thereby considerably reducing the cost, computation time and better interpretation of information from satellite images (Mohammadzadeh et al. 2009). Artificial intelligence (AI) may replicate the human intelligence model or some natural phenomena and is considered to be a part of machine learning (Alpaydin (2009). The conept of artificial intelligence has been introduced to develop human intelligence in machines (McCorduch 1979). AI is able to provide more precision with higher degree of accuracy. Enhanced capabilties for the spatial databases may also be developed with the help of AI (McKeown 1987). In AI, nature or bio inspired meta-heuristics algorithms models their behaviour upon the naturally occuring phenomena and are developed for optimization problems to provide sufficiently good results and makes sure that the computation or size are not increasing (Singh et al. 2017). AI based techniques are capable of improving the image analysis, to give precise information regarding geographic conditions and earth resources applications (Estes et al. 1986). Thus, the artificial intelligence techniques viz.- Genetic Algorithm, Particle Swarm Optimization, Cuckoo Search etc. may play a vital role to improve the quality and interpretation process of available satellite images. These techniques may be efficiently utilized to carry out various image processing operations such as image enhancement, segmentation, contrast variation, noise removal etc. [U6].

In satellite images due to huge data size, large time is consumed to interpret the correct information. Further, the quality of satellite images is affected by weather conditions. In order to reveal the finer information and to improve the visibility qualities of satellite images, there is a need for implementation of suitable artificial intelligence technique. This paper presents detailed comparisons of various swarm intelligence-based techniques such as ant colony optimization, artificial bee colony algorithm, particle swarm optimization, bat algorithm etc.

with respect to classifier, utility, images used, observations including the advantages/disadvantages and comparisons [U7]. Further, applications, advantages and disadvantages of various swarm-based techniques in numerous areas of satellite images processing have been presented. The particle swarm optimization is one of the most widely used technique, further its application areas with future research scope have been discussed [U8]. The section 2 consists of literature review on implementation of swarm intelligence in satellite imagery. Section 3 consists of discussion drawn from section 2. The last section discusses the conclusion.

Remote sensing imagery

Remote Satellite Images are like reservoir of useful and interesting information. From these images, one can find out transformations of cities, crop classification and cultivation pattern, and damage assessment in case of natural calamities such as floods, fire, storms. Satellite images act like a rich repository of information related to agricultural activities. The focus on vegetation helps in understanding the crop growth from planting to harvest along with the observation of abnormalities and the change due to season progression (Senthilnath et al. 2008). Classification in remote sensing images is also a very crucial activity and is frequently carried out for obtaining the information related to land-use land-cover. Environmental changes at global, regional and local level can also be monitored closely with the help of land-use land-cover and the changes occurring over time (Turner et al. (1994) [U9]. Visible, Infrared and Water Vapor Imagery are the three important classifications of satellite images. Sunlight disseminated by objects suspended in the air or on Earth represents Visible Imagery. Infrared Imagery identifies the clouds by measuring the heat radiation using satellite sensors. Measurement of moisture in the upper atmosphere represents Water Vapor Image. Satellite image optimization with respect to registration, enhancement, classification and segmentation is achieved by various techniques such as discrete wavelet transformation, swarm intelligence, fuzzy logic, singular wave decomposition etc.

Swarm intelligence

Swarm Intelligence is an imperative concept in Artificial Intelligence with primary aim of performance optimization and robustness. Swarm intelligence is a bottom-up approach and behaves like a multi-agent system, where there are plentiful simple beings such as birds, fish, ants etc., and these beings work in full cooperation and competition among the individuals (Liu et al. 2008b) [U10]. Collective behavior emerged from social insects' forms swarm intelligence in which social exchanges amongst the individual representative help in finding out the optimal solutions for NP-hard problems. Different sorts of swarm optimization which are being

utilized for streamlining are Artificial Bee Colony, Particle Swarm Optimization, Firefly Algorithm, Ant Colony, Bat Algorithm and so on. Swarm intelligence has been successful in solving complex problems such as network routing, pattern recognition, travelling salesman problems, data clustering and is currently a hot research topic in artificial intelligence (Liu et al. 2008b). Classification and feature extraction research based on swarm intelligence indicates that the new clasiification and intelligence computation methods helps in avoiding the imapct on classified results generated by artificial fault or deviation, improves the clasification validity and humanity and in reforming the robustness of the algorithm to operation management (Dong and Xiang-bin 2008).[U11].

Literature review

Optimization of satellite images in terms of enhancement, segmentation, classification, clustering is accomplished by actualizing swarm intelligence methods such as Ant Colony, Particle Swarm Optimization, Cuckoo Search and so on. The following section covers the review of application of various swarm intelligence algorithms. The major advancements in the area of swarm intelligence techniques initiated from the year 2006 onwards; this study considers the period from 2006 to 2018 and is divided into four sub-sections. The strings such as swarm intelligence in remote sensing, artificial intelligence in remote sensing and nature inspired algorithms have been used with prime consideration of remote sensing images. [U12].

Analysis from 2006 to 2008

Das et al. (2006) presented a hybrid framework comprising Particle Swarm Intelligence (PSO) and Rough-Set theory for image clustering. Zhong et al. (2006) worked in the area of classification of remote sensed data by the application of an innovative approach using Unsupervised Artificial Immune Classifier. In the first step, the clustering centres were randomly carefully chosen from the input images and later the classification task was carried out. Omkar et al. (2007) implemented Ant Colony Optimization and Particle Swarm Optimization for satellite image classification problem of land cover mapping. Monteiro and Kosugi (2007) presented a feature selection algorithm for remote data by implementing Particle Swarm Optimization (PSO). The method utilized swarm implementation for optimizing desired performance criteria and the count of selected features simultaneously. Senthilnath et al. (2008) implemented Particle Swarm Optimization, Maximum Likelihood Classifier (MLC) and Ant Colony Optimization in the area of crop coverage classification using high resolution satellite images. Liu et al. (2008a) have proposed the use of Ant Colony Optimization

(ACO) for improving the classification performance. It was detected that the ACO algorithm gives better accuracy and rule set as compared to See 5.0 Decision Tree process. Liu et al. (2008b) proposed a new method using Particle Swarm Optimization (PSO) for satellite image classification. PSO is capable of finding optimized cut points and has good convergence in the exploration process. Dong and Xiang-bin (2008) worked in the area of image classification of remote sensed data showing the application of Particle Swarm Optimization. The advantage of neighbourhood information is utilized by PSO and is also a robust approach and can be implemented for other kinds of image classification. The detailed comparison is summarized in Table 1.

Analysis from 2009 to 2011

Mohammadzadeh et al. (2009) applied Particle Swarm Optimization (PSO) to a mean calculation system using fuzzy, for obtaining road mean value in each band. Maulik and Saha (2009) worked in the area of image classification and proposed a modified Differential Evolution (DE) using fuzzy clustering technique and also performed statistical significance tests for establishing the superiority. Chen and Leou (2009) have used Particle Swarm Optimization (PSO) for proposing a new IKONOS imagery fusion technique for Panchromatic (PAN) and Multispectral (MS) satellite images. The visual quality and correlation coefficients were better and greater than the other methods. Juneja et al. (2009) implemented and performed relative analysis of Artificial Neural Network (ANN), Rough-Set and Fuzzy-Rough classifier. Paoli et al. (2009) proposed unsupervised classification for hyperspectral images using swarm intelligence which optimized the Bhattacharyya statistical distance between classes and the log likelihood function.

Daamouche and Melgani (2009) implemented a novel classification scheme for hyperspectral images articulating wavelet optimization within Particle Swarm Optimization structure. Chang et al. (2009) worked in the area of hyperspectral image band selection by developing Greedy Modular Eigenspaces (GME) and a novel Parallel Particle Swarm Optimization (PPSO) was offered. The proposed technique improved the computational speed with the help of parallel computing techniques and better reliable solutions as compared to GME. Ding and Chen (2009) has proposed the use of Particle Swarm Optimization (PSO) for improving the Support Vector Machine (SVM) classifier performance for hyperspectral classification. It was observed that the SVM approach has superiority over other traditional classifiers.

Papa et al. (2010) implemented a hybrid Particle Swarm Optimization- Projections Onto Convex Sets (PSO-POCS) algorithm for remote sensing image restoration. Bedawi and Kamel (2010) proposed clustering algorithm using Particle Swarm Optimization (PSO) for segmentation of high

Table 1 Comparative analysis from 2006 to 2008

Reference	Assigned Number	Classifier	Utility	Images Considered	Observation
(Das et al. 2006)	1	PSO - Rough c-means algorithm	Image Segmentation	IRS image of Mumbai and the Nomadic Super Scout II Robot	Fitness function considered by the author is Davies-Bouldin clustering validity index. Even in the existence of noise, correct segments were recognized in the image.
(Zhong et al. 2006)	2	Unsupervised Artificial Immune Classifier	Image Classification	Multispectral Landsat TM image, MODIS image, PHI image	The classification results were compared with Fuzzy k-means, ISODATA, Self-Organizing Map and K-means. Better technique as compared to other four approaches
(Omkar et al. 2007)	3	Ant Colony Optimization and Particle Swarm Optimization	Image Classification	Quick-bird high resolution images	PSO is more efficient and introduces robustness as compared to Ant-miner. Both the techniques produce promising results
(Monteiro and Kosugi 2007)	4	Particle Swarm Optimization –Continuous and Binary	Feature Selection	Hyperspectral Datasets from Soybean Fields.	Superior performance in comparison with Principle Component Analysis
(Senthilnath et al. 2008)	5	Ant Colony Optimization, Maximum Likelihood Classifier and Particle Swarm Optimization and	Crop Coverage Identification Problem	Quick-Birds Multi-Spectral (MSS) image	Comparison of MLPNN trained by PSO, MLC and Ant Miner with respect to crop coverage identification
(Liu et al. 2008a)	6	Ant Colony Optimization	Improve the Classification Performance	Landsat TM image	Improved accuracy over See 5.0 Decision Tree Method and simple rule sets was obtained for Guangzhou area of study
(Liu et al. 2008b)	7	Particle Swarm Optimization-Miner Method	Image Classification	Landsat TM image	Panyu district of Guangzhou was considered. Better accuracy was observed than the See 5.0 Decision Tree Process.
(Dong and Xiang-bin 2008)	8	Particle Swarm Classifying Optimization Algorithm	Images Classification	Remote sensing images- Jidi area	The advantage of neighborhood information is utilized by PSO and is also a robust approach and has better universal compatibility

resolution images and the output was matched with K-means. Linyi and Deren (2010) worked in the areas of image fuzzy classification by proposing the use of Particle Swarm Optimization (PSO) and evaluation was done with Genetic Algorithm (GA) and mean value method. Ari and Aksoy (2010) worked for estimation of likelihood of Gaussian Mixture Models by presenting a Particle Swarm Optimization (PSO) based method. New parameterization for random covariance matrices was also presented. Gupta et al. (2011) proposed an extension of Biogeography Based-Optimization (BBO) for image classification. The migration rate is determined by using Rank based fitness criteria. Very accurate land-cover features were extracted.

Halder et al. (2011) presented supervised and unsupervised Ant based classification and clustering methods for automatic generation of landuse map. Senthilnath et al. (2011b) applied a novel Glowworm Swarm Optimization clustering method in image classification for multispectral satellite images. Goel et al. (2011) presented an innovative Particle Swarm Optimization - Biogeography Based Optimization (PSO-BBO) hybrid approach for classification of multispectral remote images. The method is very efficient and accurate in terms of land cover feature extraction. Arora et al. (2011) has proposed the application of Particle Swarm Optimization (PSO) with morphological operators in the classification of urban features in the satellite images. Samadzadegan and Mahmoudi (2011) proposed the implementation of Firefly Algorithm (FA) and Particle Swarm Optimization (PSO) for band selection. The proposed method outperforms Genetic Algorithm (GA). Bedawi and Kamel (2011) used Particle Swarm Optimization (PSO) for classifying remote data over urban areas. The result shows the significance with high predictive accuracy. Zhang et al. (2011a) presented endmember extraction technique by employing Ant Colony Optimization (ACO) and compared the results to N-FINDR and VCA algorithms. Zhang et al. (2011b) proposed an endmember extraction method by means of Particle Swarm Optimization. Senthilnath et al. (2011a) used Discrete Particle Swarm Optimization in image registration and it turns out to be an efficient technique. Table 2 provides a detailed comparison.

Table 2 Comparative analysis from 2009 to 2011

Reference	Assigned Number	Classifier	Utility	Images Considered	Observation
(Mohammadzadeh et al. 2009	9	Fuzzy-PSO based method	Extraction of Roads	IKONOS satellite images	The effectiveness was reflected by the quality of the extracted road centerline.
(Maulik and Saha 2009)	10	Fuzzy Clustering Technique based on Modified Differential Evolution	Image Classification-l- and cover	IRS satellite and SPOT image of Calcutta and Mumbai IRS satellite image	MoDEFC constantly performed better than classical GAFC, DEFC, SAFC and FCM methods
(Chen and Leou 2009)	11	Particle Swarm Optimization	Imagery Fusion	High-resolution, low-resolution panchromatic and multispectral IKONOS satellite images	The visual quality and correlation coefficients were better and greater than the GIHS, BT, and SDF
(Juneja et al. 2009)	12	Fuzzy-Rough classifiers, Rough Set and Artificial Neural Network and	Image classification	LISS-III image	For image classification, the better choices are Rough Set and Rough-Fuzzy, as compared to ANN
(Paoli et al. 2009)	13	Particle Swarm Optimization	Unsupervised classification	AVIRIS data set	K-means and PCA (first 10 components) + fuzzy C-means were compared with the proposed technique and acquired promising performances in classification accuracy
(Daamouche and Melgani 2009)	14	Wavelet optimization within PSO structure.	Image classification	Hyperspectral AVIRIS data set	Method outperforms the Daubechies wavelets
(Chang et al. 2009)	15	Parallel Particle Swarm Optimization	Band Selection	PacRim-II Project Dataset	Proposed technique improved the computational load along with the achievement of more reliable quality of solution as compared to Greedy Modular Eigenspaces
(Ding and Chen 2009	16	Particle Swarm Optimization	Band Selection	AVIRIS 92AV3C dataset	In comparison to customary classifiers, obtained results confirmed the supremacy of the support vector machine technique and it was suggested that the further enhancements in classification precision can be achieved
(Papa et al. 2010)	17	Hybrid Particle Swarm Optimization-Projections Onto Convex Sets	Image Restoration	CBERS-2B satellite images	Compared with Wiener and Richardson–Lucy algorithms and better results were obtained
(Bedawi and Kamel 2010)	18	Particle Swarm Optimization	Clustering-based image segmentation	High resolution sensing images	Compact clustering results were generated as compared to the K-means algorithm.
(Linyi and Deren 2010)	19	Particle Swarm Optimization	Image Fuzzy Classification	Remote sensing images	Compared with genetic algorithm and mean value method. Higher accuracy was obtained and less training time was obtained.
(Ari and Aksoy 2010)	20	PSO based method for estimation of likelihood of Gaussian Mixture Models.	Unsupervised Classification	AVIRIS image	Better clustering results were obtained, as compared to the Expectation Maximization algorithm for estimating Gaussian Mixture Models.
(Gupta et al. 2011)	21	Extended Biogeography Based Optimization	Image Classification	Remote sensing images	Very accurate land-cover features were extracted.
(Halder et al. 2011)	22	Ant based algorithms	Classification and Clustering	SATIMAGE, IRS-1A, SPOT and IRS	Supervised APC performs better than Multilayer Perceptron and SVM. Unsupervised APC method performs better than MS and KM. Factors used were Rand, Jaccard and two external evaluation indices S_dbw and β

Table 2 (continued)

Reference	Assigned Number	Classifier	Utility	Images Considered	Observation
(Senthilnath et al. 2011b)	23	Glowworm Swarm Optimization	Image Classification	Thematic mapper image of Landsat 7	The performance of the model GSO was better than the MSC unsupervised method and is reasonable for classification.
(Goel et al. 2011)	24	Particle Swarm Optimization - Biogeography Based Optimization (PSO-BBO) hybrid approach	Image Classification	IRS-P6	The method is very efficient and accurate in terms of land cover feature extraction and the method can be extended to other global optimization problem.
(Arora et al. 2011)	25	Particle swarm optimization with morphological operators	Image Feature Classification	Google Earth images of Chandigarh and Saharanpur	The implemented approach has produced satisfactory results and can identify definite features such as vehicles, bridges etc.
(Samadzadegan and Mahmoudi 2011)	26	Particle Swarm Optimization and Firefly algorithm	Feature/Band Selection	AVIRIS data set	With respect to GA and PSO, FA is more superoior
(Bedawi and Kamel 2011)	27	Particle Swarm Optimization	Urban land-cover classification	University of Waterloo Map Library	PSO based rule discovery algorithm can be used in classification and was compared with Neural Network, Parzen window, linear discrimination and K-nn classifiers
(Zhang et al. 2011a)	28	Ant Colony Algorithms	Endmember Extraction	AVIRIS data set	Endemenber extraction problem can be handled correctly by ACO and achieve better results compared to N-FINDR and Vertex Component Analysis and Pure Pixel Index
(Zhang et al. 2011b)	29	Discrete Particle Swarm Optimization	Endmember Extraction	AVIRIS data set	DPSO was better than Vertex Component Analysis and N-FINDR
(Senthilnath et al. 2011a)	30	Discrete Particle Swarm Optimization	Image Registration Technique	QuickBird multispectral Image, QuickBird panchromatic image	Proposed approach is effective for multisensory satellite image registration.

Analysis from 2012 to 2014

Senthilnath et al. (2012) proposed hierarchical clustering procedure by using Glowworm Swarm Optimization (GSO), Niche Particle Swarm Optimization (NPSO) and Mean Shift Clustering (MSC), and it was observed that GSO based approach was robust and more accurate. Wang et al. (2012) presented PSO based approach for post-processing the Sub-Pixel Mapping (SPM) results obtained with the help of Sub-Pixel/Pixel Spatial Attraction Model (SPSAM). Yamaguchi et al. (2012) applied Particle Swarm Optimization (PSO) to the problem of similar image search by using the concept of transfer learning. Banerjee et al. (2012) attempted to solve the image classification land-cover problem by implementing Artificial Bee Colony (ABC) and the comparison was made with other methods. Soliman et al. (2012) worked in the field of image classification by using Support Vector Machine (SVM) and PSO, and evaluation revealed that the usage of RBF kernel function had utmost exactness ratio as well as polynomial kernel. Gao et al. (2012) implemented Ant

Colony Optimization (ACO), for endmember extraction, based on GPU and the results were evaluated. Yavari et al. (2013) presented modified Particle Swarm Optimization (PSO) in identifying the ideal terms for Rational function models (RFM). Bhandari et al. (2014a) employed Cuckoo Search (CS) and Wind Driven Optimization (WDO) along with the use of Kapur's entropy for multilevel thresholding and revealed their efficiency and accuracy. Zarrinpanjeh et al. (2013) proposed ant-agent use in the updation of road map. Satisfactory results with respect to verification, detection and extraction of roads.

Senthilnath et al. (2013) used Genetic Algorithm (GA) and PSO for flood evaluation and river mapping and proved to be an accurate and reliable approach. Bhandari et al. (2014b) presented the application of Artificial Bee Colony algorithm with DWT-SVD for the enhancement in contrast. The proposed technique is better as compared to DCT-SVD, PSO, DWT-SVD, GHE and PSO's modified versions. Ghosh et al. (2013) designed a supervised feature selection method with the help of Self-adaptive Differential Evolution (SADE). The

techniques also used the method of feature ranking. (Zhang et al. (2013) have proposed methods for improving ACO algorithm for extraction of endmember. Bhandari et al. (2014c) presented Cuckoo Search (CS) and DWT-SVD for contrast enhancement and comparison was done in terms of Standard Deviation, MSE, PSNR and Mean. Ghamisi et al. (2014) used fractional-order Darwinian Particle Swarm Optimization for multilevel thresholding. Significant improvement with respect to CPU time and fitness value was observed. Xue et al. (2014) proposed HA-PSO-SVM for image classification which improved the classification performance as compared to other technique. Zhong et al. (2014) proposed adaptive Differential Evolution for endmember extraction. The technique extracted endmember with higher precision. The detailed comparison is presented in Table 3.

Analysis from 2015 to 2018

Bhandari et al. (2015a) worked on finding the optimal multilevel thresholds by the use of modified Artificial Bee Colony (ABC) algorithm with various objective functions. The results are promising and minimized the computational time. Bhandari et al. (2015b) used Tsallis entropy function with Cuckoo Search (CS) algorithm for color image segmentation. The proposed technique selected very effectively and properly the threshold values. Agrawal and Bawane (2015) proposed new multiobjective Particle Swarm Optimization technique for determining different bands and the count of unseen layer nodes. Jayanth et al. (2015a) projected the use of Artificial Bee Colony (ABC) algorithm in satellite classification data and was compared with MLC, ANN and SVM. Ghamisi and Benediktsson (2015) proposed integrated Genetic Algorithm- Particle Swarm Organization (GA-PSO) for feature selection. It was confirmed that the approach automatically selected the most informative feature and was also tested for road detection. Senthilnath et al. (2015) used GA and NPSO for image registration and image clustering and the performance was compared to the conventional methods. Jayanth et al. (2015b) implemented Artificial Bee Colony algorithm for improving the performance of data classification. An enhancement of 5% was achieved in classification precision. Praveena and Singh (2014) presented the use of feed-forward neural networks classifier for image segmentation. Li et al. (2015) presented DPSO based flood inundation mapping- sub-pixel and comparison was done with other methods. Wang et al. (2015) proposed improved online dictionary learning involving Particle Swarm Optimization and the method had superior effect on noise suppression. Iounousse et al. (2015) developed an unsupervised technique on the bases of Probabilistic Neural Network and the accuracy results were compared with other methods. Upadhyay et al. (2010) used Artificial Neural Networks (ANN) for satellite image classification. Yang et al. (2015) developed a multi-

agent system using Artificial Bee Colony (ABC) algorithm for the extraction of endmembers. The method solved the problem in high speed computing and distributive environments. Zhang et al. (2017) analyzed the role of swarm intelligence in the extraction of endmembers from hyperspectral images. Swarm intelligence provides a reliable solution. Kusetogullari et al. (2015) proposed Parallel Binary Particle Swarm Optimization for unsupervised change detection and compared the results with other methods. Suresh and Lal (2016) have implemented CS McCulloch for image segmentation. The results were compared with various techniques and were validated against by various measures. Singh et al. (2016) has presented a comparison of Artificial Bee Colony (ABC), Particle Swarm Optimization (PSO) and Cuckoo Search (CS) along with their hybrids for image enhancement. Sood and Menon (2016) proposed a hybrid Bat Algorithm-Cuckoo Search (BA-CS) approach for the discovery of best path for robotic navigation.

Bhandari et al. (2016) performed a comparative study of various wavelet filters for de-noising satellite images using CS, PSO and ABC. Senthilnath et al. (2016) proposed the implementation of Bat Algorithm (BA) in crop classification problem and compared the result with other intelligent algorithms. Gharbia et al. (2016) proposed image fusion method using Particle Swarm Organization (PSO). The method improved the spatial information and preserved spectral resolution. Tebbi and Haddad (2017) have proposed the use of Support Vector Machine (SVM) classifier in satellite image classification and the classification error was considerably reduced. Muangkote et al. (2016) presented an enhanced algorithm for the segmentation of image using Moth-Flame Optimization. The proposed method was more accurate and effective as compared to other traditional methods. Kusetogullari and Yavariabdi (2016) proposed the implementation of Self-Adaptive Hybrid Particle Swarm Optimization-Genetic Algorithm in obtaining change detection for Landsat multi temporal multispectral images. Sarkar et al. (2016) gave a novel unsupervised classification technique using DE and maximum Rényi entropy methods.

