# Study of Evolutionary Algorithm Application:Information Retrieval

[1]Sanjay Singh, [2]Sandeep Gupta
[1]Research Scholar, [2]Associate Professor
Department of Computer Science,
SHRI VENKATESHWARA UNIVERSITY,
Gajraula, UP, INDIA,

***Abstract:*** Evolutionary Algorithm has turned into a very rich and different field of concentrate throughout the years, and the sheer number of distributions in this field can make difficulties for new individuals. To address this, the motivation behind this investigation of Evolutionary Algorithm and their application, concentrating on traditional instead of archaic usage. In this contribution, various proposals for applying evolutionary computation to the field of image optimization i.e. information retrieval will be reviewed found in the specialized literature. The EA's belongs to family of algorithms motivated by nature widespread use to solve Comprehensive issues of optimization.

***IndexTerms - Evolutionary Algorithm, information retrieval, Genetic Algorithm, Content based information retrieval***

## I. INTRODUCTION

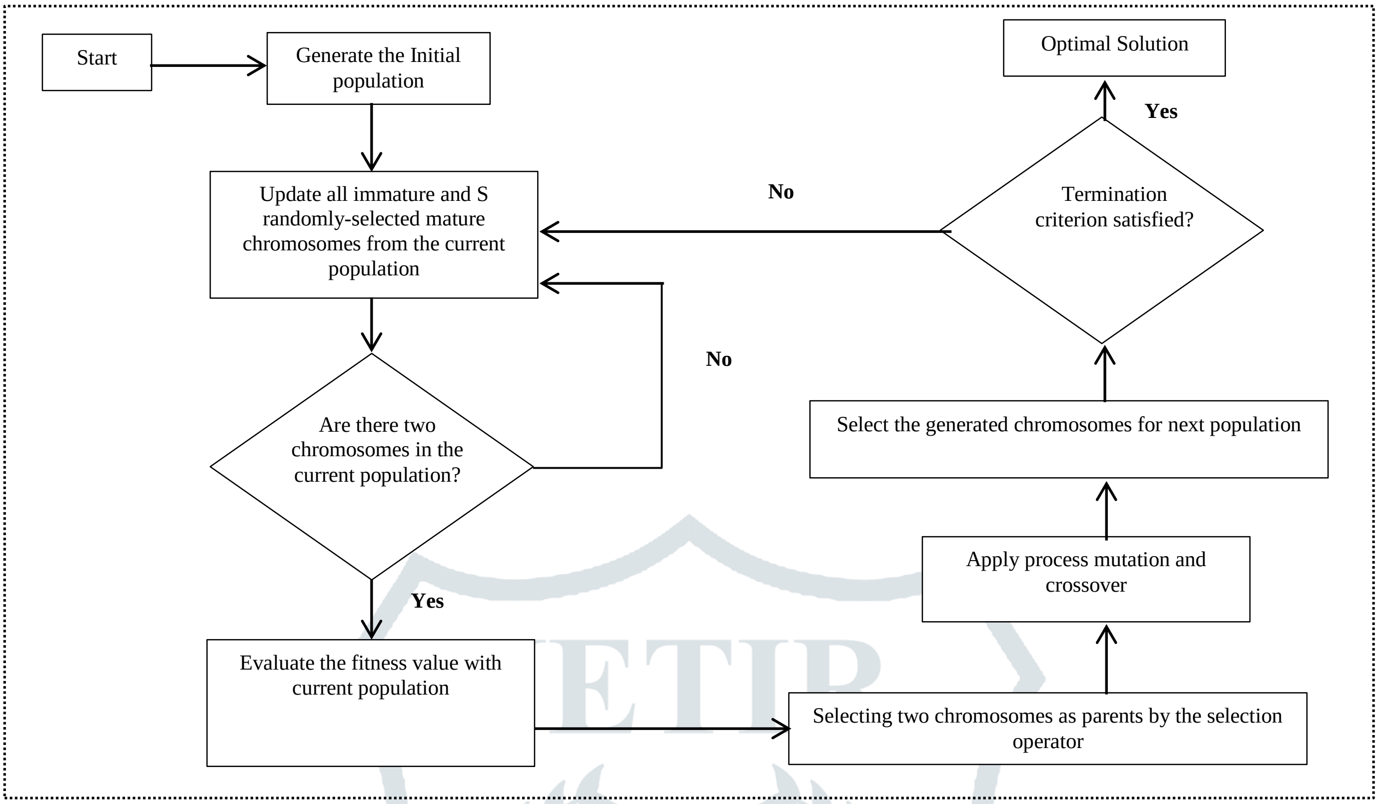
Present world is computerized with the presence of numerous gadgets that are utilized in multimedia storage. Nowadays, large quantities of images can be easily stored using processing techniques. Since the gathering of pictures and databases is quick is expanding day by day. New image recovery techniques are needed that should be efficient and quick. Unfortunately, the massive length of those databases has made good sized efforts during the last few years to retrieve beneficial records. Recovery of information (IR) attempts to make appropriate use databases, providing users with timely access to information that is really relevant [1]. A growing interest in applying artificial intelligence, is based techniques to information retrieval, has been shown over the past few years with the goal of resolving some of these shortcomings. One of the Artificial Intelligence areas with a significant development in the most recent decades is Evolutionary Computation. The various models proposed in this philosophy are normally referred to as Evolutionary Algorithms.