Tien Bui et al. (2017) implemented a novel hybrid Neural Fuzzy optimized by Particle Swarm Optimization (PSO-NF) technique for forest fire susceptibility modeling. Bhandari et al. (2017) proposed the implementation of Beta Differential Evolution (BDE) algorithm in image contrast enhancement. The results with respect to SSIM, EKI, MSE, PSNR and FSIM show the superiority over other traditional methods. Sachdeva et al. (2017) proposed a predictive model for flood susceptibility using PSO and SVM. K et al. (2016) reviewed the enactment of PSO and classifier such as Random forest to satellite images for enhancing and obtaining accurate model of Land Cover Classification. Chang et al. (2017) implemented a novel approach for dimensionality reduction. The Impurity Function band prioritization method uses PSO and

Table 3 Comparative analysis from 2012 to 2014

Reference	Assigned Number	Classifier	Utility	Images Considered	Observation
(Senthilnath et al. 2012)	31	Glowworm Swarm Optimization, Niche Particle Swarm Optimization and Mean Shift Clustering	Hierarchical Clustering-land cover mapping	QuickBird and Landsat 7 thematic mapper	GSO is computationally slow but gave better classification results. The NPSO method is slower and is also less efficient as compared to GSO
(Wang et al. 2012)	32	Particle Swarm Optimization	Post Processing of sub-pixel mapping	Indian Pine Test Site- HSI data and HJ-1 satellite image	The proposed approach provides high accuracy and also reduction in noise
(Yamaguchi et al. 2012)	33	PSO based on Transfer learning concept.	Similar Image Search	MODIS imagery	The proposed PSO method outperforms GA.
(Banerjee et al. 2012)	34	Artificial Bee Colony optimization	Land cover problem by image classification	LISS-III	Results are compared with the Minimum Distance Classifier, BBO, Maximum Likelihood Classifier, Membrane computing and Fuzzy classifier and effectiveness of the proposed method was observed.
(Soliman et al. 2012)	35	SVM and Particle Swarm Optimization	Image Classification.	ASTER satellite imagery	Results presented that the performance of projected approach using the RBF kernel function has got the highest normal overall precision and polynomial kernel function.
(Gao et al. 2012)	36	Ant Colony Optimization Endemember Extraction	Endmember Extraction	AVIRIS data set	It was revealed that the processing speed of ACOEE gets benefit by using GPU.
(Yavari et al. 2013)	37	Particle Swarm Optimization	Rational function models for extracting spatial information	IKONOS-Geo image and SPOT image	PSORFO has superiority with respect to degrees of freedom, computational time and sub-pixel accuracy
(Bhandari et al. 2014a)	38	Wind Driven Optimization and Cuckoo Search Algorithm	Image Segmentation	INSAT and five LANDAT	In multilevel thresholding problem, methods like WDO, ELR-CS and CS finds application because of their accuracy and efficiency
(Zarrinpanjeh et al. 2013)	39	Ant-Agent	Road Map Updation	GeoEye-1 pan-sharpen imagery	Observed results are satisfactory with respect to updated map production along with detection, verification and extraction of roads
(Senthilnath et al. 2013)	40	Genetic Algorithm and Particle Swarm Optimization	Segmentation and clustering	MODIS satellite images	The proposed algorithms offer accuracy and reliability in the extraction of water covered region
(Bhandari et al. 2014b)	41	Artificial Bee Colony algorithm using DWT-SVD	Augmentation in terms of contrast	INSAT and LANDSAT images	Better technique as compared to DCT-SVD, PSO, PSO's modified versions, General Histogram Equalization and DWT-SVD with respect to PSNR, mean, variance and MSE
(Ghosh et al. 2013)	42	Differential Evolution - Self-adaptive	Supervised Feature Selection	AVIRIS, Landsat, NASA Earth Observing 1 satellite	Results were compared with GA, ACO, ANTDE, DE and significant improvement was obtained in classification precision and Kappa coefficient
(Zhang et al. 2013)	43	ACO algorithm	Endmember Extraction	AVIRIS USGS library	Presented improvement over AOCEE algorithm issues
(Bhandari et al. 2014c)	44	Cuckoo Algorithm and DWT-SVD	Contrast Enhancement	INSAT images and LANDSAT images	Proposed method is superior with respect to Standard Deviation, MSE, Mean and PSNR
(Ghamisi et al. 2014)	45	Fractional-order Darwinian Particle Swarm Optimization	Image Segmentation	Multispectral Worldview Image and Hyperspectral ROSIS Image	Experimental results compared the proposed approach with the classical PSO and DPSO. A more robust method is presented along with higher prospective for locating the ideal thresholds set within less computational time
(Xue et al. 2014)	46	Support Vector Machine, Harmonic Analysis and	Hyperspectral Image Classification	AVIRIS and ROSIS data set	Proposed method improved performance, and a good balance was obtained between computational time and accuracy

Table 3 (continued)

Reference	Assigned Number	Classifier	Utility	Images Considered	Observation
(Zhong et al. 2014)	47	Particle Swarm Optimization and Adaptive Differential Evolution	Endmember Extraction	AVIRIS data set	ADEE adaptively presented better Endmember Extraction results with lower RMSE and a upper SAM and SAMSFF accuracy

Gravitational Search Algorithm for reducing the hyperspectral bands. Golovko et al. (2017) has proposed the use of convolutional neural network in low-quality satellite images for detection of solar photovoltaic panels. Google satellite images were used. Azarang and Ghassemian (2017) proposed a novel approach of image fusion for applications in remote sensing using particle swarm optimization for weight injections. WorldView-3 and QuickBird data set are considered for assessment. Kumar et al. (2016) showed implementation of PSO and K-means to cluster satellite images. The approach produced more condensed and augmented clusters than the K-means method alone. Gaba et al. (2017) developed a statistical model, which helps in learning and classifying object in hyperspectral images using combination of GSA and FODPSO. Alizadeh Naeini et al. (2018) considered satellite images of very high spatial resolution and proposed a novel object based feature selection method. Singh et al. (2017) has proposed the use of Moth Flame Optimization for image classification. The detailed comparison is presented in Table 4.

Discussion

It has been observed that different techniques have been applied in different sectors of satellite image processing. In segmentation of satellite images, Particle Swarm Optimization is most widely used followed by Cuckoo Search, Artificial Bee Colony, Differential Evolution, Wind Driven Optimization, Genetic Algorithm and Moth-Flame Optimization. Classification is covered by Particle Swarm Optimization, Unsupervised Artificial Immune Classifier, Ant Colony Optimization, Differential Evolution, Fuzzy-Rough Set, Biogeography Based Optimization, Glowworm Swarm Optimization, Artificial Bee Colony, Neural Network/ Convolutional NN, Bat Algorithm, Support Vector Machine and Moth-Flame Optimization. Feature/ Band Selection use Firefly Algorithm, Differential Evolution, Genetic Algorithm and Particle Swarm Optimization. Extraction of roads or map updation or cross-country path-finding is achieved by Cuckoo Search, Bat Algorithm, Ant Colony Algorithm and Particle Swarm Optimization. Particle Swarm Optimization discovers its use in Image Fusion or Similar Image, Image Restoration, Sub-Pixel Mapping, Rational Function Models and Online

Dictionary Learning. For Endmember Extraction, Differential Evolution, Particle Swarm Optimization, Ant Colony Optimization and Artificial Bee Colony are implemented. Genetic Algorithm and Particle Swarm Optimization are implemented for Image Registration and Change Detection. Fire and Flood susceptibility model is implemented by Particle Swarm Optimization and Support Vector machine. Contrast/ Image Enhancement are done by using Differential Evolution, Particle Swarm Optimization, Cuckoo Search and Artificial Bee Colony. Clustering uses Particle Swarm Optimization and Glowworm Swarm Optimization, whereas De-noising is using Artificial Bee Colony, Particle Swarm Optimization and Cuckoo Search. The brief analysis of various artificial intelligent techniques with their application areas in satellite image optimization is depicted in Table 5. The statistical analysis is carried out by pie and column charts as shown in Fig. 1. From Fig. 1a, this is observed that in image segmentation, Particle swarm optimization has the maximum applications (31%) followed by Cuckoo Search (23%), Artificial bee colony algorithm (15%), wind driven optimization, genetic algorithm and moth-flame optimization, differential evolution (7%). From Fig. 1b, PSO has 43% applications with respect to image classification, followed by ant colony optimization, artificial bee colony algorithm and neural networks. Similarly, Fig. 1c, d, e, f, g and h represents the applications of various swarm-based techniques in different areas of image processing. [U13].

Conclusion

Remote sensing provides coverage of large areas to collect precise information in various applications such as agricultural fields, location of floods, forest fires, landscape and regional planning etc. The quality of satellite images is weather dependent and size of data base is huge. This makes the image processing task highly time consuming and cumbersome. Thus, there is a need for application of a suitable artificial intelligence technique to improve the image quality with smaller processing time. In this study, various artificial intelligence techniques such as PSO, ACO, ABC, bat algorithm, GA etc. have been analysed for optimization of satellite image

Table 4 Comparative analysis from 2015 to 2018

Reference	Assigned Number	Classifier	Utility	Images Considered	Observation
(Bhandari et al. 2015a)	48	Modified Artificial Bee Colony	Image Segmentation	Satellite images of Pléiades and NASA Earth Observatory	Comparison done with Genetic algorithm, ABC and PSO, and the objective functions considered were Otsu, Tsallis and Kapur's. The proposed method outperforms other methods with a limitation of high complexity.
(Bhandari et al. 2015b)	49	Cuckoo Search Algorithm	Color Image Segmentation	NASA Earth Observatory images	Compared Artificial Bee Colony, Differential Evolution, Particle Swarm Optimization, Wind Driven Optimization, and Cuckoo Search. Robustness order CS > DE > ABC > PSO > WDO. Good to bad run time, DE < ABC < CS < PSO < WDO and CS approach had very less control parameters
(Agrawal and Bawane 2015)	50	Multiobjective PSO	Image Classification	Landsat satellite images of Washington, DC	Improved the classification accuracy and computation reduction in the classification stage of neural classifier
(Jayanth et al. 2015a)	51	Artificial Bee Colony	Image Classification	Multispectral data of LISS-IV IRS P6 and panchromatic image of IRS P5	The comparison is made between ABC, Maximum Likelihood Classifier, ANN and SVM methods and ABC was found out to be most effective in classification and the limitation observed is with respect to larger number of classes in feature space. No initialization of parameters is required and fast feature selection method
(Ghamisi and Benediktsson 2015)	52	Hybrid Genetic Algorithm and Particle Swarm Optimization	Feature Selection Technique	AVIRIS Indian Pines and RGB Toronto Roads data set	The proposed technique was able to locate informative bands w.r.t. classification accuracies in a tolerable CPU time. The technique can also be used in road detection and as fitness function SVM is considered. The technique can handle high-dimensional data through the use of limited quantity of training examples in comparison to various selection methods.
(Senthilnath et al. 2015)	53	Genetic Algorithm and Niche Particle Swarm Optimization	Image Registration and Image Clustering	LISS-III and SAR	Performance of SIFT-GA was better than SIFT-ANN for Image registration and also the performance of NPSO is superior as compared to ISODATA
(Jayanth et al. 2015b)	54	ABC algorithm	Image classification	L1SS-IV sensor of IRS P-6 and PAN image of IRS P-5 satellites	The comparison was made between the MLC and ABC method by making use of OCA and kappa statistics. ABC technique is more compelling for the classification. No initialization of parameters is required and fast feature selection method
(Praveena and Singh 2014)	55	Hybrid ABC-FCM algorithm	Land-cover mapping-image segmentation	Satellite images	The performance was compared with Moving KFCM, ABC algorithm, and ABC-GA algorithm and it outperformed the other methods.
(Li et al. 2015	56	Discrete Particle Swarm Optimization	Flood inundation mapping through Sub-pixel	Landsat ETM + images	Compared the proposed approach with traditional BPGDX-SFIM, BPBR-SFIM, DESDSFIM and SAM-SFIM methods. The approach dependably accomplished more precise sub-pixel mapping conclusions as far as visual and quantitative assessments
(Wang et al. 2015)	57	Particle Swarm Optimization	Online dictionary learning	Landsat-8 images and HJ-1-A	The proposed method improved the performance of ODL algorithms with respect to accuracy and has an improved effect on noise suppression

Table 4 (continued)

Reference	Assigned Number	Classifier	Utility	Images Considered	Observation
(Iounousse et al. 2015)	58	Neural Network-Probabilistic	Land use classification	LANDSAT and SPOT data set	Better performance and accurate classification with around 3.44% of error as compared to K-means and fuzzy c-means
(Upadhyay et al. 2010)	59	Artificial Neural Network	Image classification	IRS P-6 LISS-III	The results expressed that if the number of neurons are increased in ANN then the accuracy downfall is observed in classification based on neural networks.
(Yang et al. 2015)	60	Artificial Bee Colony	Endmember Extraction	AVIRIS data set	ABC algorithm implemented in MAS is capable of reducing the computing time by the introduction of additional distributed computing nodes which share the calculating cost efficiently
(Zhang et al. 2017)	61	Artificial Bee Colony	Endmembers Extraction from hyperspectral imagery.	AVIRIS data set	For endmember extraction, ACO, DPSO, and ABC have been introduced. Outstanding global optimization search ability was reflected by the good results.
(Kusetogullari et al. 2015)	62	Parallel Binary Particle Swarm Optimization	Unsupervised change-detection	USGS-Landsat satellite and Earth Resources Observation and Science Centre	Performance comparison was done with EM-, GA- MRF- and PCA- based methods. For showing the superiority, real world and semi-synthetic data sets were used for quantitative and qualitative tests
(Suresh and Lal 2016)	63	Cuckoo Search	Image segmentation	Pleiades Satellite Imagery	The performance with respect to Otsu's method, Kapur and Tsallis entropy was compared with ABC, Darwinian Particle Swarm Optimization, PSO, CS and CS Mantegna algorithm with and validation was achieved by the means of PSNR, MSE, FSIM and CPU running time
(Singh et al. 2016)	64	CS, PSO, ABC and hybrids	Image Enhancement	Satellite Images	Analytical review comparing all the techniques with respect to mean, MSE, PSNR and variance
(Sood and Menon 2016)	65	Bat-Cuckoo Search	Cross-country Path finding	Satellite images	The proposed hybrid approach can rapidly plan an optimized path in complex surroundings.
(Bhandari et al. 2016)	66	Cuckoo Search, Artificial Bee Colony and Particle Swarm Optimization	De-noising of satellite images	INSAT and LANDSAT	For noise suppression and edge preservation, Meyer wavelet filter-based CS de-noising technique was compared to other wavelet-filter methods and better performance was obtained
(Senthilnath et al. 2016)	67	Bat Algorithm	Image Classification	University of California, Irvine (UCI) repository	The performance of the algorithm was equated with existing hybrid BA with K-means, Particle Swarm Optimization and Genetic Algorithm and it was concluded that the BA can be functional to crop classification
(Gharbia et al. 2016)	68	Particle Swarm Optimization	Image Fusion	ASTER, SPOT, MODIS, ETM+ and MSS satellite	Comparison was done with the PCA and DCT. Improvement in fusion quality, spatial information and preserved spectral resolution. PSO based adaptive fusion rule can be implemented for improving performance of any image fusion technique
(Tebbi and Haddad 2017)	69	SVM	Cloud Classification	Meteosat Second Generation (MSG0 satellite images.	Elimination of non-rainy clouds by the proposed method improves the precipitation estimation
(Muangkote et al. 2016)	70	Moth-Flame Optimization Algorithm	Image Segmentation	Satellite Images	The method was more accurate and effective as compared to PSO, ABC, DE, MFO and GA.
	71				

Table 4 (continued)

Reference	Assigned Number	Classifier	Utility	Images Considered	Observation
(Kusetogullari and Yavariabdi 2016)		Hybrid Particle Swarm Optimization-Genetic Algorithm (Self-Adaptive)	Unsupervised Satellite Change Detection Method	Multi Spectral Multi Temporal Landsat Images	It augments the change detection and is able to discover the ultimate change detection mask.
(Sarkar et al. 2016)	72	Differential Evolution	Image segmentation	ROSIS and AVRIS sensors	DE based segmentation method by employing Rényi entropy is proposed. Performance is compared with ABC, CS, GA and PSO.
(Tien Bui et al. 2017)	73	Neural Fuzzy optimized by Particle Swarm algorithm	Forest fire susceptibility modeling.	Landsat-8 Operational Land Imagery	Proposed model outperforms the RF and the SVM model. Early warning system could also be established for tropical forest fires.
(Bhandari et al. 2017)	74	Beta Differential Evolution (BDE) Algorithm	Image Enhancement	Colored Satellite Images	The satisfactory results with respect to SSIM, EKI, MSE, PSNR and FSIM were compared with ABC, PSO, DE algorithms and modified ABC
(Sachdeva et al. 2017)	75	PSO and SVM	Flood Susceptibility	Remote Sensing and Geographic Information Systems	The model was compared with NN and RF, and it outperformed with an accuracy of 96.55%. District of Chamoli, Uttarakhand was considered.
(K et al., 2016)	76	PSO and classifier such as Random forest	Enhancing and obtaining accurate model of Land Cover Classification.	Satellite Images	Literature survey covering the performance of PSO and RF
(Chang et al. 2017)	77	Gravitational Search Algorithms and Particle Swarm Optimization	Dimensionality Reduction	Hyperspectral data sets	Hyperspectral data dimensionality is reduced by the proposed technique and obtained an improved classification accuracy as compared to other methods
(Golovko et al. 2017)	78	Convolutional Neural Network	Classification-detection of solar photovoltaic panels	Google Satellite low quality images	Obtained detection with high accuracy and low rate incorrect classifications
(Azarang and Ghassemian 2017)	79	Particle Swarm Optimization	Image Fusion	WorldView-3 and QuickBird	PSO is an effective method for image fusion
(Kumar et al. 2016)	80	K Means Algorithm and PSO	Satellite Image Clustering	Landsat 8 OLI	The approach produced more condensed and augmented clusters than the K-means method alone.
(Gaba et al. 2017)	81	SVM, Finite Order Darwinian PSO and Gravitational Search Algorithm,	Image Classification	Hyperspectral data set of Indian Pine	Better accuracy is obtained as compared to FODPSO
(Alizadeh Naeini et al. 2018)	82	Particle Swarm Optimization	Feature selection (Object Based-OBF)	WorldView-2 sensor	PSO approach had better performance over GA, ABC and honey-bee mating
(Singh et al. 2017)	83	Moth-Flame Optimization	Image Classification	LISS-III sensor	Proposed method was better than fuzzy classification, Membrane Computing, BBO, minimum Distance Classifier and MLC

data. The detailed analysis revealed that Ant Colony Optimization finds its applications in the field of classification and extraction of endmember from hyperspectral data. Bat Algorithm, Artificial bee colony and Neural Networks work for classification and Cuckoo search algorithm deals with segmentation. This has been observed that Particle Swarm Optimization is the most commonly utilized strategy with respect to image classification, segmentation, feature/band selection, enhancement, image fusion, registration and restoration. The major areas covered by various techniques are land use-land cover mapping, crop classification, forest fire susceptibility and flood assessment. In future, these techniques need to be applied to other significant areas such as natural calamities forecast, suggestions for quick and efficient relief

Table 5 Image optimization techniques for satellite images

Application area for image optimization	Techniques and hybrid techniques	Assigned number from Tables 1,2,3 & 4
Segmentation	Particle swarm optimization	1, 18, 40, 45
	Differential evolution	73
	Artificial bee colony	48, 55,
	Cuckoo search	38, 49, 63
	Wind driven optimization	38
	Genetic algorithm	40
	Moth-flame optimization	70
Classification	Particle swarm optimization	3, 5, 7, 8, 14, 19, 20, 24, 25, 27, 35, 46, 50, 77, 82
	Unsupervised artificial immune classifier	2
	Ant colony optimization	3, 5, 6, 22
	Differential evolution	10
	Fuzzy-rough set	12
	Biogeography based optimization	21, 24
	Glowworm swarm optimization	23
	Artificial bee colony	34, 51, 54
	Neural Network/ Convolutional NN	58, 59, 79
	Bat Algorithm	67, 72
	Support Vector Machine	69
	Moth-flame optimization	84
Feature / Band Selection	Firefly algorithm	26
	Differential evolution	42
	Particle swarm optimization	4, 26, 52, 83
	Genetic algorithm	52
Extraction of Roads/ Road Map Updation/ Cross Country Path Finding	Particle swarm optimization	9
	Ant Colony Optimization	39
	Cuckoo search	65
	Bat algorithm	65
Image Fusion/ Similar Image	Particle swarm optimization	11, 33, 68, 80
Image Restoration	Particle swarm optimization	17

Table 5 (continued)

Application area for image optimization	Techniques and hybrid techniques	Assigned number from Tables 1,2,3 & 4
Endmember Extraction	Ant colony optimization	28, 43, 36
	Differential evolution	47
	Particle swarm optimization	29
	Artificial Bee Colony	60, 61
Image Registration	Particle swarm optimization	30, 53
	Genetic algorithm	53
Contrast / Image Enhancement	Particle swarm optimization	64
	Differential evolution	75
	Cuckoo search	44, 64
	Artificial bee colony	41, 64
Clustering	Particle swarm optimization	31, 89
	Glowworm swarm optimization	31
Sub-Pixel Mapping	Particle swarm optimization	32, 56
Rational Function Models	Particle swarm optimization	37
Online Dictionary Learning	Particle swarm optimization	57
Change Detection	Particle swarm optimization	62, 71
	Genetic algorithm	71
De-Noising	Artificial bee colony algorithm	66
	Cuckoo search	66
	Particle swarm optimization	66
Fire/ Flood Susceptibility	Particle swarm optimization	74, 76
	Support vector machine	76

operations and estimation of natural resources. The use of hybrid techniques by combining one or more artificial intelligence for remote sensed image optimization can also be performed. [U15].

References

Agrawal RK, Bawane NG (2015) Multiobjective PSO based adaption of neural network topology for pixel classification in satellite imagery.
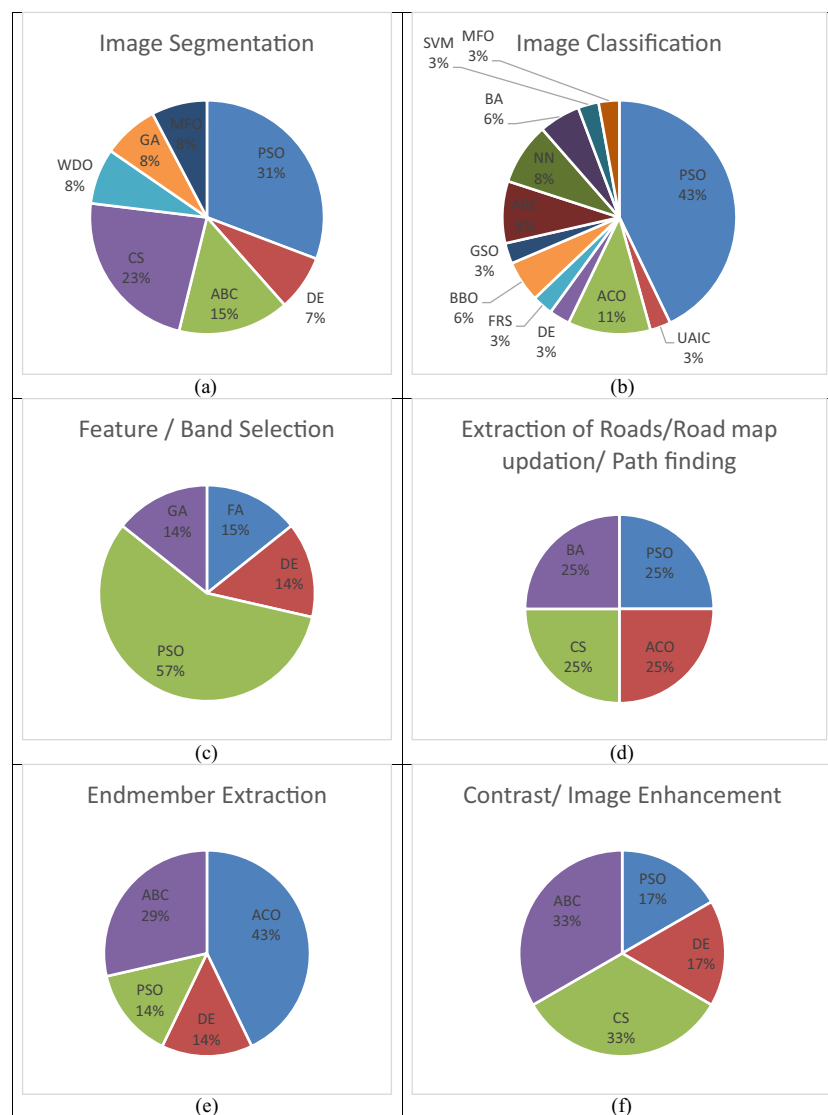
Fig. 1 Application areas of Swarm Intelligence Techniques in various fields of Image processing: (**a**) Segmentation, (**b**) Classification, (**c**) feature/band selection, (**d**) extraction of roads, (**e**) endmember extraction, (**f**) contrast/image enhancement, (**g**) De-noising, fire/flood susceptibility, clustering, registration and change detection, (**h**) other areas [U14]

Appl Soft Comput J 28:217–225. https://doi.org/10.1016/j.asoc.2014.11.052

Alizadeh Naeini A, Babadi M, Mirzadeh SMJ, Amini S (2018) Particle swarm optimization for object-based feature selection of VHSR satellite images. IEEE Geosci Remote Sens Lett 15:379–383. https://doi.org/10.1109/LGRS.2017.2789194

Alpaydin, E. (2009). Introduction to machine learning. MIT press

Ari, Ç., & Aksoy, S., 2010, July. Unsupervised classification of remotely sensed images using gaussian mixture models and particle swarm optimization. In Geoscience and remote sensing symposium (IGARSS), 2010 IEEE international (pp. 1859–1862). IEEE

Arora, P., Mishra, A., Panchal, V.K., 2011. Swarm intelligence for feature identification in natural terrain environment. Proc. - 2011 Int. Conf. Comput. Intell. Commun. Syst. CICN 2011 722–726. https://doi.org/10.1109/CICN.2011.157

Azarang A, Ghassemian H (2017) An adaptive multispectral image fusion using particle swarm optimization. 2017 25th Iran. Conf. Electr. Eng. ICEE 2017:1708–1712. https://doi.org/10.1109/IranianCEE.2017.7985325

Banerjee S, Bharadwaj A, Gupta D, Panchal VK (2012) Remote sensing image classification using artificial bee Colony algorithm 2012. Int. Jour. of Computer Science and Informatics2.3 pp:67–72

Bedawi SM, Kamel MS (2010) Segmentation of very high resoultion remote sensing imagery of urban areas using particle swarm optimization. 2010. Int Conf Image Analysis and Recognition:81–88

Bedawi, S.M., Kamel, M.S., 2011. Urban land-cover classification based on swarm intelligence from high resolution remote sensing imagery. 2011 Int. Conf. Remote Sensing, Environ. Transp. Eng. 5617–5620. 10.1109/RSETE.2011.5965626

Bhandari AK, Singh VK, Kumar A, Singh GK (2014a) Cuckoo search algorithm and wind driven optimization based study of satellite image segmentation for multilevel thresholding using Kapur's entropy. Expert Syst Appl 41:3538–3560. https://doi.org/10.1016/j.eswa.2013.10.059

Bhandari AK, Soni V, Kumar A, Singh GK (2014b) Artificial bee Colony-based satellite image contrast and brightness enhancement technique using DWT-SVD. Int J Remote Sens 35:1601–1624. https://doi.org/10.1080/01431161.2013.876518

(g)

(h)
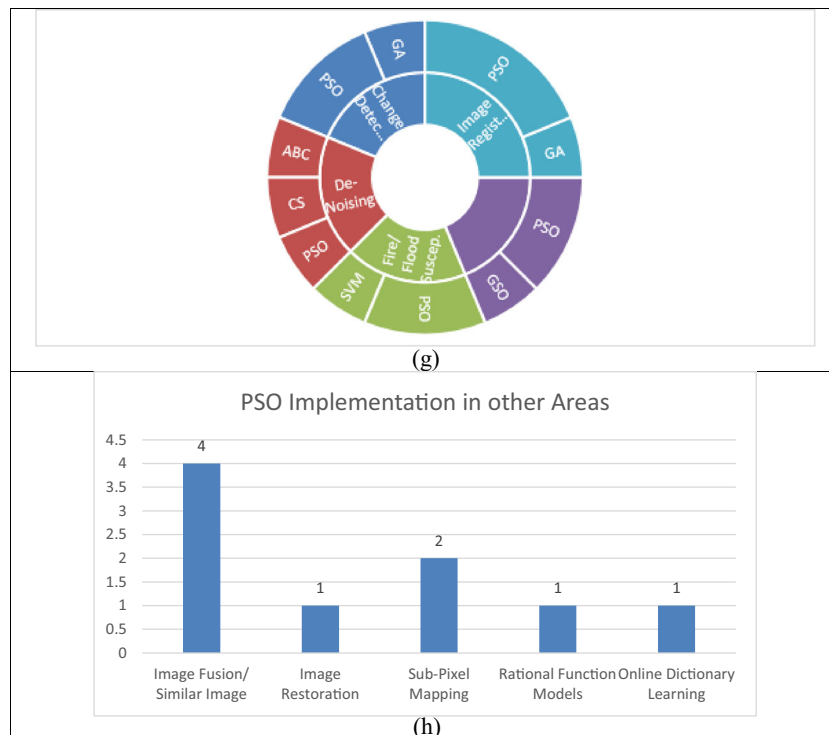
Fig. 1 continued.

Bhandari AK, Soni V, Kumar A, Singh GK (2014c) Cuckoo search algorithm based satellite image contrast and brightness enhancement using DWT-SVD. ISA Trans 53:1286–1296. https://doi.org/10.1016/j.isatra.2014.04.007

Bhandari AK, Kumar A, Singh GK (2015a) Modified artificial bee colony based computationally efficient multilevel thresholding for satellite image segmentation using Kapur's, Otsu and Tsallis functions. Expert Syst Appl 42:1573–1601. https://doi.org/10.1016/j.eswa.2014.09.049

Bhandari AK, Kumar A, Singh GK (2015b) Tsallis entropy based multilevel thresholding for colored satellite image segmentation using evolutionary algorithms. Expert Syst Appl 42:8707–8730. https://doi.org/10.1016/j.eswa.2015.07.025

Bhandari AK, Kumar A, Singh GK, Soni V (2016) Performance study of evolutionary algorithm for different wavelet filters for satellite image denoising using sub-band adaptive threshold. J Exp Theor Artif in Ligence 28:71–95. https://doi.org/10.1080/0952813X.2015.1020518

Bhandari AK, Kumar A, Chaudhary S, Singh GK (2017) A new beta differential evolution algorithm for edge preserved colored satellite image enhancement. In: Multidimensional systems and signal processing, vol 28. Springer US, pp 495–527. https://doi.org/10.1007/s11045-015-0353-4

Campbell JB, Wynne RH (2011) Introduction to remote sensing. Guilford Press

Chang, Y. L., Fang, J. P., Benediktsson, J. A., Chang, L., Ren, H., & Chen, K. S. (2009, July). Band selection for hyperspectral images based on parallel particle swarm optimization schemes. In Geoscience and Remote Sensing Symposium, 2009 IEEE International, IGARSS 2009 (Vol. 5, pp. V-84). IEEE

Chang YL, Chang L, Xu MX, Chu C (2017) Impurity function band prioritization based on particle swarm optimization and gravitational search algorithm for hyperspectral images. IEEE Geosci Remote Sens (IGARSS):825–829

Chen H, Leou J (2009) A new Ikonos imagery fusion approach using particle swarm optimization 2009. Int. Conf. Multimedia and expo. NY, pp 85–88

Daamouche A, Melgani F (2009) Swarm intelligence approach to wavelet design for hyperspectral image classification. IEEE Geosci Remote Sens Lett 6:825–829. https://doi.org/10.1109/LGRS.2009.2026191

Das, S., Abraham, A., Sarkar, S., 2006. A Hybrid Rough Set–Particle Swarm Algorithm for Image Pixel Classification. 2006 Sixth Int. Conf. Hybrid Intell. Syst. 26–26. https://doi.org/10.1109/HIS.2006.264909

Demirel H, Ozcinar C, Anbarjafari G (2009) Satellite image contrast enhancement using discrete wavelet transform and singular value decomposition. IEEE Geosci Remote Sens Lett 7(2):333–337

Ding S, Chen L (2009, December) Classification of hyperspectral remote sensing images with support vector machines and particle swarm optimization. In: Int Conf on information engineering and computer science, 2009. ICIECS 2009. IEEE, pp 1–5

Dong W, Xiang-bin W (2008) Particle swarm intelligence classification algorithm for remote sensing images. 2008 IEEE Pacific-Asia work. Comput Intell Ind Appl:566–570. https://doi.org/10.1109/PACIIA.2008.26

Estes JE, Sailer C, Tinney LR (1986) Applications of artificial intelligence techniques to remote sensing. Prof Geogr 38(2):133–141

Gaba P, Singh S, Paliwal KK (2017) Feature selection approach of hyperspectral image using GSA-FODPSO-SVM 2017. Int Conf Comp Comm and Automation (ICCCA), Greater Noida:1070–1075

Gajdhane VA, Deshpande LM (2014) Detection of lung cancer stages on CT scan images by using various image processing techniques. IOSR J Computer Engineering (IOSR-JCE) 16(5):28–35

Gao J, Gao L, Sun X, Wu Y, Zhang B (2012) GPU implementation of ant colony optimization algorithm for endmember extraction from hyperspectral image 2012. Fourth Workshop Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), Shanghai:1–4

Ghamisi P, Benediktsson JA (2015) Feature selection based on hybridization of genetic algorithm and particle swarm optimization. Geosci

Remote Sens Lett IEEE 12:309–313. https://doi.org/10.1109/LGRS.2014.2337320

Ghamisi P, Couceiro MS, Martins FML, Benediktsson JA (2014) Multilevel image segmentation based on fractional-order darwinian particle swarm optimization. IEEE Trans Geosci Remote Sens 52:2382–2394. https://doi.org/10.1109/TGRS.2013.2260552

Gharbia R, El Baz AH, Hassanien AE (2016) An adaptive image fusion rule for remote sensing images based on the particle swarm optimization. 2016 Int. Conf. Comput Commun Autom:1080–1085. https://doi.org/10.1109/CCAA.2016.7813903

Ghosh A, Datta A, Ghosh S (2013) Self-adaptive differential evolution for feature selection in hyperspectral image data. Appl Soft Comput J 13:1969–1977. https://doi.org/10.1016/j.asoc.2012.11.042

Goel S, Sharma A, Goel A (2011) Development of swarm based hybrid algorithm for identification of natural terrain features. Proc - 2011 Int Conf Comput Intell Commun Syst CICN (2011):293–296. https://doi.org/10.1109/CICN.2011.61

Golovko V, Bezobrazov S, Kroshchanka A, Sachenko A, Komar M, Karachka A (2017) Convolutional neural network based solar photovoltaic panel detection in satellite photos 2017. Ninth Int. Conf. Intelligent Data Acquisition and Advanced Computing Systems, Technology and Applications (IDAACS). Bucharest, pp 14–19

Gupta, S., Arora, A., Panchal, V. K., & Goel, S., 2011. Extended biogeography based optimization for natural terrain feature classification from satellite remote sensing images. *Int. Conf. Contemporary Computing* (pp. 262–269). Springer, Berlin, Heidelberg

Halder A, Ghosh A, Ghosh S (2011) Supervised and unsupervised landuse map generation from remotely sensed images using ant based systems. Appl Soft Comput J 11:5770–5781. https://doi.org/10.1016/j.asoc.2011.02.030

Iounousse J, Er-Raki S, El Motassadeq A, Chehouani H (2015) Using an unsupervised approach of probabilistic neural network (PNN) for land use classification from multitemporal satellite images. Appl Soft Comput J 30:1–13. https://doi.org/10.1016/j.asoc.2015.01.037

Jain AK (1989) Fundamentals of digital image processing. Prentice Hall, Englewood Cliffs, NJ

Jayanth J, Koliwad S, Ashok Kumar T (2015a) Classification of remote sensed data using artificial bee Colony algorithm. Egypt J Remote Sens Sp Sci 18:119–126. https://doi.org/10.1016/j.ejrs.2015.03.001

Jayanth J, Kumar, A., Koliwad, S., Krishnashastry, S., 2015b. Artificial bee colony algorithm for classification of remote sensed data. 2015 Int. Conf. Ind. Instrum. Control 1512–1517. 10.1109/IIC.2015.7150989

Juneja, M., Walia, E., Sandhu, P. S., & Mohana, R., 2009, July. Implementation and comparative analysis of rough set, artificial neural network (ann) and fuzzy-rough classifiers for satellite image classification. In Intelligent Agent & Multi-Agent Systems, 2009. IAMA 2009. International Conference on (pp. 1–6). IEEE

Kumar G, Sarth PP, Ranjan P, Kumar S (2016) Satellite image clustering and optimization using. IEEE Int Conf Power Electron Intell Control Energy Syst:1–4. https://doi.org/10.1109/ICPEICES.2016.7853627

Kusetogullari H, Yavariabdi A (2016) Self-adaptive hybrid PSO-GA method for change detection under varying contrast conditions in satellite images. In: Proc. 2016 SAI Comput. Conf. SAI 2016, pp 361–368. https://doi.org/10.1109/SAI.2016.7556007

Kusetogullari H, Yavariabdi A, Celik T (2015) Unsupervised change detection in multitemporal multispectral satellite images using parallel particle swarm optimization 2015. In: Jour. Of selected topics in applied earth observations and remote Sesing, pp 2151–2164

Lee ER, Kim PK, Kim HJ (1994, November) Automatic recognition of a car license plate using color image processing. In: Proceedings of 1st international conference on image processing, vol 2. IEEE, pp 301–305

Li L, Chen Y, Yu X, Liu R, Huang C (2015) Sub-pixel flood inundation mapping from multispectral remotely sensed images based on

discrete particle swarm optimization. ISPRS J Photogramm Remote Sens 101:10–21. https://doi.org/10.1016/j.isprsjprs.2014.11.006

Linyi, L., Deren, L., 2010. Fuzzy Classification of Remote Sensing Images Based on Particle Swarm Optimization. 2010 Int. Conf. Electr. Control Eng. 1039–1042. 10.1109/iCECE.2010.263

Liu X, Li X, Liu L, He J, Ai B (2008a) An innovative method to classify remote-sensing images using ant Colony optimization. IEEE Trans Geosci Remote Sens 46:4198–4208. https://doi.org/10.1109/TGRS.2008.2001754

Liu X, Li X, Peng X, Li H, He J (2008b) Swarm intelligence for classification of remote sensing data. Sci China Ser D Earth Sci 51:79–87. https://doi.org/10.1007/s11430-007-0133-6

Maulik U, Saha I (2009) Modified differential evolution based fuzzy clustering for pixel classification in remote sensing imagery. Pattern Recogn 42:2135–2149. https://doi.org/10.1016/j.patcog.2009.01.011

McCorduch P (1979) Machines who think. Freeman Inc, San Francisco

McKeown DM (1987) The role of artificial intelligence in the integration of remotely sensed data with geographic information systems. IEEE Trans Geosci Remote Sens GE-25(3):330–348

Mohammadzadeh A, Zoej MJV, Tavakoli A (2009) Automatic main road extraction from high resolution satellite imageries by means of particle swarm optimization applied to a fuzzy-based mean calculation approach. J Indian Soc Remote Sens 37:173–184. https://doi.org/10.1007/s12524-009-0021

Monteiro, S.T., Kosugi, Y., 2007. A particle swarm optimization-based approach for hyperspectral band selection. 2007 IEEE Congr. Evol. Comput. CEC 2007 3335–3340. https://doi.org/10.1109/CEC.2007.4424902

Muangkote, N., Sunat, K., Chiewchanwattana, S., 2016. Multilevel thresholding for satellite image segmentation with moth-flame based optimization. 2016 13th Int. Jt. Conf. Comput Sci Softw Eng JCSSE 2016 0–5. https://doi.org/10.1109/JCSSE.2016.7748919

Omkar SN, Mudigere D, Muley D (2007) Urban satellite image classification using biologically inspired techniques 2007. IEEE Int. Symp. ISIE

Paoli, A., Melgani, F., Pasolli, E., 2009. Swarm intelligence for unsupervised classification of hyperspectral images. Geosci. Remote Sens. Symp. IEEE Int. 2009 5, V-96-V-99. 10.1109/IGARSS.2009.5417723

Papa JP, Fonseca LMG, De Carvalho LAS (2010) Projections onto convex sets through particle swarm optimization and its application for remote sensing image restoration. Pattern Recogn Lett 31:1876–1886. https://doi.org/10.1016/j.patrec.2010.02.012

Praveena, S., Singh, S.P., 2014. Hybrid clustering algorithm and feedforward neural network for satellite image classification 2015 Int Conf Industrial Instrumentation and Control (ICIC), Pune pp. 1378–1383

Sachdeva S, Bhatia T, Verma AK (2017) Flood susceptibility mapping using GIS-based support vector machine and particle swarm optimization: a case study in Uttarakhand (India). 8th Int. Conf. Comput Commun Netw Technol ICCCNT 2017. https://doi.org/10.1109/ICCCNT.2017.8204182

Samadzadegan, F., Mahmoudi, F.T., 2011. Optimum band selection in hyperspectral imagery using swarm intelligence optimization algorithms. 2011 Int. Conf. Image Inf. Process. 1–6. 10.1109/ICIIP.2011.6108925

Sarkar S, Das S, Chaudhuri SS (2016) Hyper-spectral image segmentation using Rényi entropy based multi-level thresholding aided with differential evolution. Expert Syst Appl 50:120–129. https://doi.org/10.1016/j.eswa.2015.11.016

Senthilnath J, Omkar SN, Mudigere D, Kumar MM (2008) Crop Classification using biologically-inspired techniques with high resolution satellite image 2008. Jour. Indian Society of Remote Sensing:175–182

Senthilnath, J., Omkar, S.N., Mani, V., Karthikeyan, T., 2011a. Multi-objective optimization of satellite image registration using Discrete Particle Swarm Optimisation. 2011 Annu. IEEE India Conf. 1–5. https://doi.org/10.1109/INDCON.2011.6139338

Senthilnath, J., Omkar, S.N., Mani, V., Tejovanth, N., Diwakar, P.G., B, A.S., 2011b. Multi-spectral satellite image classification using glow-worm swarm optimization 2011. IEEE Int Geo and Remote Sesing Symp Vancouver. pp. 47–50

Senthilnath J, Omkar SN, Mani V, Tejovanth N, Diwakar PG, Archana Shenoy B (2012) Hierarchical clustering algorithm for land cover mapping using satellite images. IEEE J Sel Top Appl Earth Obs Remote Sens 5:762–768. https://doi.org/10.1109/JSTARS.2012.2187432

Senthilnath J, Shenoy HV, Omkar SN, Mani V (2013) Spectral-spatial MODIS image analysis using swarm intelligence algorithms and region-based segmentation for flood assessment. In: Int. Conf. Bio-inspired computing: theories and applications (BIC-TA 2012). Springer, India, pp 163–174

Senthilnath J, Omkar SN, Mani V, Prasad R, Rajendra R, Shreyas PB (2015) Multi-sensor satellite remote sensing images for flood assessment using swarm intelligence. Int Conf Cogn Comput Inf Process (CCIP) Noida India:1–5

Senthilnath J, Kulkarni S, Benediktsson JA, Yang XS (2016) A novel approach for multispectral satellite image classification based on the bat algorithm 2016. IEEE Geosci Remote Sens Lett 13:599–603. https://doi.org/10.1109/LGRS.2016.2530724

Shahana K, Ghosh S, Jeganathan C (2016, April) A survey of particle swarm optimization and random forest based land cover classification. In: Computing, communication and automation (ICCCA), 2016 international conference on. IEEE, pp 241–245

Singh V, Kumar G, Arora G (2016) Analytical evaluation for the enhancement of satellite images using swarm intelligence techniques. In computing for sustainable global development (INDIACom). 2016 3rd Int. Conf on IEEE 16:2401–2405

Singh A, Chhablani C, Goel L (2017) Moth flame optimization for land cover feature extraction in remote sensing images.2017. Eighth. Int Conf on Comp Comm Net Tech (ICCCNT):1–7

Soliman, O.S., Mahmoud, A.S., Hassan, S.M., 2012. Remote Sensing Satellite Images Classification Using Support Vector Machine and Particle Swarm Optimization. 2012 Third Int. Conf. Innov. Bio-Inspired Comput. Appl. 280–285. 10.1109/IBICA.2012.61

Sood, M., Menon, S., 2016 Cross-country path finding algorithm using hybridization of Bat and cuckoo search. 2016 Third Int. Conf. Computing for Sustainable Global Development (INDIACom), New Delhi

Suresh S, Lal S (2016) An efficient cuckoo search algorithm based multilevel thresholding for segmentation of satellite images using different objective functions. Expert Syst Appl 58:184–209. https://doi.org/10.1016/j.eswa.2016.03.032

Tebbi MA, Haddad B (2017) Improving infrared MSG satellite images classification, application on rainy areas detection. Proc. 2016 8th Int. Conf Model Identif Control ICMIC (2016):747–750. https://doi.org/10.1109/ICMIC.2016.7804211

Tien Bui D, Bui QT, Nguyen QP, Pradhan B, Nampak H, Trinh PT (2017) A hybrid artificial intelligence approach using GIS-based neural-fuzzy inference system and particle swarm optimization for forest fire susceptibility modeling at a tropical area. Agric For Meteorol 233:32–44. https://doi.org/10.1016/j.agrformet.2016.11.002

Turner BL, Meyer WB, Skole DL (1994) Global land-use/land-cover change: towards an integrated study. Ambio Stockholm 23(1):91–95

Upadhyay, A., Singh, S.K., Singh, P., Singh, P., 2010. Comparative study of artificial neural network based classification of 1RS LISS-III satellite images 961–965. https://doi.org/10.1109/ICGCIoT.2015.7380601

Wang Q, Wang L, Liu D (2012) Particle swarm optimization-based sub-pixel mapping for remote-sensing imagery. Int J Remote Sens 33:6480–6496. https://doi.org/10.1080/01431161.2012.690541

Wang L, Geng H, Liu P, Lu K, Kolodziej J, Ranjan R, Zomaya AY (2015) Particle swarm optimization based dictionary learning for remote sensing big data. Knowledge-Based Syst 79:43–50. https://doi.org/10.1016/j.knosys.2014.10.004

Xue Z, Du P, Su H (2014) Harmonic analysis for hyperspectral image classification integrated with PSO optimized SVM. IEEE J Sel Top Appl Earth Obs Remote Sens 7:2131–2146. https://doi.org/10.1109/JSTARS.2014.2307091

Yamaguchi T, Mori K, Mackin KJ, Nagai Y (2012) Application of particle swarm optimization to similar image search on satellite sensor data. 6th Int. Conf Soft Comput Intell Syst 13th Int Symp Adv Intell Syst:1573–1577. https://doi.org/10.1109/SCIS-ISIS.2012.6505327

Yang L, Sun X, Peng L, Yao X, Chi T (2015) An agent-based artificial bee Colony (ABC) algorithm for hyperspectral image endmember extraction in parallel. IEEE J Sel Top Appl Earth Obs Remote Sens 8:4657–4664. https://doi.org/10.1109/JSTARS.2015.2454518

Yavari S, Valadan Zoej MJ, Mohammadzadeh A, Mokhtarzade M (2013) Particle swarm optimization of RFM for georeferencing of satellite images. IEEE Geosci Remote Sens Lett 10:135–139. https://doi.org/10.1109/LGRS.2012.2195153

Zaitoun NM, Aqel MJ (2015) Survey on image segmentation techniques. Procedia Computer Science 65:797–806

Zarrinpanjeh N, Samadzadegan F, Schenk T (2013) A new ant based distributed framework for urban road map updating from high resolution satellite imagery. Comput Geosci 54:337–350. https://doi.org/10.1016/j.cageo.2012.12.006

Zhang B, Sun X, Gao L, Yang L (2011a) Endmember extraction of hyperspectral remote sensing images based on the ant Colony optimization (ACO) algorithm. IEEE Trans Geosci Remote Sens 49:2635–2646. https://doi.org/10.1109/TGRS.2011.2108305

Zhang B, Sun X, Gao L, Yang L (2011b) Endmember extraction of hyperspectral remote sensing images based on the discrete particle swarm optimization algorithm. IEEE Trans Geosci Remote Sens 49:4173–4176. https://doi.org/10.1109/TGRS.2011.2131145

Zhang B, Gao J, Gao L, Sun X (2013) Improvements in the ant colony optimization algorithm for endmember extraction from hyperspectral images. IEEE J Selected Topics in Applied Earth Observations and Remote Sensing 6(2):522–530

Zhang, B., Gao, L., Sun, X., Zhuang, L., 2017. Swarm intelligence: a reliable solution for extracting endmembers from hyperspectral imagery. Work Hyperspectral Image Signal Process Evol Remote Sens 2015–June. https://doi.org/10.1109/WHISPERS.2015.8075433

Zhong Y, Zhang L, Huang B, Li P (2006) An unsupervised artificial immune classifier for multi/hyperspectral remote sensing imagery. IEEE Trans Geosci Remote Sens 44:420–431. https://doi.org/10.1109/TGRS.2005.861548

Zhong Y, Zhao L, Zhang L (2014) An adaptive differential evolution endmember extraction algorithm for hyperspectral remote sensing imagery. IEEE Geosci Remote Sens Lett 11:1061–1065. https://doi.org/10.1109/LGRS.2013.2285476

Implementation And Performance Analysis of Proposed Security Framework For Uidai

Dr. Arpana Chaturvedi

Asst. Prof. & Department of Information Technology & Jagannath International Management School, JIMS, GGSIPU, New Delhi 110070, India

Abstract — *The development of new technologies like Hadoop, Map Reduce, used to store, manage, analyze vast amount of data, when associated with WSN (Wireless Sensor Network), the risk factor for information security increases. When it is applied in various applications of government like Aadhaar, DigiLocker etc., chances of hidden security issues increased. In this paper AES-XTS encryption mechanism and digital signature technology is used with AODV (Ad hoc On-Demand Distance Vector) routing protocol to get rid of various issues like DoS (Denial of Service), eavesdropping, imitation, coaxing etc. The theoretical analysis is shown using NS2 simulator and implementation assures to provide better data security, reliability, transmission and energy efficiency. This implementation shows that SAODV(Secure-AODV) routing protocol when used in routing layer in these applications and system, it increases its own self defensive ability to fight against various challenging hidden security issues.*

Keywords - *AODV, SAODV, WSN, AES-XTS, UIDAI, DoS.*

I. INTRODUCTION

The UIDAI (Unique Identification Authority of India) data, both at rest or on move, stored at CIDR (Central Identities Data Repository) is at high risk. Due to the advent of new Hacking technologies and advanced tools, the risk factor has increased. The government is enforcing citizen of India to link Aadhaar number with different citizen centric services so that right citizen can avail the benefits. There is need to safeguard the data stored in different data centres which might be secured through a strict encryption technique applied on application process in parallel mode. In this chapter, an encryption process to be performed in parallel mode is discussed. The proposed approach is AES in XTS Mode in Map Reduce paradigm which supports parallel programming in the distributed environment. The findings after reviews of results show that it provides better security against external attacks and overcomes the shortcomings of Kerberos. Encryption followed by compression on various datasets provides better result and protection from vulnerabilities and threats.

With the development of wide application of WSN (Wireless Sensor Network) and advanced technologies, the risk to the information security in various government and private sector has increased. Here we integrate the AES encryption standard in XTS mode and digital signature technology to improve classic AODV (Ad hoc On-Demand Distance Vector) routing protocol. The resultant SAODV (Secure Ad hoc On-Demand Distance Vector) routing protocol provides better information security and achieves energy efficiency as well. In this chapter, we reviewed different observations and it is concluded that the algorithm is an appropriate choice.

The chapter is divided into further four sections. Section II explains the requirement of strict encryption algorithm in current scenario. The concept is suggested with respect to the requirement of security approaches for the UIDAI System. Section III describes the importance and benefits of AES-MR technique highlighting the XTS encryption Mode. Section IV shows the experimental results are explained using implemented simulation environments. It highlights the benefits of using this approach to secure sensitive information. The chapter is finally concluded in Section V.

This document is a template. An electronic copy can be downloaded from the conference website. For questions on paper guidelines, please contact the conference publications committee as indicated on the conference website. Information about final paper submission is available from the conference website.