The evolutionary algorithms usually show robust performance on the complex problems of optimization. In fact, the evolutionary algorithm is a family of algorithms inspired by nature's evolutionary principles. The EA's family such as Genetic Algorithm(GA), Genetic Programming(GP),Evolutionary strategies(ES), evolutionary Programming(EP), differential Evolution(DE), Swarm Intelligence(SI),Particle swarm optimization(PSO), ant colony optimization(ACO), and Binary Bat algorithm also regard as new member of the EA family. Although in a number of respects these algorithms differ from one another, every one of them depends on a similar center procedure. Every one of them has a search point population (known as candidate solutions, individuals, chromosomes or agents in different ways) These are typically randomly generated and then iteratively evolved over a series of generations by applying and selecting variation operators. Variation operators generate changes to the population members, i.e. they perform movements through the search space. The objective value (or fitness) of each search point will be calculated after each generation. Then selection removes the search points with the lowest target a value, which means that only the best search points are maintained and new search points, are always derived from them. This combination of maintaining a search point population and selecting search points distinguishes EAs from most other metaheuristics. In spite of the extraordinary technical information adopted in exceptional EAs. Most of them percentage a common place framework as given in Fig 1. Each generation within the fundamental loop of an ordinary evolutionary algorithm includes the subsequent components: reproduction, fitness evaluation and selection. A range of recent works have been devoted offering learning - enabled EAs.

In this article we investigate the utility of Evolutionary computation to information retrieval, investigating the family of evolutionary algorithm, study the various types of information retrieval problems solved by the evolutionary algorithm and to evaluate the results obtained.

For this purpose, the article is structured as follows: Some preliminaries are introduced in section 2 by examining the information retrieval basis and the EAs family. In section 3 different applications of EA's is analyzed. At last, we have a couple of finishing up comments.