II. REQUIREMENT OF STRICT ENCRYPTION OF DATA

Since ages data collection has increased enormously with high velocity. 80% of the total data is collected just during the last two years. This extraordinary growth in data generated is due to miscellaneous activities performed by common man. The advent of technologies like cloud computing, active use of social media, mobile computing, internet of things, sensor-based network etc. requires large and efficient storage with security. This advancement in technology increases the generation of data exponentially and need strict security

strategies to be implemented in the system. The sensitive data are stored in various datacenters. The government initiated the concept of Digital India and enforces end users to link their Aadhaar number with various services. UIDAI system and related infrastructure have already implemented strong security authentication and compliances. With the pace, speed, and momentum the citizens are enforced to link their Aadhaar number with various services to avail benefits, their concern about the breach of privacy of their data has increased. Citizens of India have submitted their demographic and biometric details to the UIDAI system. To avail benefits, their personal data need to be shared among different agencies. They need assurance that the data will not be misused. Many issues in past were raised regarding misuse of private information, leakage of private information and theft of data. IBM estimates that 90% of the world's data was generated in the last few years alone which has increased the reliability and security challenges to keep safe and secure the enormous data. Major challenges are related to reliable storage, efficient processing, data integrity, and recovery. It is necessary to modify the existing security compliances, legal provisions, and auditing policies.

There are so many technologies and algorithms available which might be appropriate to handle security issues. The study concludes that the implementation of AES-XTX with map reduce parallel programming will be a cost-effective solution to process such a large user-generated vital and sensitive data.

An easy way to comply with the conference paper formatting requirements is to use this document as a template and simply type your text into it.
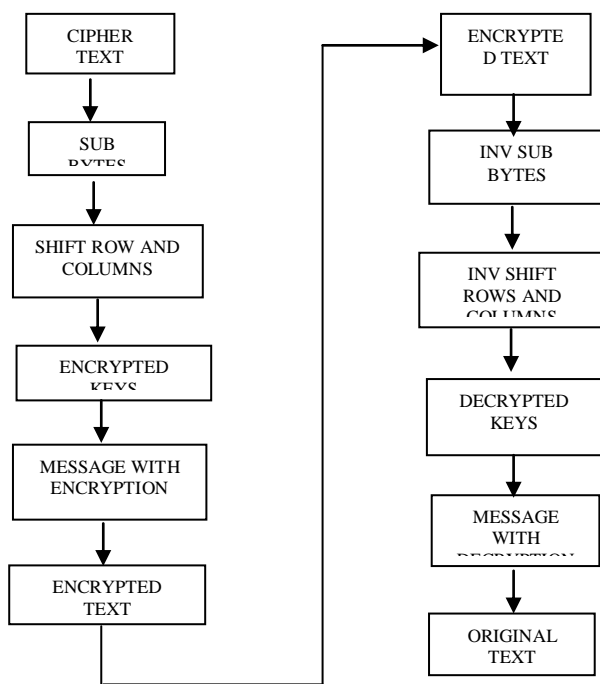
Fig 1 Encryption and Decryption Process

III. AES-MR

The combination of Advanced Encryption Standards (AES) and Map Reduce (MR) is termed as AES-MR. It is the proposed algorithm for this research work which is related to the case study done for UIDAI to provide better data security. This encryption algorithm is proposed to provide Data level security. Map Reduce is a parallel programming language and AES is the encryption algorithm suitable for longer messages. It is suggested that the best features of both techniques should be used together to provide much stricter security features in the introduced security layer.

Map Reduce with AES-XTS has the capability to compose applications that generates endless data during runtime. It has the ability to adapt non-critical failures and can perform better planning, testing of information. The failed jobs if encountered in the clusters of machines, it re-executes them.

PROPOSED ENCRYPTION ALGORITHM-AES (XTS)-MR

In this proposed technique AES is used with XTS mode which is supported by IEEE 1619-2007 standards [5]. The XTS modes contain XEX-TCB-CTS (XTS) mode where XTS stands the XEX Tweakable Block Cipher with Cipher Text Stealing. The XTS mode performs parallel executions and allows pipelining in respective executions. Data Encryption Standards which were used earlier is vulnerable to Brute Force attacks. It was due to the small size of the key (53 to 2054 bits) DES uses for encryption. US Government Agency NIST (National Institute of Standards and Technology) selected Rijndael's Algorithm as Advanced Encryption Standard. It is being a better security standard now becoming an Industry Standard.

AES is designed to accept 128, 192, 256 bits' size of keys. These various sizes of keys are capable of encrypting different types and variety of information in bulk. The performance of AES algorithms varies on different 32 bit and 64-bit CPU's based on key sizes. This technique will provide better security (Fig.2) in case of the UIDAI system where the large data sets are generated for storing sensitive information of residents and generating UID numbers of a citizen of India.

The block contents can perform parallel processing in various modes of operations. These operations can handle fixed block as well as variable block encryption with the help of single key or different keys.

XTS ENCRYPTION MODE

The Electronic Cook Book (ECB) and XTX are used with AES to increases the effectiveness of an algorithm and can be improved by the use of it. [5] The XTX supports parallel encryption mode with Symmetric Block Cipher encryption mode. It was designed to protect data lying at rest on storage devices.

It uses a fixed size of data units to perform cryptographic protection of data at rest. The operation of AES-XTX Mode with two different keys is shown in Fig. 3 [12] [5]. The XTS-AES mode is an enhanced concept of Rogaway's XEX (XOR Encrypt XOR) Tweakable Block Cipher, improved with a method called "Cipher Text Stealing". It expands the range of possible different types of data inputs. XEX can only encrypt sequences of complete blocks of any data type. This input data should be an integer and necessarily be a multiple of 128 bits. In XTS-AES, the data string consists of one or more complete blocks which are followed by a single, non-empty partial block ().
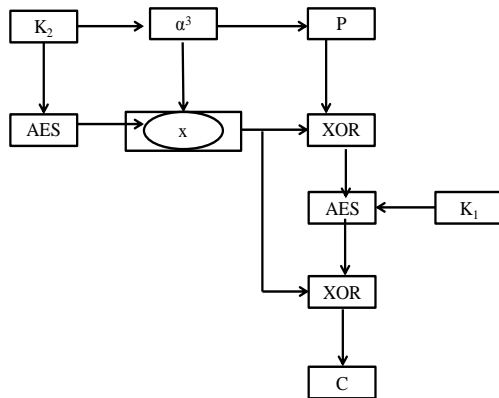
Fig.2 Operation of AES-XTX Mode with two different keys

The working of AES-XTS is shown in Fig. 3. The XTS-AES is composed of two keys, first one is an encryption key and second one is tweak key. It incorporates the logical position of the data block into the encryption [11]. The output produced by XTS is independent which leads to parallelization. XTS, an instantiation of the tweakable Block Cipher class. It is capable to implement ciphers in parallel and pipeline modes. It enables the encryption of the last incomplete block also.

IV. EXPERIMENTAL RESULT AND ANALYSIS

The experimented results are drawn are done using various randomly created sizes of datasets. This work integrates the international popular AES-XTS encryption standard to improve existing AODV routing protocol, and uses improved, energy efficient SAODV routing protocol to provide better information security. The details of the simulation environment used to show the proposed implementation work is shown in Table 1.

CPU type	Intel® Core™ i5 4210U
Clock speed	CPU ®@2.40 GHz
Ram size	16 GB RAM x-64- based processor
Operating System	Ubuntu Server Basic, 64
	Bit Windows 7 Home.
Framework	Hadoop Framework
Simulation Tool	NS2
Tools	Ms. Office, MS Excel
Language	Java, , JDK1.8,

Table 1 Details of Simulation Environment

The theoretical analysis and simulation shows the better performance of SAODV by adding proposed security mechanism to AODV in the routing layer. The steps of implementation work are shown in Fig. 3.

Implementation Work
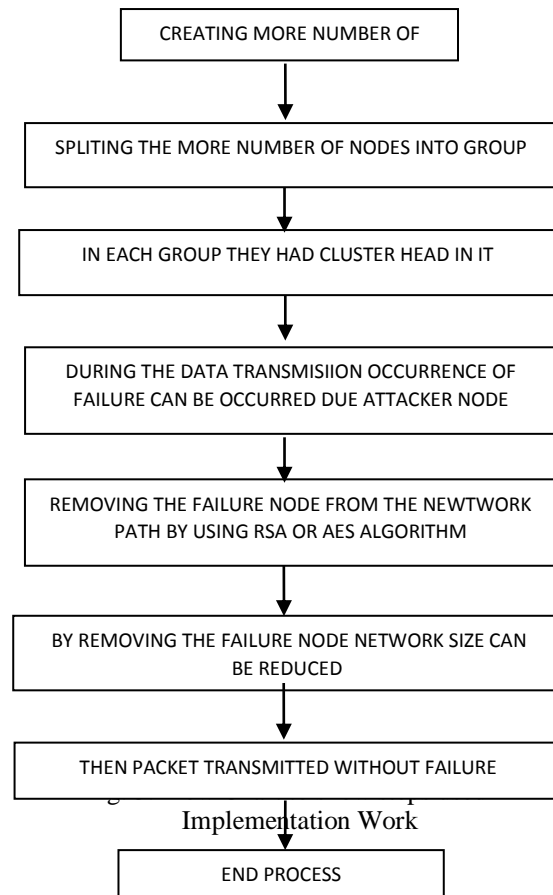
DESCRIPTION OF ROUTING PROTOCOL AND ITS ASSIGNED PROTOCOL

The routing protocol used in this simulation environment to show implementation and other details are shown below in the Table 2 as:

Routing Protocols	Remarks
Number of nodes used	31
Node act as source node	Node 2
Node act as destination node	Node 21, 15, 19
Node act as cluster head	Node 4, 9, 16, 26
Mobility speed	10mps
Simulation time	25ms
Transmission range	300m
Mobility movement	Random path
Transmission range	2packets/sec
Number of connection	5 connections
Buffer size	128 packets
Number of graphs	6

Table 2: Description of Routing Protocol and assigned protocol

USING AES ALGORITHM

In the below shown screenshot (Fig. 5), it clearly gives the explanation of the process of encryption and decryption by using AES method has been explained. Thus in this first we used plain text after encryption process can be occurred and convert the message into cipher text.

At final receiver want to receive the original message thus the decryption had been occurred and convert cipher text message to original message.

Fig. 5. Screenshot of command used to execute the implementation work

OUTPUT OF IMPLEMENTATION WORK

The various output of the implementation work is shown below in the sequence of implementation of work performed.

A. NODE FORMATION

The formation of nodes is shown here. In this the total number of nodes used are 31. Among these 31 nodes, four nodes are used as a cluster node. These are node 4,9,16 and 26. The node which acts as a source node is 2. The nodes which acts as a destination nodes are 21,15 and 19. The mobility speed 10 mpbs. The simulation time is 25ms and the transmission range is 300m. The total number of connections used in the implementation are 5 with random mobility movement. The transmission range is 2 packets per second with buffer size of 128 packets. The four channels ch1 as node 4, ch2 as node 9, ch3 as node 16 and ch4 as node 26 are used with Key exchange node as node 7. The total number of graphs shown here are 6.
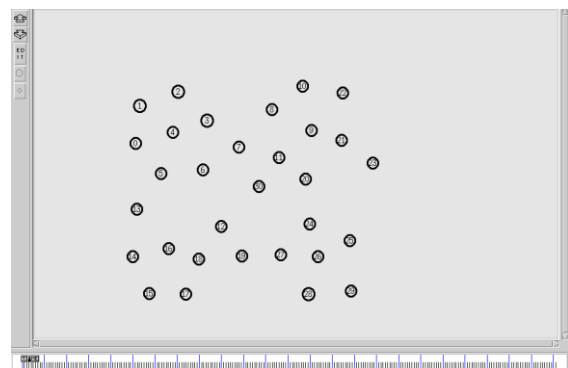
Fig. 7. Screenshot which show Node creation

In the above screenshot (Fig. 7) node creation has been occurred. In this network number of node created is 30.

B. ALLOCATION OF SENDER AND RECEIVER FOR VARIOUS NODES

In the below (Fig. 8) screenshot for the formation of group of nodes it had been assign the cluster head with source and destination.
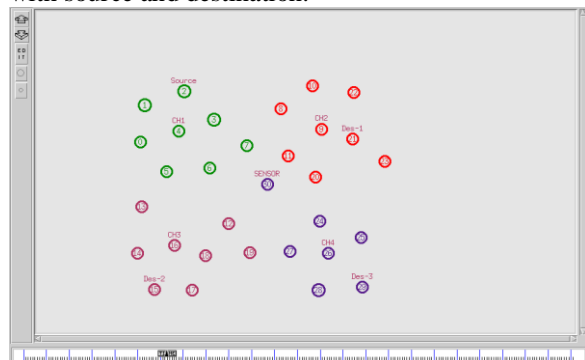
Fig. 8. Assign the cluster head with source and destination

C. DATA TRANSMISSION

In this graph (Fig. 9) it has shown that how TCP/IP protocol is selected among various channels. The total channels used here are 4. These are as Ch1, Ch2,

Ch3 and Ch4 as node 4, node 9, node 16 and node 26 respectively.
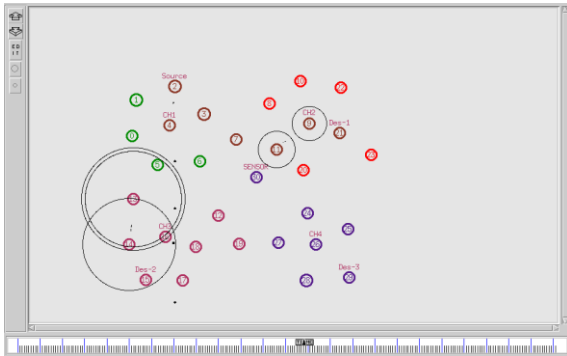
Fig. 9. Selection of Channel by TCP/IP Protocol

In the above after the assigning of cluster head, source and destination the process of transmission had been occurred. While transmitting the data loss may be occurred due to this security of the network had been occurred problem.

D. TRANSMISSION WITH KEY EXCHANGE

In this graph (Fig. 10) the key exchange node is identified as node 7. It is responsible for exchange of keys while transmission of packets from single source to multiple destinations. It performs encryption and is known to sender and receivers only. As during transmission only, the key exchange node is identified, it is very difficult for hackers to identify it. In this way it provides better security implementation.
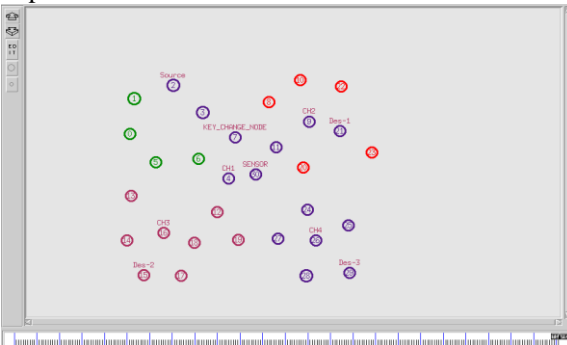
Fig. 10. Transmission of Information with Key Exchange

In the above screenshot due to the occurrence of loss we want to reduce that loss and improve the security of network. We can use the AES algorithm in which while transmitting the data, key had been used to improve the security without any loss in the network.

GRAPH REPRESENTATION

The various output obtained after implementation of proposed technique is shown below in the form of graph as:

A. DROP IN DATA

In this graph (Fig. 11) the packet drop has been showed between both existing and proposed method. Packet drop has been calculated by number packets drop in network while transmitting the data between the different nodes.
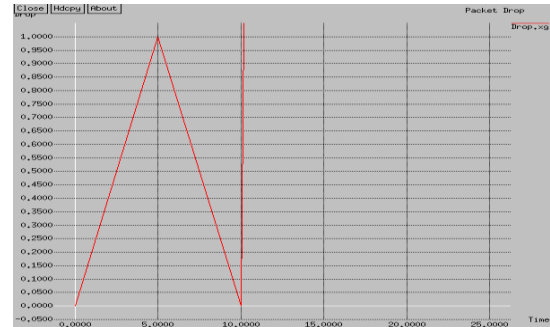
Fig. 11. Packet drop between the Existing and Proposed Method

B. DATA TRANSMISSION IN ENCRYPTION

In this graph (Fig. 12) the filtering of nodes and packets are shown. It checks the packets which are manipulated while transmission by some intruders, hacker's r or by eavesdropping. It does not allow the corrupted packets as removes them and passes rest of the filtered packets. It also checks the nodes. If some nodes through which packets are transmitted are not authorized and registered, it removes it from the channel.
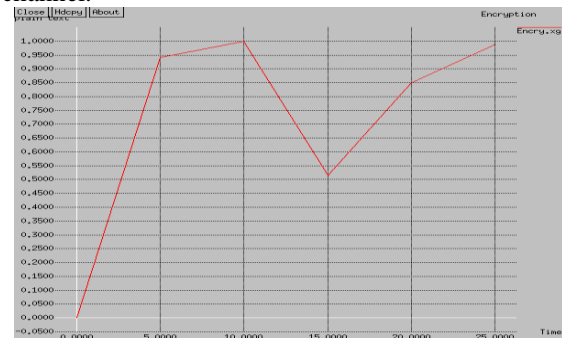
Fig. 12. Filtering of Nodes and Packets

In this graph the data encryption has been showed while transmitting the data. When the sender sends the original data by using encryption process data can be converted to cipher text. During the process encryption of the data in network has been calculated.

C. DATA DECRYPTION WHILE DATA TRANSMISSION

In this method data is decrypted while transmission of data to provide better security using AES technique.
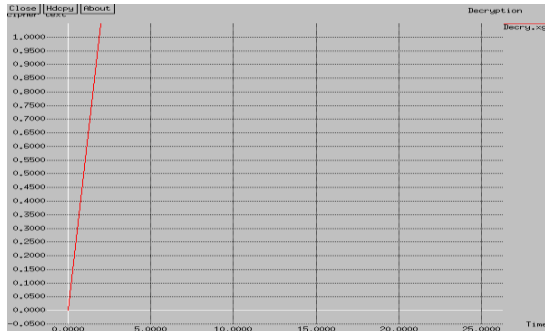
Fig. 13. Data Decryption while Transmission of Information

In this graph (Fig. 13.) the data decryption has been showed while transmitting the data. Thus after the process, of encryption data want to decrypt to plain text for the receiver view. Thus data transmission has been calculated in decryption.

D. COMPARISON FOR LOSSES IN PACKETS IN EXISTING VS PROPOSED

The loss of data has been showed while transmitting the data. It is identified by particular amount of data loss and filtering of nodes.

Fig. 14. Loss of Data while Transmission of Information

In this graph the loss of data has been showed while transmitting the data. When the data transmitted from sender to receiver, the particular amount loss in data has been occurred. Thus it has been calculated and shown in the graph in the relation of existing and proposed method

E. DATA PACKET TRANSMISSION IN VARIOUS NODES

The data transmission has been showed while transmitting the data in various nodes. Sender sends the data packets n various nodes and it is clear through graph (Fig. 15) that more packets are transmitted in case of proposed methodology in comparison to the existing one. It has shown a 95% of difference in number of data packets transmitted in various nodes and increases the speed of transmission also.
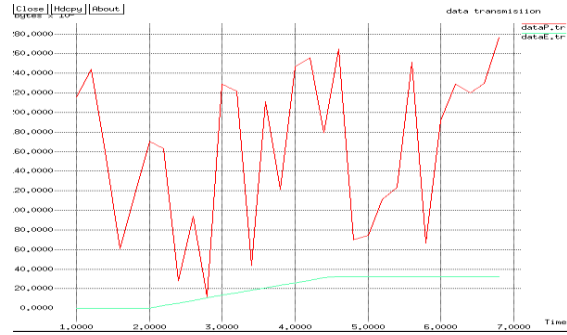
Fig. 15. Increased speed when Data Packets transmitted through Various Nodes

In this graph the data transmission has been showed while transmitting the data from sender to receiver.

F. Passing of Data with Individual Key Values

In this graph (Fig. 16) it's clearly shown that when data passed or transmitted through different channels in the form of packets, using AES-XTS encryption methodology and using key exchange according to this encryption technique, provides better transmission in comparison to existing encryption techniques. The comparison shows the 95% of betterment and difference in transmission with individual key values. This proposed encryption technique results in better security implementation in comparison to existing mechanisms
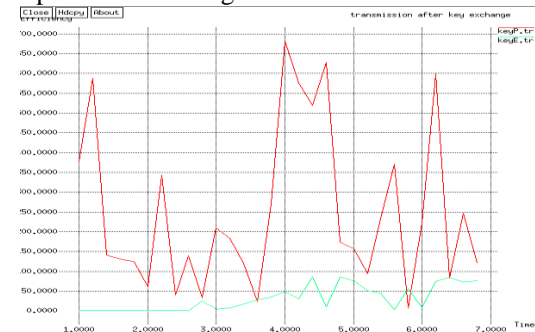
Fig.16. Comparison of Security implementation in Existing and Proposed Mechanism

In this graph the data transmission has been showed while transmitting the data from sender to receiver in the presence of key.

V. Conclusion

In this work we suggested AES-MR encryption technique to be implemented to solve the issue of data security at the storage level i.e. on HDFS (Hadoop Distributed File System). It is done by encrypting the data using AES-MR(XTS) along with compression at Mapper and Reducer phase. AES-MR(XTX) will not only enhanced the security of important data at HDFS level, but with the help of parallel processing we can do it with a faster manner as well.

With the help of simulation using NS2 simulator it is concluded that as the security is key anxiety of the wireless network which transfer data from nodes

towards other. In the existing work, the performance of MANET is considered by using different parameters such as Network Overhead, Delay and Throughput. There are a few nodes which increase the system transparency, Delay and reduce Network throughput they are famous as malicious nodes. In proposed work, we used AES algorithm in WSN and then encryption to expand the demand or of network. In the future this approach can also be used to get better security in other networks like VANET, SPANs etc. it has been concluded that AES encryption standard when used in XTS mode will give better results in terms of security, Speed of transmission, better filtration of packets and Nodes. This technology can be used in various services provided by government now-a-days where it became mandatory to use Aadhaar card to link with the services to avail benefits. The block size used in AES is comparatively larger and varied, depending upon the requirement. The encryption process of XTS mode in AES is also complex that hackers cannot easily hack the transmitted crucial information. We can conclude that as AES-MR (XTS) helps to attain all the levels of security that too at a faster speed, it is a good approach.

ACKNOWLEDGMENT

REFERENCES

[1] Kadre V., Chaturvedi S., "AES – MR: A Novel Encryption Scheme for securing Data in HDFS Environment using Map Reduce", www.ijcaonline.org/ research/ volume129/ number12/kadre-2015-ijca-906994.pdf

[2] Mehak, Gagan, "Improving Data Storage Security in Cloud using Hadoop ", ISSN: 2248-9622, Vol. 4 , Issue 9(Version 3), September 2014, pp.133-138,http://www.ijera.com/papers/Vol4_issue9/Version%20 3/U4903133138.pdf

[3] Weeks B., Bean M., Rozylowicz T., Ficke C.," Hardware Performance Simulations of Round 2 Advanced Encryption Standard Algorithms, National Security", https:// csrc.nist.gov/CSRC/ media/ Projects/ Cryptographic-Standards-and-Guidelines/documents/aes-development/NSA-AESfinalreport.pdf doi=10.1.1.35.6941

[4] Clunie D., Public Comments on the XTS-AES Mode," https://csrc.nist.gov/csrc/media/projects/.../comments/xts/collected_xts_comments.pdf

[5] Public Comments-Modes Development - Block Cipher Techniques, "https://csrc.nist.gov/Projects/Block-Cipher-Techniques/BCM/Public-Comments-Modes-Development",Comments submitted to Encryption Modes @nist.gov.

[6] Dr. Hawthorne, NY, Computing Arbitrary Functions of Encrypted Data Craig Gentry IBM T.J. Watson Research Center 19 Skyline,cbgentry@us.ibm.com https://crypto.stanford.edu/craig/easy-fhe.pdf.

[7] Desai Spark Y., Gao J., Sang-Yoon Chang, Chungsik Song, "Improving Encryption Performance Using Map reduce", Published in: High Performance Computing and Communications (HPCC), IEEE 17th International Conference, ISBN: 978-1-4799-8937-9, http:// ieeexplore. ieee.org / document/7336355/

[8] G. Sujitha, M. Varadharajan, B. Raj Kumar and S. Mercy Shalinie ,"Provisioning Mapreduce for Improving Security of Cloud Data ",http:// scialert.net/ qredirect.php? doi = jai.2013.220.228& linkid=pdf, International Journal of Computer Science and Applications, © Technomathematics Research Foundation Vol. 13, No. 2, pp. 89 – 105, 2016.

[9] Alexander Uskov, Adam Byerly, Colleen Heinemann," Advanced Encryption Standard Analysis with Multimedia Data on Intel® AES-NI Architecture", Research Institute Bradley University, 1501 West Bradley Avenue Peoria, Illinois 61625, U.S.A. auskov@bradley.edu http://www.tmrfindia.org/ijcsa/v13i26.pdf

[10] Demir L., Thiery M., Roca V., Jean-Louis Roch, Jean-Michel Tenkes, "Improving dm-crypt performance for XTS-AES mode through extended requests ", Nov 21, 2016 The 4th International Symposium on Research in Grey-Hat Hacking - aka GreHack, Nov 2016, Grenoble, France https://hal.inria.fr/hal-01399967

[11] Philip Derbeko, Shlomi Dolev, Ehud Gudes, Shantanu Sharma" Security and Privacy Aspects in map reduce on Clouds: A Survey", www.https://arxiv.org/abs/1605.00677, www.iiste.org, ISSN 2224-610X (Paper) ISSN 2225-0603 (Online) Vol 2, No.2, 2012

[12] Liskov M., Mine Matsu K., "Comments on XTS-AES" September 2, 2008 This is a comment in response to the request for comment on XTS-AES, as specified in IEEE Std. 1619-2007 September 2, 2008, https://csrc.nist.gov/csrc/media/ projects /block-cipher-techniques/ documents/ bcm/ comm. ents/xts/xts_comments-liskov_minematsu .pdf.

[13] Kirat Pal Singh, Shiwani," An Efficient Hardware design and Implementation of Advanced Encryption Standard (AES) Algorithm ", https :// eprint.iacr.org/ 2016/ 789.pdf.

[14] Vaidyaa M., Dr Shrinivas Deshpandeb ," Study of Performance Parameters on Distributed File Systems using map reduce ", www.sciencedirect.com ICISP2015), 11-12 December 2015, https://ac.els-cdn.com/S18770509 16000399/1-s2.0-S1877050916000399-main.pdf ?_tid=78 c69a4a-e233-11e7-8cfd-00000aab0f01& acdnat=15134098 15_ d18af66cf2c2e5fa578411397b06ce28

[15] Wei Li, Ming Chen, Mingming Li, " Information Security Routing Protocol in the WSN", Published in: 2009 Fifth International Conference on Information Assurance and Security, ISBN: 978-0-7695-3744-3, https://ieeexplore.ieee.org/document/5284242

[16] Sonkar Abhilash., Aggarwal Abhishek, "Enhancement of Security using greedy approach and encryption in Mobile Ad Hoc Network", Published in: 2017 International Conference on Trends in Electronics and Informatics (ICEI), https://ieeexplore.ieee.org/document/8300887

Relationship of Information Technology with Management In Unified Communication – Importance Towards E-commerce & Digital Marketing

Dr. Ruchi Singhal *, Ms. Rama Rani**

*Associate Professor, Institute: Jagannath International Management School, New Delhi.
(Affiliated To GGSIP University, New Delhi)
Research Scholar Institute: Galgotia University, Noida*

Abstract

In today's era, the Unified communication is being proven a significant role in the timing of e-commerce & Digital Marketing. Basically, some collections of communication tools which are going to be very important tool for office work is known as Unified i.e. Amalgamated Communication. There are some famous and common tools of Unified Communication such as Scheduling of team work through Email, Telephonic Conferencing, Video calling and Calendaring & some real-time communication tools. A browser based interface and/or a thick client are being used for using unified communication tools by a user. The video conferencing, Sms Capability, Fax Services, Text Messaging & Mobility Services are an essential part of amalgamated Interaction System. Today's largely spilt exchanging modes are integrated by the Unified communication. In today's era, the amalgamated communication has a main feature in each organization as Increase in employees efficiency, decrease in IT costs & increasing the swiftness & elasticity are same motives for all of the organization, So, unified communication platform should be within affordable prices and easy to manage and surely reliable to accomplish the motives of the organization. The motive of the research is to specify the necessity of using amalgamated communications among executives and users of an organization which can use all of their communiqué tools from a single edge as this communication is very fruitful while in case of not well or inclement weather or somehow during travelling. This communication is much better and fruitful for the environment for office sitting persons. By using amalgamated interactions which is independent of distances from the working area, we can have better user occurrence and by using these amalgamated communication tools the employees are motivating for their work with greater potential.

Keywords: *Mobility Services, unified, browser-based, inclement weather, Swiftness, amalgamated interactions.*

Introduction

According to the ecological meaning UC or Unified Communication is the amalgamation of online or real time communiqué facilities like online quick messaging i.e. chat, telephonic discussions, data distribution, speech recognition, video conferencing, call control and some unreal time communiqué amenities like Unified Messaging which includes emails, sms, fax and integrated voice mail. It is not just one but an amalgamation or set up products that provides a regular united customers friendly interface end users or customers familiarities across numerous equipment or devices

We can also say or we can define Unified Communication is open in each and every context of communication that are mixed or exchanged between two persons through a set up connection which includes added ways of interactions. Such type of communications allows a personage to deliver or exchange message through one source or object and to receive exactly the exact message on different sources. As an example, we could have an individual can receive a message through voice mail and thinks of accessing the same message via email or a device such as cell phone. If the person who is sending the message is online and currently is attending information and accept the call, a retort can be sent on that moment through a message, video call or a text chat, or it might be sent as a virtual communication that could be easily accessed via multiplicity or different devices and mediums.

Meaning of U C:

So many definitions of Unified Communication are there. A very basic and entirely fruitful meaning is "Interactions with Communication devices incorporated to optimize business processes with high profits within rational timings".

UC is a sprouting collection of tools and techniques which allows functioning and unifies the relation between human being or gadgets, in a universal framework or the occurrence. With the help of these tools there is improvement and optimization of company's ways and proceedings by enhancing human communications through dipping latency, running flows by eliminating device equipment and dependency on media.

From where does UC originated and its working:

It is a technology of progression of grouping the sustaining technology. Firstly, business telephone systems were considered to be main model for the development of UC. Phone companies used the analog and digital circuits to transfer phone calls from a main office to the users. The key telephonic system named as PBX was used during that time which acknowledges the call routed to the suitable line or extension facade on the cell phones and the consumer office. When the era of 1980s existed the audio mail SMS with Interactive voice response features were considered and treated as a tool to convey the peculiar information for phone workforce. Up to that time there was not any family tree of mobiles and the creation personal computer. There was a grow in Electronics in around 1985. Electronic mails reading features were made accessible for sending confident voice mails. A phrase used "Unified Communication" came to the general public following the amalgamation of texting and real time communication. The delayed 1990s IPFX developed a company invention which allowed the users to see the position of contemporary and made decisions on the way to get in touch with them and let users describe how their communication was handled based on their personal occurrence which resulted in a residual, recurring cost or penalty to customers. Over the time, the PBX became additional privatized and domestic work force was asked or taken help from to manage these systems. Companies characteristically undertook this so that they could come up with the money for to bring such type of ability in their work environment and thereby diminishing the necessity to advise the telephonic companies or their confined The key telephonic system (PBX) vendors as each time a change was required in the systems. This change in escalating and accelerating privatization of an organization provoked the expansion of special powerful devices or objects that improve the utility of the structure. The companies began to arrange Internet Protocol networks in the surroundings in which they are working and began to put out voice instead of depending on conventional telephonic network circuits. There are few purveyed or twisted circuit cards or packs for their key telephonic system i.e. PBX systems that can interrelate their exchange system to the internet protocol networks. Some other vendors produced tackle that could be situated in routers to carry audio calls across the organization's association from different location. The execution of key telephonic system ie. PBX circuits and networks which are to be carried across an integrated system and deliver to another device/phone arrangements which is frequently refer to as VOIP on voice over internet protocol. This drawing required a particular type of hardware install on equal end of the network model apparatus so that it can provide the killing and release it at every location. This internet protocol or IP explanation is a software ambitious only and thus do missing by means of the necessitate "for switching" apparatus at a client's location. This fashioned a new technology known as internet protocol or IP telephone. It is a system which uses internet protocol based telephonic services only rather than a heritage private branch exchange a key system which is called IP telephone explanation. With its arrival, the IP telephone did not work like a digital machine which was a synonym form of PBX. While bearing in mind the labors of amalgamated connections solution provider, our all over objective does not for being focussed harshly on the telephonic segment on a daily basis interaction. Such amalgamation of all communiqué equipment and devices inside a single stage provide the power of mobility, existence and get in touch with the capabilities which help in expanding a way from telephone to all plan a human being wants to use or have at their discarding. The objective of the discussed open sources Unified Communication is

to allow all the open foundation society of developers and users to tell exactly the same in what UC means.

Technology

The UC is occasionally confused with Unified Messaging but they are different the amalgamated communication refer to or include non - real and real delivery of connection depending the most favorable technique and receiver's site. Amalgamated Messaging through different ways such as emails, voicemails and at the time of recovery. UC works on a human being to ensure and recover a voice mail or an email from any communiqué mechanism at desired time. This expands beyond voice mail services to data interactions and video facilities.

Components

Along with amalgamated connections, several ways of business related communications are integrated. Unified Connections are not only a single invention but a collection of various elements which includes:

- Presence

Presence is considered to be the foundation of Unified Communication Solution. It is considered to be the enabler of UC and the present will be "the dial tone of the expectation". It provides instantaneous warning of users existing accessibility and capability of commune. The servers of presence meet existing information which is extracted from a variety of sources and supply presence information to the key users or application. In a Unified Communiqué Environment and world whenever we talk about the presence, we are trying to depart further than the simple instantaneous point presence which enables all interactions, together with telephony. Nearly all switch vendors now a day also provide and offer their own existing servers and capabilities and put together along with existence capabilities. The major confront today is to be deficient in of confederation or the capability of such presence system to effort jointly to permit the consumers or the users on single existence systems so that they can observe the existent status of a collaborator or client on different systems.

- **Instant   messaging**
  instant messaging (IM) is an additional serious constituent of a Unified communiqué solution. In commerce environment, Instant messaging (IM) usually requires an enterprise-grade Instant messaging (IM) system, relatively than a public Instant messaging (IM) service such as America Online (AOL) or Yahoo. Venture Instant messaging (IM) systems suggest safety and isolation that public Instant messaging (IM) services cannot. As with existence servers, however, most enterprise Instant messaging (IM) systems do not interoperate with others.
- **Amalgamated   messaging**
  Amalgamatedi.e. Unified messaging (UM) is the amalgamation of voice, fax and email communication ways and message warning. Unified messaging (UM) allows users to access any of these messages, anyplace, anytime, from any mortal. Building on Unified messaging (UM)'s store-and-forward capabilities, most Unified messaging (UM) goods add a multiplicity of superior call and message administration function, together with desktop call viewing of inbound calls, locate me/go behind me, survive respond or call revisit, and cross-media messaging. New occurrence capabilities mean that the necessitate for Unified messaging (UM) systems to act as answering apparatus is being condensed, and the worth of Unified messaging (UM) is affecting on the way to enhanced, immediate connectivity with persons.
- **Speech   access   and   personal   supporter**
  Using speech instructions, personal assistant (or practical assistant) allow user to access their inbox, calendar and directory and so on. Individual assistant supply bright transmission and filtering of messages and let users find the way their agenda, calendar, contacts, outbound dialing and so on, in adding together to their Unified messaging (UM) system.

- **Conferencing and collaboration**
Conferencing and collaboration includes audio, video and Web conferencing, as well as collaborative capabilities such as common workspaces, white boarding, file allocation and file distribution. The fastest-growing knowledge in the two-way assortment is net conferencing, which bring association to the desktop via a network browser and an Internet link, allowing participant to analysis presentation and additional credentials while participating in an instantaneous meeting. Voice communications can take place over the Internet or through a separate audio conferencing bridge. Another collaboration component is shared workspaces, which allow participant to outlook, allocate, check over and keep credentials and documents.

- **Mobility**
In the Unified messaging world, additional voice and video connections will be launched from beginning and linked to mobile wireless devices. Integrating the mobile users' voice and real-time interactions services with core venture transportation let them do their job in spite of site. Enabling Unified Communication (UC) features like click-to-conference and Instant Messaging (IM) for mobile users allows them to contribute in conference and to access some of the similar capability they can access in the place of work, civilizing their efficiency when mobile phone.

- **Business process integration**

  A significant constituent of a Unified message (UM) solution is amalgamation of business process and also workflow application. major goal of integration of business process is to remove "human latency" -- caught up by the necessitate to hang around contribution or communication. Now a day in many situations, work come to a languish awaiting offer information looked-for proceeding. Unified message (UM) can lower this wait by contact the next individual in a series of ladder, or by starting an ad hoc assembly or consultation call to resolve a matter. By communiqué- or voice-enabling business functioning and application, interactions can be started within the application, warn and interrelate among solve a situation.

- The primary application is to be communiqué enable is support office application such as Customer relationship management (CRM), Enterprise resource planning (ERP), sales force automation and supply-chain management. Order completion and client check willingly lend themselves to generalization through occurrence and exchanges consciousness.

- An illustration of basic business process mixing is a procedure that uses MS-Office i.e. Microsoft Office application. Anyone on review a file or a worksheet and has added things about biographer can just click on the writer, check the present situation of his/her, and click-to-call to start an instantaneous talk. The same can be completed in dedicated application. A mechanized exemption structure, for example, can perceive a problem and repeatedly notify the suitable people (i.e., excellence assurance expert, wangle and administrator) by means of any interactions mode, be it cell phone or desk phone, electronic mail or instantaneous text.

Scope in Future:

Now days, the outlook is very brilliant as at the premature stage of unified communiqué deployments. Both the dealer and the venture clients are harassed to know what such new earth hold is very clear: The prospect the start the shift Unified message (UM)completion will have a viable edge over those who holdup.

Shortcomings

Unified communiqué use knowledge to amalgamate voice, information, and commerce application. Its mechanism is by keeping away barrier so that it is achievable to admission and to make use of connections by all workers. Its working is on a multiplicity of plans exclusive of look upon to site. When allowing for unified exchanges compensation and disadvantage, its clear that one of these reimbursements is the cost investments achieved during the consolidation of tools and services. Instead

of running's pli network for information and video tape as well as voice, such unified interactions architectures permit for the middle management of network. This can give enhanced telecommunication opportunity. Amalgamated telecommunications offer business a means to make use of telecoms in additional to a price center. In a disturbed economy, like the one we have at the present, unified interactions can increase information distribution, lesser expenses, and to recover returns opportunity, which could propose a spirited advantage to a trade.

One more "professional" in support of unified communication(UC) come with the capability to allow all human resources, together with smart mobile phones and those operational at distant dispersed location, in the company of total industry telephony capability. With a ordinary consumer boundary, with no observe for the machine or mobile phone type in use, it provide the chance for augmented efficiency and worker contentment with the system provided.
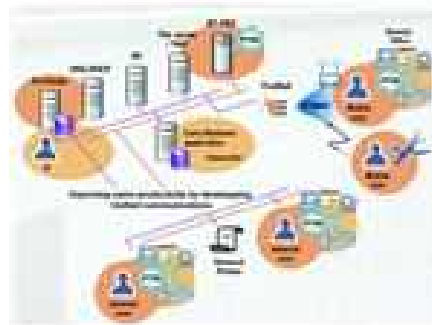
Several vendor offer diverse unified communiqué solution, but the situation sandwiched between both of them can be a main concern because no one dealer can give all the technicalities and capability that descend below the unified exchanges caption. This means that solution have to turn into software based. So the interoperability harms caused by hardware need to be defeat. However, those still have to take place.

A solution is that hardware self-determining is significant for a trade. A unified connections answer, in order to turn into a key constituent of the organization's process, cannot simply alter hardware models or offer added dispensation control or memory when necessary. It is obvious that solutions that advise a gain the direction of solitary seller architecture get the wrong idea what are the main compensation of unified exchanges. They are the elasticity and long-standing monetary profit extracted by having the capability to change the set of connections as time and chance needed by the business houses. However, in the intervening time, maintaining a set of high excellence big business applications for all the users is significant.

Conclusion

We gain many profits from amalgamated communications. The return and disadvantage must both be measured. A unified communiqué solution can be a brilliant industrial service for a production. It acknowledges cost investments as well as an ordinary crossing point for human resources to use. This makes it more creative as the employees do not have to expend time in getting knowledge how to work with dissimilar interface in spite of the machinery they have to use. However, this in turn brings to brain some of the harms connected with this amalgamated communication's approach. It seems that eventually the solution must be software based and rely less on hardware. A stele-communications turn out to be more refined, the interface and the connections must be accessible to all users in spite of what kinds of devices are accessible.

All they want to say is that as the communiqué necessities become more assorted, the hardware must hold those technological software fundamentals. Business enterprises will not be able to change hardware and software or upgrade at the same time. Practically talking, it will become a solid hardware with a software platform that is workable.

Bibliography:

[1]   http://www.brighthub.com/office/collaboration/articles/90048.aspx#imgn_1

[2]   www.webtorials.com/main/resource/papers/.../ucdummesseconded.pdf

[3]   www.cns.vt.edu/docs/UC_NISadvisoryCommittee_presentation.pdf

[4]   www.slideshare.net/sarwarj/unified-communication-presentation

[5]   docwiki.cisco.com/.../Unified_Communications_in_a_Virtualized_Environment

[6]   www.networkworld.com/category/uc-voip

[7]   searchunifiedcommunications.techtarget.com/.../Unified-communications.

[8]   searchunifiedcommunications.techtarget.com › ... › Email and messaging

[9]   https://en.wikipedia.org/wiki/Unified_communications

OPINION ANALYSIS OF TWITTER DATA USING MACHINE-LEARNING TECHNIQUE

Garima saini[1], Shweta Roy[2], Preeti Kaushik[3], Kirti saraswat[4], Pardeep Kumar[5], Sudeshna Chakraborty[6]

[1,3,4,6]CSE Department, Sharda university, JIMS Science and Engg.,Lingayas Vidyapeeth ,Faridabad,India

[2]JBKP CSE / BCA Department, Faridabad

[5]Galgotias University

Garima.saini@jagannath.org[1], sudeshna2529@gmail.com[2]

Abstract: Online Micro blogging on casual associations have been used for exhibiting sentiments about certain substance in very short messages. Existing notable small scale websites like twitter, facebook and so forth, in which twitter accomplishes most prominent proportion of thought in the field of examination locales related to thing, film reviews, stock exchange, etc. The assessment on feeling investigation has been going for a long time. Assumption investigation in present days transforms into the major issue in field of examination and advancement. In view of bit by bit increase in the amount of clients on the long range casual correspondence destinations, huge proportion of data conveys as substance, sound, video and pictures. There is have to do estimation analysis as writings in type of messages or presents on discover the whether the opinion is negative, positive or unbiased. We had separated information from twitter for example movie surveys for feeling forecast utilizing machine-learning calculations. We associated managed AI counts like Support vector machines (SVM), most extreme entropy and Naive Bayes to describe data using unigram, bigram and cross variety for instance unigram + bigram features. Result demonstrates that SVM outperformed different classifiers with striking precision of 87% for filmaudits.

Keywords: Opinion mining, movie reviews, machine learning, SVM

I. INTRODUCTION

Social networks are the gathering of people in explicit gatherings. It might be a political or religious gathering or a gathering of college understudies, youths, all sharing data about their interests, basically on the web. Twitter, MySpace or Facebook are a portion of the free and simple to- get to interpersonal interaction locales. Almost certainly, this connection incorporates kinship, families, aggregate connections, and opinional connections. Interpersonal organizations help individuals make new companions create individual connections and keep in contact with family very effectively. Due to the expansive number of individuals who interface with the system locales, the quantity of connections increments step by step. The long range informal communication highlights consolidated in a site are: client gatherings, most recent data about music gatherings, video and photograph areas, websites, individual profile and the sky is the limit from there. Interpersonal interaction locales additionally help individuals keep up and create business contacts with them. LinkedIn is the best model, as it might be a better than average spot to discuss business and meet specialists. It is simpler and quicker to engage with new business users. Web is the first and the primary correspondence innovation to change the social association of individuals. Since the mid 1990s, Web appropriation has developed quickly. For instance, in 2003, 63% of Americans had utilized the Web. During the 1990s, data innovation specialists anticipated that the Web should be thrownever.

Web has turned into a basic piece of our lives. Numerous sites have offices for individuals to keep in contact as informal communities. Interpersonal interaction destinations are the best approach to communicate with new individuals and make associations, just as offer photographs, recordings and exercises with one another.

Opinion Mining: Sentiment Mining is to choose the assumption of customer identified with some occasion or the declaration depict the sentiment of the customer for instance what he/she feel about it. Customers share the things about their propelling life, talk about current issues and mix of points. Allowed to write in any structure

without keeping concludes that puts this more on the map than dynamically settled blogging regions. Movies and thing reviews viably available now a days or considerations on strict and policy centered issues, so it ends up central wellsprings of customer thought and feeling. Information that we use in our investigation are from twitter, it contains tremendous number of messages by broad number of clients made without any other individual. Messages can change from popular conclusion to individual idea. For instance some post from twitter can be appeared Table 1.1.

These micro blogging destinations are immense wellspring of data and it is very simple to state that there is a need of computerizing the opinion examination process as there is excessively work engaged with preparing this data physically. Different methodologies are drilled for the mechanization of this procedure like machine learning and Normal language handling. Clients are extending bit by bit as the masses and example of using miniaturized scale blogging goals are growing, so the data can be used in investigate purpose behind inclination examination and evaluation mining.

For instance, film creators keen on following inquiries:-

•   What is gathering of spectators want from our film?(whether the film is pleasant or not)

•   How everybody responded to our film?

•   Whether the film is swing to be extraordinary or horrendous?

In this period of race each news channel appear the leave overviews of each ideological gatherings, so every ideological gathering ready to realize what number of are in help and with the help of small scale blogging goals people will give the decisions regarding inclinations of the social occasion. Such assessments will surely help gatherings to build their vote banks.

The data we are using in this assessment are film reviews. We have assembled around 17000 film overviews from the twitter. The film review contains reviews of distinguish motion pictures. Overviews can be requested in three different manner:

1.   Positive examinations: texts in which human beings favored the feature film.

2.   Negative examinations :texts in which human beings not favored the feature film.

3.   Neutral examinations: texts in which human beings doesn't have any inclination or subject to insignificant truth.

Fig 1: Opinion Mining of Social Media

II. LITERATURE SURVEY

Assumption Analysis is the raised examination of how appraisals and points of view can be identify with ones tendency and state of mind appears in conventional language significant regard to an occasion. Consistent occasions show that the assumption mining has come to upto wonderful accomplishment which can beat the positive versus negative and supervise entire field of direct and expressions of love for various frameworks and core interests. In the field of assumption mining utilizing specific strategies mind boggling extent of exploration has been practiced for want for relational associations.

Ache and lee [1] suggest the structure where an inclination can be certain or dimissive was found by extent of constructive words to signify words. Later in 2008 the makers made system in which tweet result can be picked by term in the tweet. Appear differently in relation to baselines that are made by individuals, the results are genuinely extraordinary when AI methodology are used. SVM gave best result as stand out from Naïve Bayes. In spite of using particular sorts of features the makers didn't accomplish needed exactnesses over subject based request.

Jiang *et al.* [2] centrearound target-subordinate opinion grouping. Here target-subordinate methods whether the opinion is certain, negative or unbiased relies upon nature of the inquiry that is inquired. The creators proposed to improve target- subordinate conclusion arrangement by joining highlights of target-subordinate and thinking about related tweets. The creators likewise recommended that there is need of thought current tweets to the related tweets by utilizing chart based streamlining. As guaranteed by creators trial results, the chart based streamlining builds the execution.

Tan et al. [3] said that the clients that regular relative sentiments are likely going to be related. The makers proposed the model that were created from either by following the framework that has been made by marking various customer with the help of "@" or by analyzing the arrangement of twitter supporter/adherent. The makers explained that by using information of association of twitter there will be improvement in customer level notion assessment.

Chen et al. [4] utilized the feed-forward BPN system and utilizations assessment introduction to figure the outcomes at every neuron. The creators proposed a system dependent on neural system. The proposed system is blend of machine learning classifiers and semantic introduction files. So as to get proficiency in strategy, semantic introduction files utilized as contributions for neural system. The proposed procedure beats other neural systems and customary methodologies by expanding proficiency in both preparing just as arrangement time.

Malhar and Ram [5] used oversaw AI methodology and fake neural frameworks to arrange twitter data nearby relevant examination of Presidential and Assembly races which results SVM beats each and every other classifier. The makers proposed a procedure to envision the aftereffect of race results by utilizing the customer sway factor. To finish decline in estimation the makers merged the Principle Component Analysis with SVM.

Anton and Andrey [6] inspected the current strategies and built up a model for programmed estimation examination of twitter messages utilizing unigram, bigram and mutually for example half breed highlight. The motivation behind the creators is to investigate and deliver approaches for breaking down the emphasize of the messages in online networking. The creators looked into existing programmed conclusion examination approaches and so as to keep up the setting of developing techniques the character highlight of internet based life proclamations were considered.

Pak and Paroubek [7] perform phonetic analysis and manufacture a feeling classifier to decide positive, negative and nonpartisan estimations for a report. The creators built up an assumption classifier, which gives unbiased, negative and positive articulations of an archive. So as to prepare opinion classifier the creator proposed a methodology that gathers corpus consequently. So as to break down the uniqueness in dissemination among impartial, negative and positive sets, the creators utilizedTreeTagger.

Kopel and Schler [8] explain that it is basic to use fair-minded messages to get extraordinary data of limit. The maker furthermore communicates that positive and negative messages alone won't give authentic understanding about fair-minded messages. Considering objective messages clear the qualification among positive and negative messages.The makers found that in one of the corpus having most of the impartial reports gives no thought which can be used as counter to test both vitality and opposition of arecord.

Go et al. [9] presented an approach for programmed estimation arrangement of twitter messages. Individual of question term messages were delegated negative or positive. Here creators utilize far off supervision to show the consequences of estimations of twitter posts with the assistance of the machine learning calculations. The calculations, forexample,MaximumEntropy,SVMandNaïve

Bayes are connected to preparing information which contains emojis, gave exactness above 80%. The creators likewise talk about pre-processing steps that was gotten higher precision. The creators concocted a thought for inaccessible managed picking up utilizing tweets that containemojis.

Christianini and Taylor [10] distributed and shared the information about SVM which is machine learning calculation. The creators figure out how to give profound comprehension about calculation and how to approach the SVM calculation so as to actualize it to tackle the down to earth issues. The methodology will be hypothetical as when the book was distributed, the exploration was on going on eachfield.

Burger et al. [11] Since, In this period the PC have turned out to be sufficient amazing that can deal with extensive scale application which gives design acknowledgment and factual estimation of certifiable issues. The creators presented a n approach for measurable demonstrating dependent on most extreme entropy. By utilizing instances of issues in common language preparing, the creator demonstrates most extreme probability system for programmed development of greatest entropy models. Here the creators portrayed the standard of most extreme entropy. This rule chooses the model with most noteworthy entropy among all the reliable models. By augmenting the probability of preparing information we can acquire ideal estimations of givenparameters.

Romero et al. [12] found that hash tags turns into the normal component of twitter utilized in each message and new terms are made and changing on everyday schedule which impacts the general importance of the first term. The creators likewise discovered basic contrast among issues and get familiar with the structure of generally utilized distinctive kinds of hash tags. The creators likewise created generative and recreation based models to examine the connection between plan of most recent adopters on which hash tag extends and appropriationelements.

III.AIMS &OBJECTIVES

The main goal of the thesis is to compare the results that are implemented with the help of supervised classifier

The methodology followed is:

1.   We‟ve ordered an assortment of positive, negative, and impartial tweets utilizing the Twitter4j Twitter API. The size of our body can be very enormous.

2.   Secondly, we have expelled the unfilled words from the ordered corpus with the goal that the substance is liberated from commas, stand-stills, and soon.

3.   Next, we at first apply the AI counts to our readiness set, and a short time later we test and dissect theresultsUsing the outcomes, we assess which programmed learning calculation is best for arranging sentiment investigation.

The objective of the thesis is discussed in the followingpoints:

1.   Explore, analyze and study the existing detection analysis techniques in the online micro blogging network.

2.   Learn how to generate tweets from Twitter using the PythonAPI.

3.   Implement and analyze the results obtained after the application ofthe

4.   Learning classifiers for thedataset.

IV.METHODOLOGY

1.   Data Gathering: We gather information from the Twitter API called Twitter4j utilizing Netbeans. Searches are finished utilizing #Hashtag followed by the name of the film, for example, #FAN, #Gully Boy, # URI-The Surgical Strike, etc. About 23,000 tweets from a few film tweets have been incorporated. Pundits can likewise be found on #Hash marks, trailed by celebrities, chiefs, creation organizations and individual record organizations. On Twitter, hash labels become the image expected to discover something and offer the client a restriction of 120 words to communicate his perspective and disposition.

2.   Normalization: We have found that to get the ideal outcomes from the classifier, we have to ensure that the tweets can be prepared accurately. Since tweets can be in the client's language, we have to tidy up random information. The accompanying things that may not be applicable to the dataare:

3. URL: The URLs of the message will be trivial in light of the fact that they just occupy the outcome from the classifier.

4. Client Name: Deleting the client name might be important for cleaning purposes as it might dishonestly influence our outcomes.

5. RepetitiveWords:

5.Machine learning procedures: We use extremity based order strategies that utilization a lot of positive, negative, and unbiased tweets gave by the Twitter4j API. Extremity is given by the likelihood connection that a word shows up in a lot of positive or negative confirmations that make the word positive or negative. The classifiers we use depend on the idea of extremity.

$$\text{Polarity} = \frac{\text{P(Postive\_Words)/P(Total\_Words)}}{\text{P(Negative\_Words)/P(Total\_Words)}}$$

If the feature is independent and based only on Standard English Dictionary then only this technique works. This method fails when we tried to record the opinion shown with respect to comparison. Further, the polarity based technique also fails to record query related opinion. In order to fulfill the requirement of classification we involved machine learning techniques.

The machine techniques comprised of following supervised classifier that are given below:-

- NaïveBayes

- Support Vector Machines(SVM)

- Maximum Entropy(MaxEnt)

6. SupervisedClassifiers:

Naïve Bayes: The Naïve Bayes classifier in one of the easiest probabilistic model works emphatically on text order and utilized on Bayes rule with self-supporting component assortment [3] works decidedly on text classification and utilized on Bayes rule with self-supporting element assortment [3]. It is adaptable in method of taking care of with any number of classes or traits. For a given tweet d, C* is a class variable which defines the opinion given by $C* = arg(C|D)$

Bayes Probabilty PNB (C|D) described as

$$PNB\ (C|D) = ((c)\Sigma P(f|c)ni(d)\ mi = 1\ )$$

In the event that the character is rehashed more than twice,it may incorporate another word, yet the significance is the equivalent. In this manner, we should take out this word and make it genuine. For instance, you can compose goooooood. In the event that the message contains a word that has showed up more than twice in succession, it must be changed twice. For instance, a decent film can turn into a decent film.

6. Elimination of stop words: Stop words will be words, for example, "an", "is", "el", "and so on.", and so on. These words have nothing to do with feeling, so they should dispose of the message. The subsequent stage is to prepare the information utilizing a regulated classifier.

$P(d)$

Support Vector Machines: SVMs are going on to be very cultivated at text classification, broadly beating Naïve Bayes (Joachims 1998). We analyzed enormous edge classifier to achieve powerful precision of order process [9]. SVMs uses a function called kernel which are machine learning classification methodology in which the data is not

separable linearly in the new area which it is to locate to area of data points, with allocation for classification of erroneous.

Support Vector Machines are the members of the family of classifiers which are linear. The main objective of the linear classifier is to find a hyperplane which is linear in nature of a feature area that divides all other entities in form of two classes. The main function of the SVMs is find out the hyporplane which is separating that has distance maximum from the nearest points to feature area init.

Searching hyperplane in sample of linear separable, the equation can be consider as problem of optimization:

½ ‖ω‖2 → min (ω, b)

yi (wTxi + b) ≥ 1, j = 1, … ,,

Maximum Entropy: MaxEnt is another order procedure broadly utilized part in uses of normal language preparing [10]. MaxEnt not in every case except in some cases beats the Naïve Bayes classifier for content order [11]. MaxEnt is the most uniform model incline toward for the grouping reason [12].

In the scenario of 2-class, to search for distribution over the both classes it is likely the same thing as using the logistic regression. Regarding independence of feature, it does not make any assumption. Due to this we can add features and phrases such as bigrams and to MaxEnt without affecting overlapping of the features. Let"s take an example in which we have two features such as "good" and other one is "very good" , then in case of Naïve Bayes their probabilities will be taken as independent even when the both of this are overlapping but not in case of MaxEnt. The equation for this model can be given as:

P(C|D) = 1 ⁄ Z(d)exp (Σi λi,cFi,c(d,c))

Here, c indicates class, d indicates a single tweet, λ indicates vector of weight,

normalization   function   Z(d),   *Fi,cφ*   is   aclass/feature function for class c and feature *fi* defined as follows:

fi,(d,c′)={1,ni(d)> 0andc′=c0, ot ☐ erwise

Performance Measure: To compute the exactness of classifier we required measure on which precision can be acquired. There are two measures on which accuracy can be dependent:

- Precison

- Recall

- Accuracy

V. IMPLEMENTATION

1.   Installation:

Tweepy: It is the python customer for the official Twitter API. Introduce it utilizing following pip order:

pip install tweepy

TextBlob: It is the python library for handling printed information. Introduce it utilizing following pip order:

pip install textblob Also, we need to install some NLTK corpora using following command:

python -m textblob.download_corpora(Corpora are nothing but a large and structured set of texts.)

2.   Authentication:

So as to get tweets through Twitter API, one needs to enroll an App through their twitter account. Follow these means for the equivalent:

- Open the connection and Create NewApp

- Fill the application subtleties. You can leave the get back to url fieldempty.

- Once the application is made, you will be diverted to the application page.

• Open the „Keys and Access Tokens"tab.

• Copy „Consumer Key", „ConsumerSecret",

„Access token" and „Access Token Secret".

We follow these 3 significant strides in our program:

• Authorize twitter APIclient.

• Make a GET solicitation to Twitter API to get tweets for a particularquery.

3. DataExtraction

As a matter of first importance, we make a Twitter Client class. This class contains all the techniques to communicate with Twitter API and parsing tweets. We useinitfunction to deal with the confirmation of APIclient.

In get_tweets function, we use:

fetched_tweets = self.api.search(q = query, count

= count)

• *to call the Twitter API to fetchtweets.*

• *In **get_tweet_opinion**we use textblob module.*

analysis = TextBlob(self.clean_tweet(tweet)) TextBlob is actually a high level library built over top of NLTK library.

4. Pre processing: - In the first place, we call the clean_tweet strategy to expel joins, exceptional characters, etc. Tweet utilizing straightforward standard articulations.

5. Then, when we pass the tweet to make a TextBlob object, the accompanying procedure is executed on the content utilizing the textbloblibrary:

6. Tokenize the tweet, that is, partition the words in the group of the text.

7. Eliminate the vacant expressions of the chips. (Stop words are regularly utilized words that are not important in the investigation of a book, for example, I am, you, are, and so forth.)

8. Label the tokens (some portion of the voice) and select just huge highlights/tokens, for example, descriptive words, adverbs, etc.

9. Pass the tokens to an assessment classifier that orders the assessment of the tweet as positive, negative, or unbiased by allotting an extremity between - 1.0 and 1.0.

10. Here's the manner by which the emotions classifier is created:

11. TextBlob utilizations a lot of film survey information in which updates have just been labeled as positive or negative.

12. The positive and negative characteristics are removed from every positive and negative examination, respectively.

13. Training information currently comprises of positive and negative labeled highlights. These information are framed in a Naive Bayes classifier.

14. Then, we use sentiment. polarity method of TextBlobclass to get the extremity of tweet between -1   to   1.

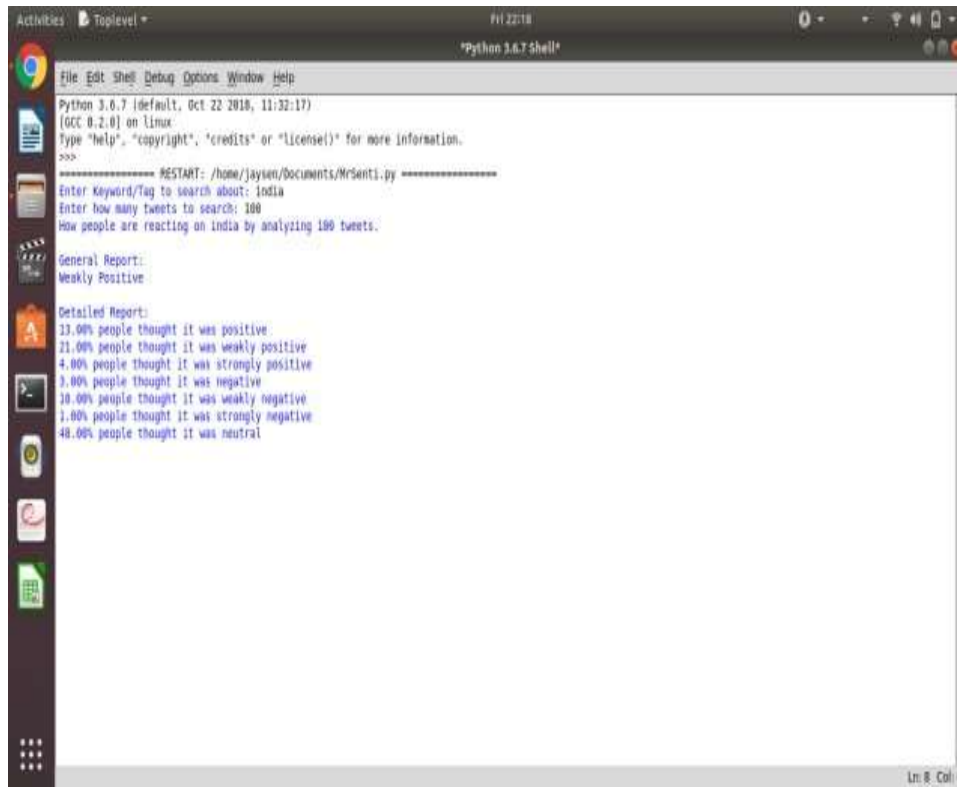At   that   point,   we   group   extremity   as:
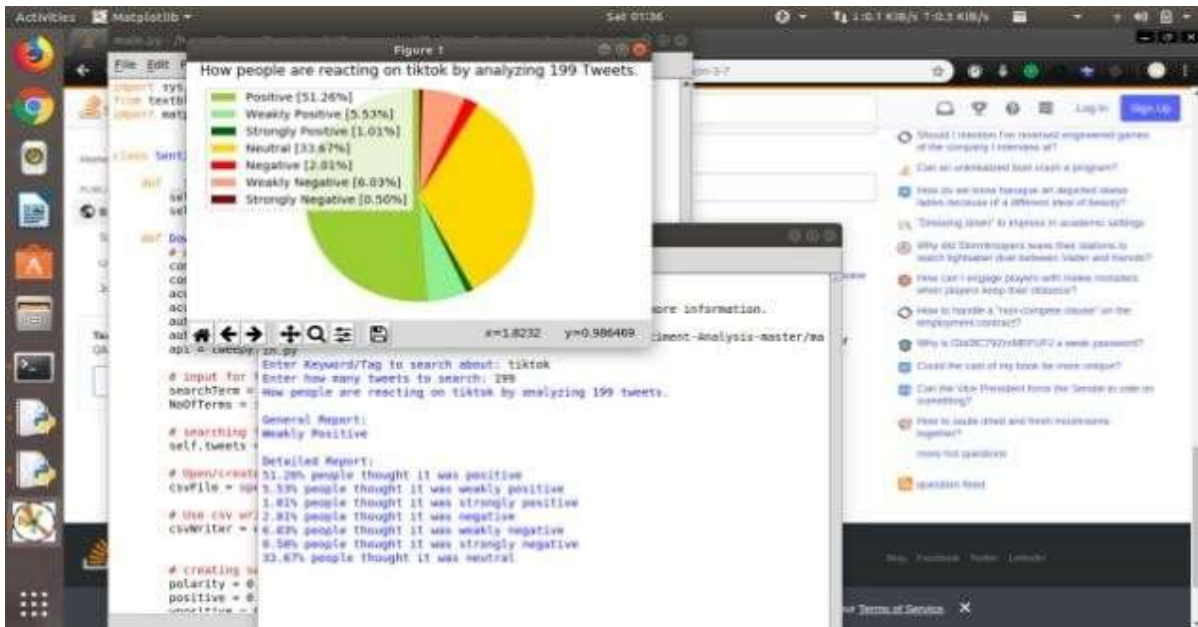
Fig: 1 Initial Console System Interface

Fig 2: During Execution of the algorithm

{

if analysis. sentiment. polarity> 0: return 'positive'

else if analysis. sentiment. polarity == 0: return 'neutral'

else:

return 'negative'}

Finally, parsed tweets are returned. Then, we can do various type of statistical analysis on the tweets. For example, in above program, we tried to find the percentage of positive, negative and neutral tweets about aquery.

VI. RESULTS

We are utilizing Python language for execution. Python language offers most extreme help with regards to AI procedures. AI methods can be effortlessly executed in Python language.
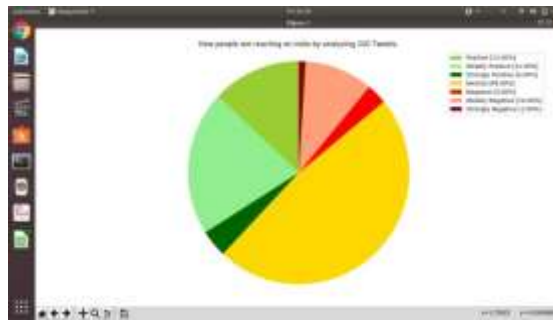
Fig 3: Results of Classification of Tweets

Table 1: Precision and recall

Algorithm	Performance Measure	
	Precision	Recall
Naïve Bayes	0.78	0.76
Support Vector Machines	0.87	0.73
Maximum Entropy	0.74	0.71

VII.   CONCLUSION

In this proposition, we played out a similar investigation of administered classifiers, for example, Naïve Bayes, bolster vector machines, and most extreme entropy utilizing the unigram, bigram, and half and half capacities (unigram + bigram). It is important to lead assessment dissects as writings as messages or distributions to decide if the sentiment is negative, positive or impartial. We had removed information from Twitter, that is, film audits for the forecast of feelings utilizing AI calculations. We originally separated the twitter information utilizing the twitter API. At that point, during preprocessing, we tidy up the information and make it accessible for preparing utilizing classifiers. We aggregated 15,000 tweets for the preparation set and 2,000 tweets for the test set. The SVM that utilizes the cross breed work outperforms every single other classifier and the determination work

with an exactness of 87%, while Max Ent surpasses Naïve Bayes with the bigrama work. MaxEnt, in some datasets, gives preferred outcomes over Naïve Bayes. It is inferred that SVM performs superior to other classifiers.

VIII. REFERENCES

[1] **B**. Pang, L. Lee and S. Vaithyanathan, "Thumbs up: opinion classification using machine learning techniques", *Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10*, pp. 79--86,2002.

[2] L. Jiang, M. Yu, M. Zhou, X. Liu and T. Zhao, "Target-dependent twitter opinion classification", *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, vol. 1, pp. 151--160, 2011.

[3] C. Tan, L. Lee, J. Tang, L. Jiang, M. Zhou and P. Li, "User-level opinion analysis incorporating social networks", *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 1397--1405,2011.

[4] L. Chen, C. Liu and H. Chiu, "A neural network based approach for opinion classification in the blogosphere", *Journal of Informetrics*, vol. 5, no. 2, pp. 313-322,2011.

[5] M. Anjaria and R. Guddeti, "Influence factor based opinion mining of Twitter data using supervised learning", *Communication Systems and Networks (COMSNETS), 2014 Sixth International Conference on*, pp. 1--8,2014.

[6] A. Barhan and A. Shakhomirov, "Methods for Opinion Analysis of Twitter Messages", *12th Conference of FRUCT Association*,2012.

[7] A. Pak and P. Paroubek, "Twitter as a Corpus for Opinion Analysis and Opinion Mining.",*LREc*, vol. 10, pp. 1320-- 1326,2010.

[8] M. Koppel and J. Schler, "THE IMPORTANCE OF NEUTRAL EXAMPLES FOR LEARNING OPINION", *Computational Intell*, vol. 22, no. 2, pp. 100-109,2006.

[9] A. Go, R. Bhayani and L. Huang, "Twitter opinion classification using distant supervision", *CS224N Project Report, Stanford*, vol. 1, p. 12,2009.

[10] A. Andrew, "An Introduction to Support Vector Machines and Other Kernel-based Learning Methods20016Nello Christianini and John Shawe-Taylor. An Introduction to Support Vector Machines and Other Kernel-based Learning Methods. Cambridge University Press, 2000. xiii + 189 pp., ISBN: ISBN 0-521-78019-5 Hardback: £27.50", *Kybernetes*, vol. 30, no. 1, pp. 103-115,2001.

[11]A. Berger, V. Pietra and S. Pietra, "A maximum entropy approach to natural language processing", *Computational linguistics*, vol. 22, no. 1, pp. 39--71, 1996.

[12]K. Nigam, J. Lafferty and A. McCallum, "Using maximum entropy for text classification", *IJCAI-99 workshop on machine learning for information filtering*, vol. 1, pp. 61--67, 1999.

[13]D. Romero, B. Meeder and J. Kleinberg, "Differences in the mechanics of information diffusion across topics: idioms, political hashtags, and complex contagion on Twitter", *Proceedings of the 20th international conference on World wide web*, pp. 695--704, 2011.

[14]S. Tan and J. Zhang, "An empirical study of opinion analysis for Chinese documents", *Expert Systems with Applications*, vol. 34, no. 4, pp. 2622-2629, 2008.

[15]J. Martineau and T. Finin, "Delta TFIDF: An Improved Feature Space for Opinion Analysis", *Proceedings of the Third International ICWSM Conference*, vol. 9,2009.

[16] FA. Nielson. "A new ANEW: Evaluation of a word list for opinion analysis in microblogs", *arXiv preprint arXiv:1103.2903,*2009.

[17] SM. Mohammad, S. Kiritchenko and X. Zhu, "NRC- Canada: Building the State-of- the-Art in Opinion Analysis of Tweets", *arXiv preprint arXiv:1308.6242,*2013.

[18]E. Kouloumpis, T. Wilson and J. Moore, "Twitter Opinion Analysis: The Good the Bad and the OMG!", *ICWSM*, vol. 11, pp. 538—541,2011.

[19]K. Denecke, "Using SentiWordNet for Multilingual Opinion Analysis", *Data Engineering Workshop, 2008. ICDEW 2008.IEEE 24th International Conference on IEEE,* pp. 507—512, 2008.

[20]B. Gokulkrishnan, P. Priyanthan, T. Ragavan, N. Prasath and A. Perara," Opinion Mining and Opinion Analysis on a Twitter Data Stream", *Advances in ICT for Emerging Regions (ICTer), 2012 International Conference IEEE,* pp. 182—188, 2012.

[21]F. Neri, C. Aliprandi, F. Capeci, M. Cuadros and T. By," Opinion Analysis on Social Media ", *Proceedings of the 2012 International Conference on Advances in Social Networks Analysis and Mining,* pp.919—926, 2012.

[22]T. Wilson, J. Wiebe and P. Hoffman, "Recognizing Contextual Polarity: An Exploration of Features for Phrase- Level Opinion Analysis", *Computational Lingusitics, MIT Press,* vol. 35, no. 3,pp. 399—433, 2009.

[23]N. Godbole, M. Srinivasaiah and S. Skiena," Large-Scale Opinion Analysis for News and Blogs", *ICWSM,* vol. 7, no. 21, pp. 219—222,2007.

[24]F. Benamara, C. Caserano, A. Picariello, DR. Recupero and VS. Subrahmanian," Opinion Analysis: Adjectives and Adverbs are better than Adjectives Alone", *ICWSM,* 2007. [25]D. Boyd and N. Ellison," Social Network Sites: Definition, History, and Scholarship", *IEEE Engineering Management Review,* vol. 3, no. 38, pp. 16—31,2010.

[26]A. Agarwal, B. Xie, I. Vovsha, O. Rambow and R. Passonneau, "Opinion Analysis of Twitter Data", *Proceedings of the workshop on languages in social media, Association for Computational Linguistics,* pp. 30—38,2011.

[27]T. Nasukawa and J. Yi, "Opinion Analysis: Capturing Favorability Using Natural Language Processing", *Proceedings of the 2nd international conference on Knowledge capture, ACM,* pp. 70—77,2003.

[28]H. Wang, D. Can, A. Kazemzadeh, F. Bar and S. Narayanan," A System for Real- time Twitter Opinion Analysis of 2012 U.S. Presidential Election Cycle", *Proceedings of the ACL 2012 System Demonstrations, Association for Computational Linguistics,* pp. 115—120, 2012.

[29]T. Wilson, J. Wiebe and P. Hoffman,"Recognizing Contextual Polarity in Phrase- Level Opinion Analysis", *Proceedings of the conference on human language technology and empirical methods in natural language processing, Association for Computational Linguistics,* pp. 347—354, 2005.

[30]H. Kanayam and T. Nasukawa, " Fully Automatic Lexicon Expansion for Domain- oriented Opinion Analysis", *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics,* pp. 355—363,2006.

[31]Y. Choi and C. Cardie, "Learning with Compositional Semantics as Structural Inference for Subsentential Opinion Analysis", *Proceedings of the Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics,* pp. 793—801, 2008.

[32]P. Melville, W. Gryc and RD. Lawrence, " Opinion Analysis of Blogs by Combining Lexical Knowledge with Text Classification", *Proceedings of the 15thACM SIGKDD international conference on Knowledge discovery and data mining, ACM,* pp. 1275—1284, 2009.

[33]G. Paltoglou and M. Thelwall, "A study of Information Retrieval weighting schemes for opinion analysis", *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics, Association for Computational Linguistics,* pp. 1386—1395,2010.

[34]J. Fernandez, Y. Gutierrez, J. Gomez and P. Martinez- Barco, " GPLSI: Supervised Opinion Analysis in Twitter using Skipgrams", *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014), number SemEval,* pp. 294—299, 2014.

[35]T. Mullen and R. Malouf, "A Preliminary Investigation into Opinion Analysis of Informal Political Discourse", *AAAI Spring Symposium: Computational Approaches to Analyzing Weblogs,* pp. 159—162,2006.

[36]A. Harb, M. Plantie, G. Dray, M. Roche, F. Trousset and P. Poncelet, "Web opinion mining: How to extract opinions from blogs?",*Proceedings of the 5th international conference on Soft computing as transdisciplinary science and technology, ACM,* pp. 211—217,2008.

[37]SM. Kim and E. Hovy, "Determining the Opinion of Opinions", *Proceedings of the 20th international conference on Computational Linguistics, Association for Computational Linguistics,* 2004.

[38] A. Martalo, N. Novielli and F. De Rosis," Attitude Display in Dialogue Patterns",*Proc. AISB 2008 Symposium on Affective Language in Human and Machine,* 2008.

[39] E. AL-Daoud, "Integration of Support Vector Machine and Bayesian Neural Network for Data Mining and Classification", *World Academy of Science, Engineering and Technology,* vol. 64, 2010.

[40] K. Yessenov and S. Misailovic, "Opinion Analysis of Movie Review Comments", *Methodology,* pp. 1—17, 2009.

Meta Search Engine using Semantic Similarity and Correlation Coefficient

Naresh Kumar, Deepak Sharma, Nripendra Narayan Das

Abstract- This paper aims to provide an intelligent way to query and rank the results of a Meta Search Engine. A Meta Search Engine takes input from the user and produces results which are gathered from other search engines. The main advantage of a Meta Search Engine over methodical search engine is its ability to extend the search space and allows more resources for the user. The semantic intelligent queries will be fetching the results from different search engines and the responses will be fed into our ranking algorithm. Ranking of the search results is the other important aspect of Meta search engines. When a user searches a query, there are number of results retrieved from different search engines, but only several results are relevant to user's interest and others are not much relevant. Hence, it is important to rank results according to the relevancy with user query. The proposed paper uses intelligent query and ranking algorithms in order to provide intelligent meta search engine with semantic understanding.

Keywords: Meta Search Engine, Scrapping unit, Intelligent Query System, Rankings

I. INTRODUCTION

A human can see and understand words and sentences. But to make the words and sentences understandable by machines, these have to represent them in form of vectors. A machine will be able to get similar words based on particular word by representing the words as vectors. A machine will be able to do so by using some similarity metrics that can be cosine similarity or dot similarity. There are various pre-trained models that can be used to convert word into vectors like word2vec and glove2vec.

These are the models which can convert word to vectors. But these models can't give us the encodings for the sentences directly. Author find encoding for the sentences indirectly by finding out the vector of each word in the sentence using either of these models using word2vec or glove2vec and averaging them.This will give us the vector which will show the meaning of sentence in a vector form.

But here is one major problem and that is while finding the vector form of sentence author have not considered the word semantic order inside the sentence. This semantic word order is very important in consideration for conversion into vector form of sentences. For example there are 3 sentences:
1. How old are you?
2. What is your age?
3. How are you?

Naresh Kumar, Associate Professor, Department of CSE, Maharaja Surajmal Institute of Technology, New Delhi, India. E-mail: narsumsaini@gmail.com
Deepak Sharma, Assistant Professor, Department of IT, Jagannath International Management School, Vasant Kunj, New Delhi, India, E-mail: deepaktech@hotmail.com
Nripendra Narayan Das, Associate Professor, Department of Information Technology, Manipal University, Jaipur, India. E-mail: nripendradas@gmail.com

It is obvious that 1 and 2 are semantically the same even though 1 and 3 have more common words. A good sentence encoder will encode the three sentences in such a way that the vector for 1 and 2 are closer to each other than 1 and 3. So to eliminate this problem Universal Sentence Encoder by Google which is a very powerful model for encoding the sentences which will preserve it's semantic word order.

Universal Sentence Encoder is able to embed not only words but also phrases and sentences. It takes variable length English text as input and outputs a 512-dimensional fixed vector.

II. LITERATURE SURVEY

Authors of [1] consider the architectures and features of Meta Search Engines (MSE) for extracting documents from one or more domains on the web. It analyzes two MSE i.e. general MSE and Special purpose MSE. Authors of [2] proposed a MSE on the basis of clustering and ranking to find the relevant results. It have user interface, relevancy calculator, cluster generator and webpage adjuster as a main modules. It took top 10 results Search Engine (SE) –Bing, Google, and Alta Vista) and tested the proposed work on 30 different queries. Traditionally, MSE matches the query with the webpage and then provide the result which is as keyword matching. According to [3], a MSE uses interfaces of self forward query to SEs produce its results from the Internet. MSE consume input form end user and simultaneously send out this query to multiple SEs for getting results. Then it formats the received result and present it to end user.Set of keywords in a single query and Word Net ontology is used to provide the most suitable query to SE through the MSE is proposed in [4]. Some authors used optimization techniques [5], some are using query semantics and some are using, clustering techniques [6] to improve the architectures of MSEs but still they have challenges [7]

III. ISSUES AND CHALLENGES

Problem in [4] is that the author is making the query set by extracting the similar words to the given query using Word-net Ontology. Problems is that:
A. **Large number of permutations with synonyms**: Taking the words from the sentence, finding the synonyms of words and making a query set using synonyms of words. The number of sentences formed by using these synonyms will be very high. Permutations of synonyms to make the sentence is very high.

Meta Search Engine using Semantic Similarity and Correlation Coefficient

B. Semantic issues: By taking the synonyms of words and put them in the sentence somehow is not semantically correct. Every word has its own meaning in a sentence. Proper words should be used while making sentence.

C. Lack of sentence level work: Working with words in not very efficient way of dealing with queries as it is hard to maintain the semantically correct order in order to make a meaningful sentence. Work should be done directly on sentences in order to maintain the word order.

D. Result Aggregation: The extraction of results from different Search engines is not efficient, but also involves the concept of result merging, i.e., removing the duplicate results and irrelevant results or sponsored results. Thus, an efficient algorithm is needed for merging results from a no of search engines.

IV. OBJECTIVES

The proposed MSE has the following objectives over existing algorithms:

(i) Author will be use Universal Sentence Encoder Model instead of Word2Vec model. This model will help us to get similar semantic sentences. We will get the similar queries for a given query using Universal Sentence Encoder model. This helps us to maintain word semantic order in the sentence.

(ii) Proposed system will rank pages according to the correlation between content of pages and the original query.

V. PROPOSED ARCHITECTURE

The architecture of proposed MSE is shown in Figure 1 and it is explained module or block wise.

A. Intelligent Query System: This block will take the input query from the user and gives corresponding similar queries so as to get more accurate and relevant results. This Intelligent Query System will contain Universal Sentence encoder model which will be used to generate the encodings of the input query and this encoding will be compared with all the encodings of the questions using correlation coefficient with the threshold value of 0.8. After comparing the threshold values we will retrieve top most 2 questions which has highest correlations with the user's query.

B. Scraping Unit: Scraping Unit consists of 3 search engines (Google, Bing and Yahoo). For a given query we retrieve top 3 results from each search engine and remove duplicates of results if any and then return the corresponding results.

C. Ranking of Results: Here we will rank the pages with the given user input according to the relevancy. This will be achieved by finding out the relevancy using correlation coefficient between the content of pages and given input query.

D. Results shown to user: Here we finally show the results to the user after ranking the pages.
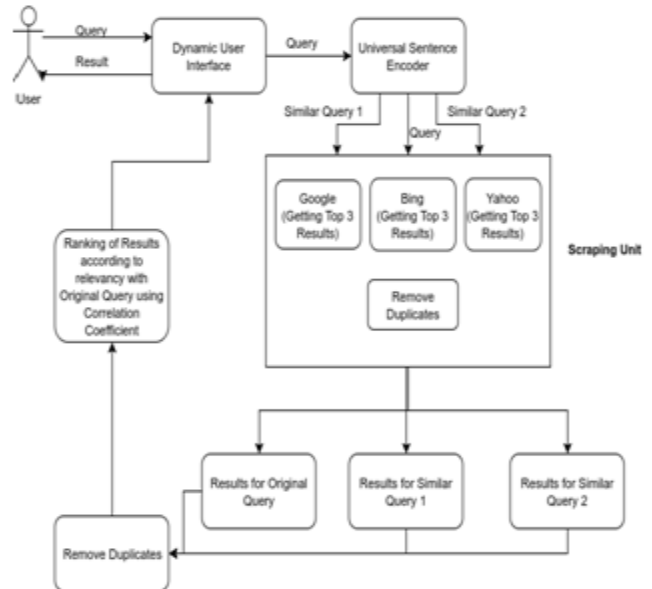

Fig. 1. Proposed Architecture

VI. IMPLEMENTATION AND RESULT DISCUSSION

We have used python for backend where we are using tensor-flow library for universal sentence encoder model. First the user will write down the input query on the frontend. After submitting the query, the query will be send to the server. Inside server, this input query will be fed to the universal sentence encoder model. Universal Sentence Encoder model will generate encoding of input query which is a 512 dimensional vector. These encoding is compared with the encodings of questions that are already stored in the database. We are comparing the 2 encodings of 2 sentences using correlation distance factor. Here we have taken threshold value for correlation distance factor to be 0.8. After comparing all the questions inside the database with the current original query, we will take the top 2 questions to the given input query using this correlation factor. Now we have a total of 3 queries, one is input query and the other two are similar queries. Now all these queries will be fed to the Scraping unit one by one. Scraping unit is a unit consisting of 3 search engines (Google, Bing, Yahoo). Each search engine will give us top 3 results which gives a total of 9 results for 1 query. Now we will remove the duplicate pages from these 9 results. So basically, we are sending all 3 queries one by one into the scraping unit which will give us the results from each query. Now these results also may have duplicates so we have removed the duplicates from here. So after this step we get all the results which we have to show to the user, but one last thing remains is to do ranking of each page. We have send the all the result pages along with its content and the original query to the ranking unit. In ranking unit all the results are ranked according to their relevancy with the given actual query. Relevancy is calculated by taking the content of the page which is converted to the corresponding 512 dimensional vector using universal sentence encoder and finding out the correlation distance factor between original query encoding vector and the content page encoding vector.

We will do this for all the result pages and sort pages according to the correlation distance factor. After sorting results will be shown to the users.

Dataset used & Correlation Distance Factor: This paper used quora questions pair's dataset for finding out the similar queries. This dataset can be taken from the kaggle website. This dataset was published on 2016. There are over 400,000 lines of potential question duplicate pairs. This makes up total of 800,000 questions. After this each question is input to our universal sentence encoder model to get the corresponding 512 dimensional encoding vector which is stored in the database. It is a correlation coefficient between 2 vectors which will tell us how much these vectors are similar to each other. This factor will give us a value in between (0-1). If this factor is approaching to 1 then that means 2 vector are very similar to each other, otherwise they are not. Usually we can have to take a threshold value for this factor which can be 0.9 or 0.8 to distinguish between similar and not similar vectors. Here, in our implementation we have taken the threshold value to be 0.8. Universal Sentence Encoder will give us the encoding of vectors which is a fixed dimensional vector of size 512 irrespective of the size of input, that means whether the word, sentence or even paragraph is fed into the encoder it will give us a fixed length vector which will explain the input (of any length) in 512 dimensional space. With the help of this we can easily find the relevancy between the queries and between the query and pages. Output of Encoder which is 512 fixed dimensional vector will help us to main word semantic order in sentence.

Results and Discussion: We have taken 5 queries from 5 domains (Person, Music, Education, Social Media, General). Queries that we have taken are:

(i)   Who is Barack Obama? (Person)
(ii)   Which are the best rap songs? (Music)
(iii)   What is Web Applications? (Education)

How can we use twitter for money? (Social Media) Is creativity important? (General) Results for each query are shown in the table below. Table contains 4 columns (query, similar queries, results, relevancy score for each result). Each column is described below in brief:
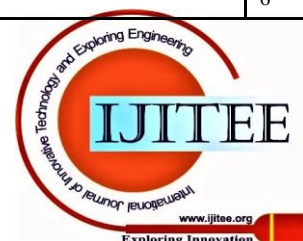
(i)   Query: this is the query input to the MSE.
(ii)   Similar Queries: Similar queries of a given Query from database.
(iii)   Results: Final extracted resultant pages
(iv)   Relevancy Score: Score that represents correlation between original query and contents of resultant links.

Table 1 shows the following:

(i)   Shows the results for the query "Who is Barack Obama?"
(ii)   Shows the results for the query "Which are the best rap songs?"
(iii)   Shows the results for the query "What is Web Applications?"
(iv)   Shows the results for the query "How can we use twitter for money?"
(v)   Shows the results for the query "Is creativity important?"

S.No.	Query	Similar Queries	Result	Relevancy Score
Table I: Results for different queries with relevancy scores				
1	Who is Barack Obama?	Who is the president of America?'	https://www.biography.com/people/barack-obama-12782369	0.41393143
		Who is the President of America now?	https://www.history.com/topics/us-presidents/barack-obama	0.41218776
			https://barackobama.com/	0.3825521
			https://en.wikipedia.org/wiki/Barack_Obama	0.2819289 6
			https://en.wikipedia.org/wiki/President_of_the_United_States	0.2513864 6
			https://en.wikipedia.org/wiki/Bill_Clinton	0.2350767 1
			https://en.wikipedia.org/wiki/Presidency_of_Donald_Trump	0.2248516 7
			https://www.britannica.com/biography/Barack-Obama	0.2017445 6
2	Which are the best rap songs?	What are some good rap songs to dance to?	https://www.thetoptens.com/rap-songs/	0.6975456 5
		What are some of the best rap songs?	https://digitaldreamdoor.com/pages/best_rap-songs.html	0.5860768
			https://www.funadvice.com/q/dance_148283	0.5260219
			https://www.thoughtco.com/best-rap-songs-of-all-time-2857834	0.4922760 4
			https://www.thoughtco.com/top-rap-songs-of-the-90s-2858039	0.4637395 7
			https://www.thoughtco.com/best-hip-hop-dance-songs-2858028	0.4627532 4
			http://www.madmanmike.com/dance_songs.html	0.4337399 6

| #	Original Query	Modified Query	URL	Score
			http://colemizestudios.com/blueprint-for-writing-rap-songs/	0.42528862
3	What is Web Applications?	What is web application?	https://rubyonrails.org/	0.5580404
		What is the web application framework?	https://en.wikipedia.org/wiki/Web_application	0.5060072
			https://en.wikipedia.org/wiki/Web_framework	0.44203725
			https://docs.microsoft.com/en-us/aspnet/whitepapers/add-mobile-pages-to-your-aspnet-web-forms-mvc-application	0.441998808
			https://en.wikipedia.org/wiki/Web_application_framework	0.44408918
			https://searchsoftwarequality.techtarget.com/definition/Web-application-Web-app	0.43899176
			https://www.geeksforgeeks.org/top-10-frameworks-for-web-applications/	0.42710933
			https://searchsoftwarequality.techtarget.com/.../Web-application-Web-app	0.4082357
4	How can we use twitter for money?	How do I use Twitter as a business source?	https://www.jeffbullas.com/37-ways-to-use-twitter-for-business/	0.48076215
		How can I use Twitter for business?	https://www.socialmediaexaminer.com/how-to-use-twitter-for-business-and-marketing/	0.47589707
			https://www.businessnewsdaily.com/7488-twitter-for-business.html	0.46570903
			https://www.lifehack.org/articles/money/7-creative-and-effective-ways-make-money-twitter.html	0.388535955
			https://analytics.twitter.com/	0.36602813
			https://www.youtube.com/watch?v=t9Si-UEczvM	0.36224222
			https://www.wikihow.com/Earn-Money-Using-Twitter	0.33447188
			https://business.twitter.com/en/basics/intro-twitter-for-business.html	0.32550812
5	Is creati...vity important?	Is Creat...ivity Important?	https://www.elitedaily.com/money/entrepreneurship/creativity-important-quality	0.5917218
		Why is creativity important?	https://www.linkedin.com/pulse/importance-creativity-innovation-business-siyana-sokolova	0.39924747
			https://alistemarketing.com/blog/reasons-why-creativity-is-important-to-decision-making/	0.29024532
			https://en.wikipedia.org/wiki/Creativity	0.28962886
			http://www.edudemic.com/creativity-in-the-classroom/	0.25812262
			http://www.innovationmanagement.se/imtool-articles/why-diversity-is-the-mother-of-creativity/	0.24503651
			https://tscpl.org/art/why-is-creativity-important-in-everyday-life	0.20352465
			https://www.dictionary.com/browse/creativity	0.11143326
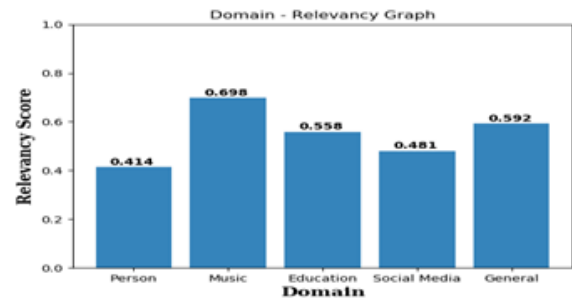
Fig. 2. Domain vs. Relevancy scores

VII. CONCLUSION

In proposed MSE authors have minimized the issues of intelligent query system by using Universal Sentence Encoder instead of Word2Vec model. This will help us to get similar semantic sentences. This helps us to maintain word semantic order in the sentence. In this MSE extraction of similar queries to the original query using correlation factor and then get the results from 3 search engines (Google, Bing, Yahoo) for each query. After getting results, ranking is performed according to the relevancy score. Relevancy score is calculated using correlation between contents of pages and the original query.

REFERENCES

1. A. Madhavi and K. Harisha Chari, "Architecture Based Study of Search Engine and Meta Search Engines For Information Retrieval", International Journal of Engineering Research & Technology (IJERT), Vo.2, Issue 5, ISSN: 2278-0181, May – 2013
2. N. Kumar and R. Nath, "A Meta Search Engine Approach for Organizing Web Search Results using Ranking and Clustering," International Journal of Computer, vol. 10, issue. 1, pp.1-7, ISSN: 2307-4531, 2013.

3. The Wikipedia website [Online]. Available: http://en.wikipedia.org/wiki/Meta-search_ engine at 4:00PM (IST) on23/09/2018.
4. V. Sivakumar, "Semantic Meta Search Engine Using Semantic Similarity Measure", International Journal of Information System and Engineering, Vol. 3, ISSN: 2289-7615, November, 2015
5. Mr. K. P.Raghuvashi, "An Empirical Study on the Meta- Search Engine Optimization Technique Based on Keyword: A Review", IBMRD's Journal of Management and Research, Vol. 3, Issue-2, ISSN: 2277-7830, September 2014
6. N. Kumar, Sonali and S. Gupta, "Non Overlapping Clustering based Meta Search Engine", International Journal on Future Revolution in Computer Science & Communication Engineering, ISSN: 2454-4248, Volume: 3 Issue: 8, 172 – 177.
7. The Bing website [Online]. Available: https://blogs.bing.com/search-quality insights May-2018/Towards-More-Intelligent-Search-Deep-Learning-for-Query- Semantics at 4:00PM (IST) on 23/09/2018.

AUTHORS PROFILE

Naresh Kumar holds a Ph. D. from Kurukshetra University, Kurukshetra and M. Tech. (Computer Science & Engineering) degree from YMCA University of Science and Technology, Faridabad. He is currently working as an associate professor at Maharaja Surajmal Institute of Technology, New Delhi since 2011. His area of research interest includes web crawlers, search engines and meta search engines. He has published more than 44 research papers in reputed journal and conferences.

Deepak Sharma is currently working as an Assistant Professor in the Department of Information Technology, Jagannath International Management School, Vasant Kunj, New Delhi. I am pursuing his PhD(CS) from Jagannath University, Jaipur. Also obtained M.Tech(CS) and MCA degree. I am having more than 14 years of experience in academics and industry. Main interest areas include Programming, Data Structure, Data Mining and Machine Learning, Search Engines and Web Crawlers. During the tenure of my service, I hold various important positions and proved myself as valuable asset for the organization. I was awarded "Best Employee Award" by the institute & also received appreciation letter from the Education department of Kandahar province, Afghanistan.

Nripendra Narayan Das received PhD from Gautam Buddh University, UP, India. Currently He is working as an Associate Professor in Department of Information Technology, Manipal University Jaipur, India. He has published more than 30 papers in National and International Journals. He has more than 24 Years of experience in Industry as well as in teaching.

Technological Convergence: A Study Based on Policy Challenges in Indian Communication Industry

Shruti Chopra

Assistant Professor
Jagannath International Management School
Affiliated to G.G.S.IP.U, New Delhi

Abstract :　The aim of the present paper is to analyze the technological convergence of digital media as a part of industrial restructuring and thereby, to get the insight of related challenges as far as policy formulation is concerned. The paper aims to look at the evolvement of communication technologies in the history of India and ongoing convergence due to digital platforms; the study attempts to analyze the role of regulatory bodies in the regulation of the ongoing scenario and thus, will look at the need for integrated approach required to regulate ICT and Broadcasting in India. The study aims to understand the convergence in Indian context and the policy challenges related to the convergence of technology. It will attempt to suggest the appropriate regulatory choices in Indian context.

IndexTerms - **Convergence, Regulation, Co- Regulation, Self – Regulation**.

I. INTRODUCTION

India had entered into the age of electronic media during the British Raj with terrestrial radio broadcasting in 1920s but after independence, the AIR network had only six stations: Delhi, Bombay, Calcutta, Madras, Lucknow, and Tiruchi; the total number of radio sets at that time were about 275,000. Television broadcasting began in Delhi in 1959 as part of AIR, but was separated from radio as Doordarshan on April 1, 1976. During that period India took up the experiment called Satellite Instructional Television Experiment (SITE) for one year, i.e. from August, 1975 - 1976 in collaboration with United States to meet various technical and development objectives. This was the first time when India broadcasted through the Satellite using United States ATS – 6 spacecraft and later moved to have its own broadcasting satellite INSAT, commissioned in 1983. It was a joint venture of Doordarshan, All India Radio, India Meteorological Department, Department of Space and Department of Telecommunication. So, the need for the regulation of industrial convergence (across various sectors) can very well be traced back in 1980s.

The early decade of 1980s saw remarkable development in Television broadcasting. National telecasts were introduced first time in 1982 and the same year colour TV was introduced in the Indian market with the live telecast of the Independence Day speech by the then prime minister Indira Gandhi on 15 August 1982, followed by the 1982 Asian Games which were held in Delhi. On the other hand, Indian Telecommunication services, which began in 1965, were going through various changes with the establishment of Mahanagar Telecom Nigam Limited (MTNL) in 1986. Since then, regulation of services has become a matter of consideration among the public and private stakeholders, which will be discussed in the later sections of this paper. Moreover, in order to discuss the existing policies and the policy measures for future, it is important to understand the regulatory aspects of past.

II. REVIEW OF LITERATURE

The term convergence has become a buzzword to indicate the developments that are taking place globally, combining the various aspects which were once separate. Mark Hukill (2000) puts forward that since 1980s, the ubiquity of computer technology and the advancement of digital technology have accelerated the convergence of the sectors which were previously separating the communication industry. Moreover, organizational orientations have begun to converge not only cooperatively but through new alliances and mergers, and as well as competitively through the intrusion into each other's market. Hukill (2000) has seen the convergence at three levels, a) at the level of technological developments, b) industrial restructuring and c) convergence in the regulatory measures. He has given a broad and inclusive definition of convergence from the technological perspective as 'the overlapping and ubiquitous use of computer systems, network infrastructure and other electronic communication systems and devices for the production, organization, storage, retrieval, interlinking and dissemination of digitally encoded information (text, audio and visual) in all communication industry domains including, but not limited to broadcasting, and cable media, print media as well as telecommunications and information technology' (p.15).

Michael Latzer (2009) has also pointed out that convergence can take place at various levels and the leading role is played by the technological convergence, which he defines as 'a universal digital code, for common (IP) protocols, which are used for different technological (hybrid) platforms (fixed-wire and mobile communication, WLAN, broadcasting, wi-max and so on). Convergence creates a 'digital modular construction system', which offers great flexibility for innovatively assembled services. This can lead to service-integrating devices, such as TV-capable mobile phones. Convergence has also resulted into increased flexibility on the supply side, and to increased product variety since the previously rigid combination of technology and content (services) has been dissolved (p.414). It would be inappropriate and misleading to reduce convergence to technological level alone, rather technological convergence is leading to the integration of services as well. This type of convergence will be the focus of discussion in this paper. Further, Latzer has talked about corporate convergence, that is, the same companies are now active in both sectors and in the third sector, the internet – keyword triple play and he gives the example of the newly emerging enterprises as Google, ebay.[1] Latzer has talked about social-functional convergence, wherein, telecommunication is now increasingly being used in the private-entertainment sector and broadcasting is increasingly used for business communication (e.g. Internal corporate business TV). At

---

[1] These enterprises have emerged as consequence of convergence and Latzer has called them as convergence enterprises.

the same time, shifts, substitutions and combinations in the application of services are taking place. This is also called receptive convergence, since it is about the change in reception patterns, a convergence of usage patterns (Latzer 2009, pp. 14 – 15).[2]

Lastly, Spatial Convergence can be defined as the globalizing effect of ever-increasing cross-border services and uniform technology – as well as a regulatory convergence, affecting the coordination and integration of regulatory systems for media and telecommunications. Considering all these aspects of convergence, Latzer has coined a term, "mediamatics" by which he means 'the computer sector serves as connector between the formerly separate sub-sectors of telecommunications and the mass media. Seen chronologically, convergence has taken place in two steps - data communication and the digitalization of telephony, which marked the arrival of computer technology (informatics) into telecommunications (= telematics). This pattern has been followed till the end of the 20th century by convergence of mass media with telematics (=mediamatics)' (Latzer, 2009, p.415).

Mosco and Mckercher (2006), locating the term convergence as integration of technologies, arenas and institutions, has defined the technological convergence as ' the integration of devices that the industries use as well as information they process, distribute and exchange over and through these devices' (p.734). According to them, by integrating computer and telecommunication, internet has become an iconic example of technological convergence and is also responsible for the convergence of once separated industries into common arena providing electronic information and communication services.

Therefore, while defining various types of convergence, the study is limiting the discussion by broadly classifying the convergence into two categories - a) Convergence due to the technological developments and b) Vertical and Horizontal convergence of the services which is taking place due to convergence of technological developments. This study is not only limited to see the internet as convergence between the telecommunication and computer, but has rather gone far beyond that to illustrate the implications of the more recent examples in Indian context. As far as convergence of technology is concerned, the best example to illustrate would be a mobile handset which is having features of camera, radio, and mp3 player. Thus, the convergence of technology has made devices more compact and advantageous.

Due to the technological developments, the convergence is taking place in terms of services which requires an in-depth analysis in order to come to any conclusion regarding the regulatory choices to have optimum policy outcomes. But prior to dealing with possible regulatory measures, it is important to understand the horizontal and vertical convergence in Indian context. Vertical Convergence can be understood as the convergence taking place within a sector, for instance; convergence of services in telecom sector, broadcasting of television including cable and satellite television and Conditional Access System (CAS), which is, digital mode of transmitting TV.

Hereby, the paper will be discussing the vertical convergence of services in telecom sector. As mentioned above, the history of telecommunication in India can be traced well back to the establishment of Indian telecommunication services in 1965 and the establishment of MTNL in 1986 was proved to be a watershed. In 1991, liberalization of Indian economy led to the privatization of the erstwhile governments owned sectors. Subsequently, the National Telecom Policy of India was introduced in 1994 to regulate the telecom sector. The policy was aimed at 'improving India's competitiveness in the global market and rapid growth of exports. Another element of this policy was to attract foreign direct investment and stimulate domestic investment' (TRAI, NTP 1994. The corollary of the policy was the entry of foreign private players in Indian market to provide services (Airtel) and to set up manufacturing units (Nokia, Alacatel). This period was marked with rapid growth in the development of technology and the services. The examples to illustrate the vertical convergence in telecom sectors are Short Message Services (SMS), Multimedia Message Services (MMS), call transfer and roaming facility. The other example is mobile number portability, wherein, the user without changing the number can switch to another network. The important aspect which is to be discussed in later sections is the regulation of the sector wherein multiple players are involved and are sharing services or infrastructure (for e.g. signal towers) to provide the services to consumers.

The other parameter to be discussed here is horizontal convergence, means that the convergence of technology and services is taking place across sectors. This can be illustrated through various examples, like on one hand there is integration of camera with mobile and on the other hand, the files or photographs and videos can be seen at Television or Computer by attaching the mobile with it. This is one aspect of technological convergence across sectors, but the other aspect is important from this paper's point of view; which is that today we can avail the broadcasting services on our mobile network. This type of convergence involves a trade – off between the mobile network service provider and the television service providers (individual channels).[3]

Network Service Provider   Service Provider of another
of one sector   sector
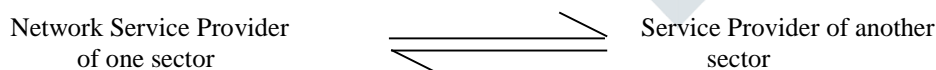
TRAI has defined convergence as 'convergence of mediums or technologies facilitating provision of all services by using a given facility or network and vice versa. It also means convergence of services at the provider's end as well as the consumer's end implying that a service provider should be able to provide a whole range of technologically feasible services and a consumer should be able to receive all services through a given terminal at any time and place of his choice' (TRAI Draft Bill, 2000; p.73).

III. OBSERVATIONS

As discussed above, Indian media industry has gone through various facets in terms of regulatory measures. But as far as convergence is concerned, it is seen as the fastest growing industry. This became evident after liberalization of Indian economy when private players entered into the Indian market both in telecommunication as well as in broadcasting sector. Eventually, a trend can be seen over decade where both the sectors have converged to provide the new technological services.

---

[2] Latzer has quoted Höflich, 1999; Hasebrink, 2003; Wagner et al., 2004 to define receptive convergence.

[3] 3G JADOO of MTNL used to providing services to watch Television channels on Mobile by paying either individually for a single channel on per day basis, or by paying a monthly amount. There are only 12 channels in the list that user can watch, so a trade – off between MTNL and the satellite channel is quite evident. Similarly, JIO, Airtel and all other private service providers have been converging to provide satellite services in the digitized form.
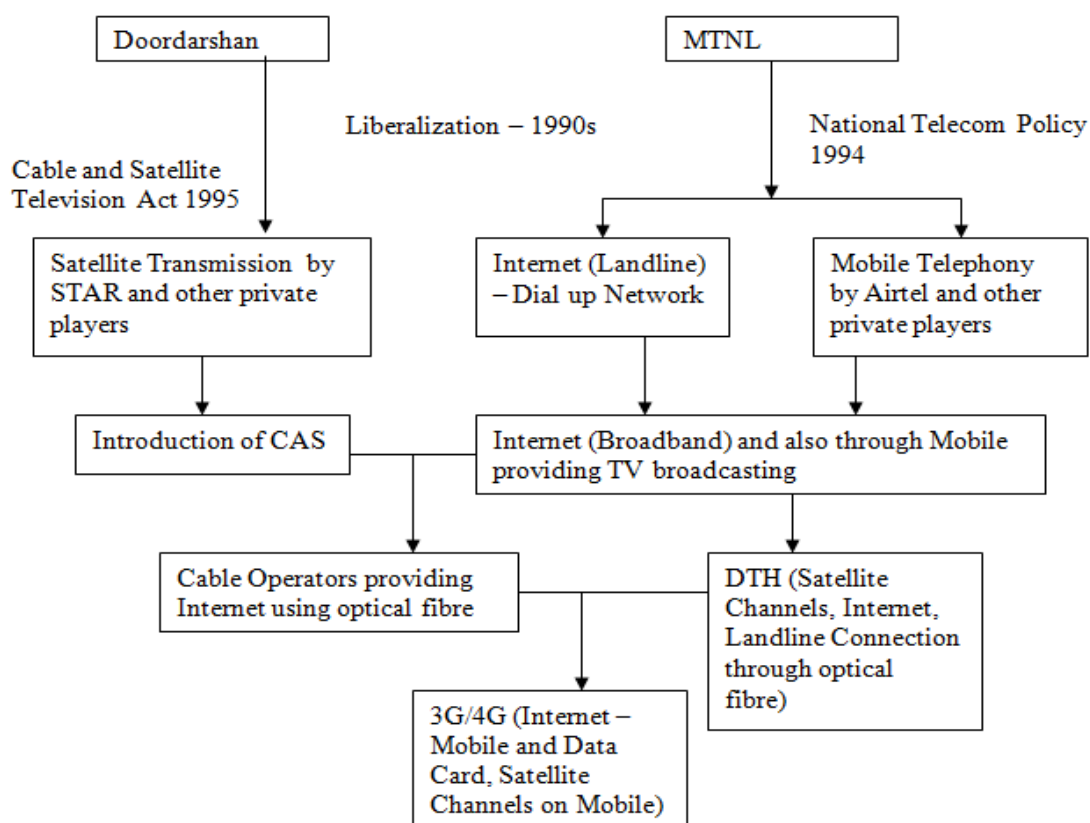
Figure 1: Trajectory of Telecommunication and Broadcasting Sector in India

(Source: Self Compiled)

The linkages between two sectors show the trade – off between these sectors and the convergence between the two. This convergence is not only limited to the above given example, but rather can be seen even in Radio broadcasting. So, under such circumstance, there are questions like who will regulate, what will be regulated, how will be regulated? These questions are still required to be answered. This study is an attempt that can help us in understanding the situation more clearly, also might help us in choosing a viable regulatory option.

3.1 Regulatory Measures

Before looking at the regulatory measures, we should first understand the term regulation and its need in the converging scenario. The term regulation and intervention is most of the time used interchangeably and is misconceived as an instrument in government's hand to control the economy and society. But the need of time, considering the convergence of technology, we require a viable mechanism to regulate the industry and society in an efficient manner. Thus, the converging technologies require new meanings of the term regulation. Jordana and Levi-Faur has used the Baldwin et al.'s view to define regulation as the following:
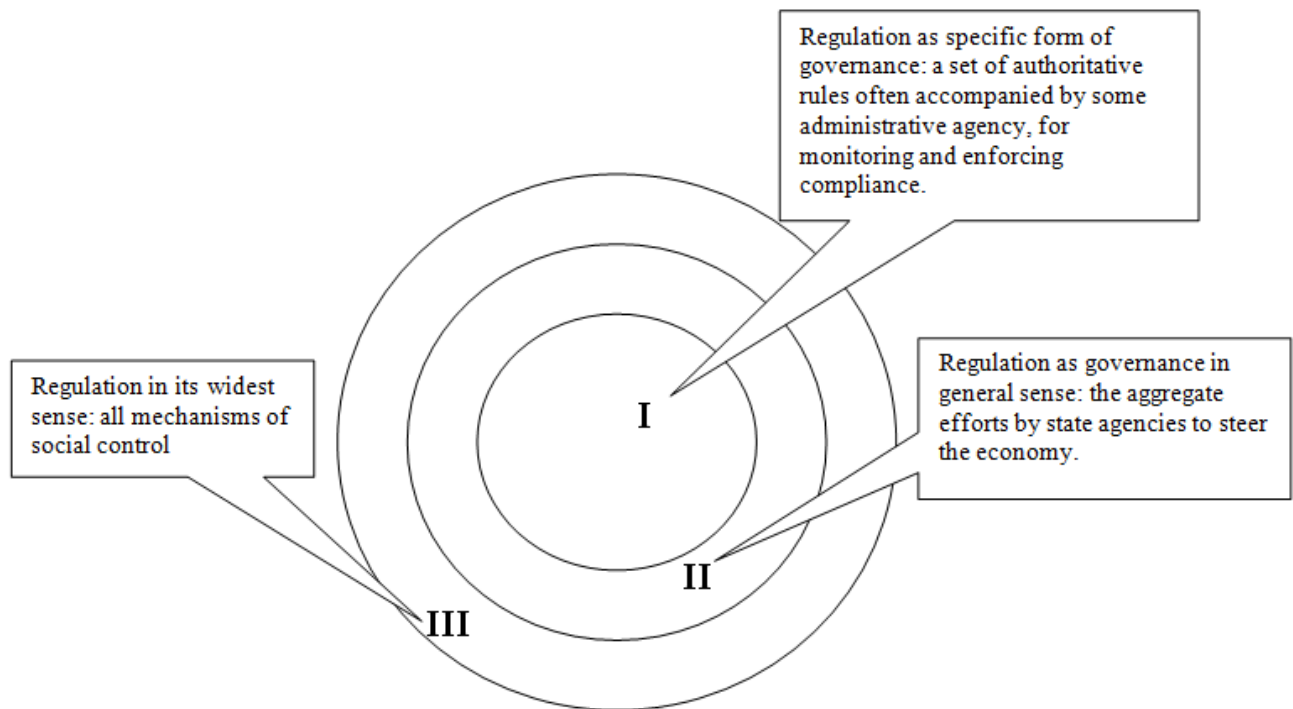
Figure 2: Definitions of Regulation

(Source: Baldwin et al. In Jordana and Levi-Faur, 2004)

Regulation has been defined with the help of three concentric circles; the first circle can be defined as a promulgation of an authoritative set of rule, accompanied by some mechanism, typically a public agency for monitoring and compliance with these rules. The second circle refers to all efforts of the state agencies to steer economy and the third circle encompasses all mechanism of social control, including unintentional and non – state processes.

According to Jordana and Levi-Faur (2004), until the end of the 1980s, scholars outside the United States tended to employ the word 'regulation' to denote the general instruments of government for the control of the economy and society (meaning II). But in United States the meaning of regulation was different, from broadest to narrowest due to the large number of independent regulatory agencies and the consequent crystallization of regulatory practices into theory of governance (meaning I); it has been proclaimed that after 1990s, there is a global spread of the wave of regulatory reforms, and especially the establishment of independent regulatory institutions, in various sectors has led to the convergence in the meaning of regulation from broader which is more general to the narrow one. This movement was strengthened by a shift in the way some economists used the notion of regulation. Therefore, each type of definition suited to people from different discipline. And so, it is important to understand the need of the particular situation in which it can be regulated as there are exhaustive and consensual definitions of regulation across different disciplines and research agendas, but we should focus for a specific context and goal that shape the particular meaning of the notion of regulation.

3.2 Role of TRAI

The Telecom Regulatory Authority of India (TRAI) is the apex regulatory body in India formed to meet the objectives of New Telecom Policy, 1999 and is aimed to encourage the competition in telecom sector together with better quality and affordable prices. Broadcasting and Cable Services were also brought within the definition of 'telecommunication sector' in terms of section 2(k) of Telecom Regulatory Authority of India Act, 1997 as amended by TRAI (Amendment) Act, 2000.

The aims and objectives of TRAI include increasing tele-density, making services of best quality available at affordable prices, social equity that includes Universal Service Obligation fund (USO) under which both the public and private companies are required to give 5% of their revenue as Access Deficit Charge (ADC), preparing grounds for smooth transition to an era of convergence of services and technologies, protecting interest of consumers, promoting the growth of coverage of broadcasting, increasing consumer choice in reception of TV channels and choosing the operator who would provide television and other related services.

So, depending upon the needs of new emerging technologies and to have more viable regulatory mode, we have been analyzing the trends of convergence in regulation as well. But the difficulties are evident for TRAI, there is need for integrated approach to regulate ICT and Broadcasting in India, but, this move has been opposed by broadcasters in India so far. And, in recommendations of National Telecommunication Policy (2018), TRAI states that that there should be an integrated regulation of ICT and broadcasting sector and that TRAI itself should be restructured as a converged regulator for both.

3.3 Need for Convergence Act in India

TRAI drafted a Convergence bill in 2000 and the bill says that convergence is a means or the provision of different kinds of services over the existing infrastructure and the enhancement of existing technologies so as to provide a wide variety of services which is resulting into the blurring of borders between telecommunications, computing and media.

The continuous development of new technologies results in an inability to predict the future evolution of convergence viz. the development of new services like web-casting, Internet Telephony etc. Resulting in the need for regulations which does not aim to

predict the future, but aspires to be flexible enough to accommodate and propagate any permutation and combination of technologies and services.

It predicts regulation of convergence as an effort to facilitate the convergence. It says that its aim is to provide for a regulatory mechanism, which facilitates convergence and therefore, remains valid over a period of time and says that in order to regulate the convergence, there are four categories of licenses. These are as follow:[4]

(a) Network infrastructure facilities;
(b) Network services;
(c) Application services; and
(d) Content application services.

The licensing structure has hence been broken into its different elements which lead to a four layered hierarchical structure where each layer is dependent on one or more of the earlier layers for the provision of services. The structure results in four different service providers namely;

a) **Network infrastructure facility provider**; where network infrastructure facilities refer to the provision of physical infrastructure which would be utilised by other licenses for providing various services.
b) **Network service provider**; who will utilise the infrastructure set up by one or more network facility providers to carry various kinds of services.
c) **Application service provider (ASP)** is the one who provide services to the end consumer using the services of one or more network service providers.
d) **Content ASPs** are the ones that provide content to the end users using the services of one or more network service providers.

3.4 Regulatory Choices

The Indian media industry has already experienced the statutory regulation, under which all the sectors were completely controlled by government, and in fact, the broadcasting of television and radio became the instruments in the hands of government to the extent that they were being misused during the period of emergency to formulate the public opinion. The misuse of television or radio has been evident since its use by Indira Gandhi, the Prime Minister of India during the period emergency in 1987. The radio was misused to such an extent that people started calling All India Radio as 'All Indira Radio'.

Therefore, various committees for formed in order to have optimum policy outcomes, the committees actively involved were Chanda Committee, Verghese Committee, a committee headed by P.C. Joshi. Almost, all of these committees mentioned in their report that Indian media required decentralization and autonomy. And so, many attempts have been made in this direction but the political parties always subjugated the matter and at last the bill which came in 1997, the Prasar Bharti Act, it had all the clauses different and modified from the ones which these committees asked for. So, it may be said that there were so many developments going on in terms of technology or convergence, but media did not have the autonomy till that period. Therefore, there is need of shifting from vertical to horizontal regulation, i.e. sector specific to integrated media market which involves institutional change of existing separate regulatory structures for telecommunication and media, the integration of political responsibilities at the ministerial and parliamentary level, as well as the harmonization of respective laws and regulation for telecommunications and media (Just and Latzer, 2004).

The change in technology definitely require a viable mode of regulation as statutory regulation was failed to fulfill the requirements, in such a situation there is need to rethink upon the valid regulatory options or if there is possibility of having universally applicable regulation rather that the sector-specific; as from economics point of view there is 'normalization' of sectors taking place, i.e. an adaptation to other sectors, is taking place with the liberalization of telecommunication and broadcasting markets and the 'economization' of media markets.

Hence, due to globalization and rapid technological change there is need for a shift from statutory regulation to co – regulation or self regulation, the distinction between the two forms of regulation refers to the varied intensity of intervention by the state government.

3.4.1 Co - Regulation

Just and Latzer (2004), Palzer and Scheuer (2003) argue that regulation takes place on a continuum between the pure state regulation at one end and pure self regulation at the other end. The co-operative arrangement of private and public is often conceived as co – regulation. Prefix "Co-" refers to the degree of state involvement in the regulatory process, Just and Latzer (2004) calls co-regulation as self-regulation with public oversight or ratified by the state, it is self-regulation with a legal basis. Palzer and Scheuer (2003) puts forward that the term co-regulation denotes co-operative forms of regulation that are designed to achieve public authority objectives – the co-operation being performed by public authority and civil society. This scheme combines elements of self-regulation (and self-monitoring) as well as of traditional public authority regulation to form a new and self-contained regulatory system.

Palzer and Scheuer (2003) say that there can be different models of co-regulation, they have talked about the two models that broadly covers many other forms of regulatory models, one, is the possibility that state would integrate an extant self regulatory system into a public authority framework. And another possibility would be initiation of a co-regulatory system by the state. In this

---

[4] The classification is technology-neutral and service sector neutral. Setting up an infrastructural facility and its use is not linked to the provision of a particular service by using a particular technology. Similarly, services can be provided by using any facility and any technology.

case, the public authority would lay down a legal basis for the co-regulation system, so that it could begin to function. The choice of deciding the elements as the foundations of a co-regulation framework depend in particular on the task to be performed. One common feature will exist in each case: the pursued aim will be a public one.

A key element of a co-regulatory regime is the self-contained development of binding rules by the co-regulatory organization and its liability for these rules, the latter being one of the main differences between co-regulatory systems and self-monitoring systems (Ibid.). In respect of the distinction between self- and co-regulatory schemes, an important criterion is the voluntariness of participation. In a co-regulatory-system, non-compliance with the given rules is directly or at least indirectly (e.g., in the form of possible revocation of a license) sanctioned by the state (public authority). Thus, the market players concerned are not actually free in their decision to participate in the system. In fact, in a functioning self-regulatory system, there is also some pressure to participate; although this pressure is not exercised by the state but by the public, the customer – in short, by societal institutions.

The greater the public authority's involvement in a co-regulatory model, the less participation in the inclusive co-regulatory organization can be considered to be voluntary. This leads to the distinction between co-regulation and state regulation: in this context, the main criteria can be seen in the degree of autonomy of the co-regulatory organization from state influence, e.g., the extent to which it can make its own decisions, or whether representatives of the public authority can exert influence over the rule- or the decision-making of the co-regulatory body.

3.4.1.1 Advantages and disadvantages of co-regulation[5]

Due to new technological developments, especially their speed and the growing Convergence, public regulation or the statutory regulation is deemed to be no longer able to solve some of the problems. At present, the manner of bringing an audiovisual product to the consumer is decisive for implementation of standards for the protection of minors. This "artificial" distinction already causes some problems. If convergence becomes reality in the mid-term perspective, providing for different regulatory regimes according to merely technical considerations (what represents the status quo) will become increasingly difficult – if not deficient and therefore unacceptable.

Against this background, the prospects and advantages of co-regulation become visible. The more stakeholders take the initiative for responsible handling of relevant concerns, such as the protection of minors, the more efficient and prompt the regulatory framework can react to new technologies. On the other hand, in the context of co-regulation, the state or competent authority will play a significant role, setting the legal framework and monitoring the functioning of the system by assuming responsibility for initially checking self-regulatory bodies, having a say on the monitoring of results and, if necessary, requesting that adaptations be implemented. Thus, the achievement of public policy goals is not relinquished to societal control entirely; the responsibility remains with the state, which is often even under an obligation to guarantee this achievement. With regard to youth protection in the media, where such fundamental public goods are at stake, the state cannot exercise complete restraint in view of its responsibility for safeguarding the public interests involved. Furthermore, democratically-founded legislation will have to establish, in most cases, criteria according to which co- regulatory systems should work, addressing such issues as complaint procedures, sanctioning powers in view of members, organization and representativeness, conditions for accreditation, etc.

Bearing in mind that state authorities may intervene in the case of an alleged malfunctioning of a co-regulatory institution, this will necessitate, at least to some extent, the doubling of institutional structures, on the side of the organization in charge and, in addition, on behalf of a competent state authority. Therefore, one may doubt that in the short-term co-regulation will also show prospects for more efficiency, in particular in terms of costs.

3.4.2 Self Regulation

The prefix 'self' can be categorized into two — in an individual sense and in collective sense, by individual sense it means that one company sets its own rules and by collective sense it means that an industry groups regulates the conduct of its members (Just and Latzer, 2004). Self regulation can be defined as the system which is situated at the other end of the "regulatory scale". Under this system, social groups (producers, providers, etc.) draw up their own regulations in order to achieve their objectives and take full responsibility for monitoring compliance with them (Palzer and Scheuer, 2003). Ang and Pramanik (2008) articulate that the term "self-regulation" is often juxtaposed against government regulation so that where is there self-regulation; there is no need for government regulation. Traditionally, it has been taken as industry regulating industry, where government delegated most of its regulatory powers to industry, typically an industry association, while reserving powers of the ultimate, i.e. the strongest, sanctions. Self-regulation is therefore a form of regulation, a delegated regulation at that.

This form of regulation may take the form of technical or qualitative standards, potentially combined with codes of conduct defining good and bad practice. Codes of conduct may also contain rules on out-of-court mediation and on the structures of the relevant complaints bodies. These rules may be laid down by a self-regulatory organization created by the parties concerned (ideally involving other interested parties, such as consumers). This body may also monitor compliance with the rules and impose any sanctions, if provided. Such a model might even be considered the preferable one, because the rule-making – "legislative" – power is separated from the rule-applying – "executive" – power (Palzer and Scheuer, 2003).

3.4.2.1 Advantages and disadvantages of self-regulation

Ang and Pramanik (2008) have said that the success of self regulation can be determined by certain conditions; the success is seen in the following cases:

a) **Motivated industry**, i.e. 'industry is not dragging its feet and feeling compelled to have to self-regulated' so, there should be incentives that might motivate people to participate. So, on the one hand this voluntariness can be seen as a positive point whereas on the other hand it might be seen as a negative point (Palzer and Scheuer, 2003). Positive aspect can looked

---

[5] Just and Latzer (2004) have talked about the potential incentives, risks, and success factors of self and co – regulation in a generalized form without making any distinction between the pros and cons of both the regulatory models separately. The mentioned description is based upon Palzer and Scheuer (2003).

as that the members, voluntarily submitting to the system, are willing to comply with the rules. Normally (when the system is running well) they obey without coercion; they are convinced that conformity of their behaviour with these rules is the best way to solve the given problems. This attribute can work well as in the case of internet where voluntary self- regulation, implemented ideally by all, if not by most of the stakeholders, may provide for a solution. On the other hand, it might become a weak point, i.e. stakeholders cannot really be forced to comply with the rules. If, for example, the management of an undertaking changes and the new leaders do not subscribe to the idea of self-regulation, the heaviest sanction that can be imposed is exclusion from the system – the effect being that the failing undertaking is no longer subject to the rules.

b) **A small number of large players** — this attribute definitely helps in increasing competition and thus making the services at affordable rates available to all consumers.

c) **A government regulatory backstop,** i.e. government must be willing and able to put in place a regulatory mechanism to address the recalcitrant offender.

d) **Maturity in market** — the rules for market behaviour have yet to stabilize. A matured market will therefore make it easier to self-regulate.

The above mentioned conditions determine the success rate of this regulatory model. Apart from that another positive aspect is that self-regulation as a self-contained regulatory system often receives support. This is due to the fact that the industry is perceived as most experienced and best placed to evaluate the risks of their products and services. Professionals may react to new challenges in an easier and faster manner than do the makers of public regulation (Palzer and Scheuer, 2003).

The disadvantage of self-regulation is that it suffers from a lack of democratic legitimation. It originates from economic players or groups with their own specific interests – interests that may contribute to or even is partly congruent with the general interest; nevertheless these special interests do not necessarily coincide totally with the general interest. Thus, there will always be a tendency to allege that self-regulatory organizations pursue their own policies rather than general policy goals.

Since the state is not involved in this form of regulation, public authority sanctions cannot be imposed, but only those provided by civil law, particularly the articles of association(s). According to this, the most severe sanctions are financial penalties or exclusion from the relevant association that has adopted the self-regulatory system. Thus, the self-regulatory approach is not primarily based on enforcement by punitive or exemplary sanctions. Based on agreement, the conviction that the parties concerned have common objectives should ensure the effectiveness of this system. The key element of the above-defined self-regulatory system is the voluntary nature of the participation of those who are subject to regulation.

IV. CONCLUSION

It may be concluded that regulatory approaches and instruments change with the subject of regulation, in our case with the communications industry and the societal communications system respectively. As argued above, the convergence of telecommunications and mass media at the corporate level--which challenges the regulatory telecommunications-(mass) media-dichotomy plus the liberalization and growing globalization of mediamatics markets-which challenge the former dominance of national regulations--are causing a political control crisis and pose new challenges to the regulatory system. Self-and co-regulation are considered as tools of great promise in this situation, especially by the industry.

Therefore, a lot of engagement is required to understand which definition of regulation follows in India Context and which regulatory mechanism will be appropriate for India. But the policy challenge for the regulatory bodies is to consider the socio – economic framework as well for optimum policy outcomes; they should keep in mind that their objective is to facilitate the convergence in such a manner so that policy should attend the vast majority of poor and uneducated people of the county.

REFERENCES

[1] Ang, Peng H. and Pramanik, A. 2008. By the Industry, of the Industry and for the Industry: The Possibilities and Limits of Self - Regulation of Indian Broadcasting. paper presented at International Seminar, Contours of Media Governance, Jamia Millia Islamia

[2] Hukill, M., Ono, R. and Vallath, C. 2000. Electronic communication Convergence: Policy Challenges in Asia. Sage Publication: New Delhi.

[3] Jordana, J. And Levi-Faur, D. 2004. The politics of regulation: Institutions and Regulatory reforms for the age. United States: Masschusetts.

[4] Just, N. And Latzer, M. 2004. Self- and Co-Regulation in the Mediamatics Sector: European Community (EC) Strategies and Contributions towards a Transformed Statehood. Knowledge, Technology, & Policy. 17(2): 38-62.

[5] Mosco, V. And Mckercher, C. 2006. Convergence Bites Back: Labour Struggles in the Canadian Communication Industry. Canadian Journal of Communication. 31: 733-751

[6] Palzer, P. and Scheuer, A. 2003. Self, Regulation, Co-Regulation and Public Regulation' Promote or Protect? Perspectives on Media Literacy and Media Regulations. Yearbook: Gotherburg.

[7] Telecom Regulatory Authority of India. 2000. The communication Convergence Bill 2000

[8] Department of Telecommunication. 2018. National Telecommunication Policy 2018