Fig. 1. *Evolutionary Algorithm Flow chart*



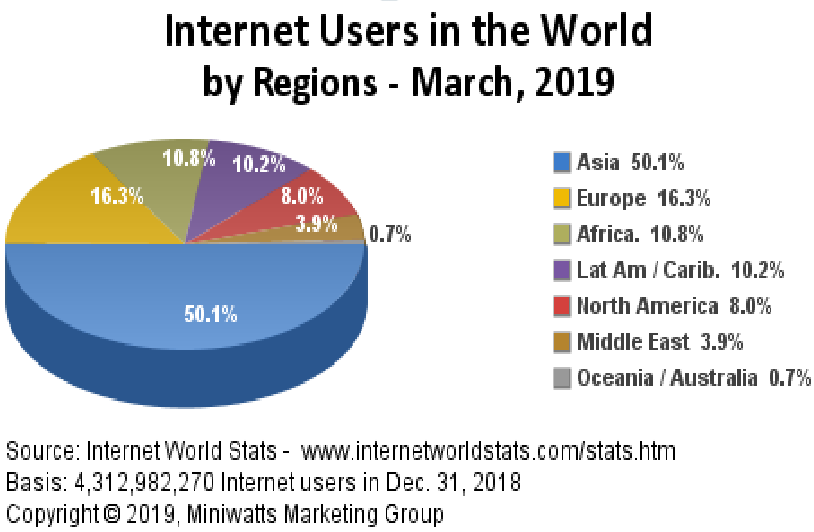

## II. Related Work

### *Information Retrieval*

In general, in response to user search questions, IR can be defined as the problem of selecting documentary storage information [2]. Information Retrieval frameworks (IRSs) manage narrative bases containing literary, pictorial or vocal data and procedural user inquiries aimed at enabling the user to access relevant data in a timely manner. The diverse worldwide web indexes, for example, Google are these days the fundamental instances of IRSs. First, the user must translate this need for information into a query that the IR system will process. This translation provides a set of keywords in its most general form to summarize the description of the user information needed. Given the user question, the key objective of IR framework is to recovered data which might be applicable to the user.

Nowadays, using the Internet to gain access to WWW information has become an important part of human life. The current population of the world is about 7,753,483,209 billion out of which 4,312,982,270 billion people (55.6%) use Internet [3] (see Fig. 2). As in Asia internet user increased 50% in March 2019. The data is taken from the source internetworldstats.com, which shows that all the information of the world is available on the internet. It is act as the repository of information. As on other words we can say that today's world is known as

Fig. 2. *Internet Users in the World Regions*



Digital world and all the digital data are available on internet. Now the question arises that how we can access the right information at the right time. For this we need a strong information retrieval system form which we can retrieve the information (it can be image, data, text, audio, video any type information which can be stored in digital format) form internet repository.

*Table 1. World Internet Usage and Population Statistics –March, 2019 [3]*

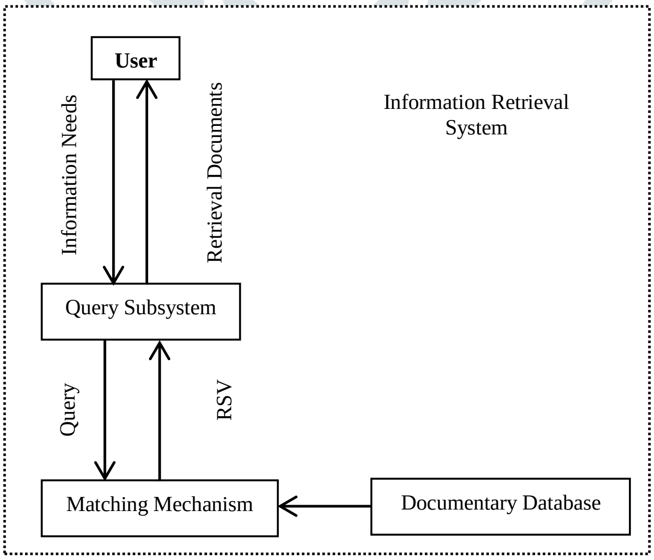
| WORLD INTERNET USAGE AND POPULATION STATISTICS | | | | | | |
|---|---|---|---|---|---|---|
| MARCH, 2019 - Update | | | | | | |
| World Regions | Population ( 2019 Est.) | Population % of World | Internet Users 31-Dec-18 | Penetration Rate (% Pop.) | Growth 2000-2019 | Internet Users % |
| Africa | 1,320,038,716 | 17.00% | 4649,23,169 | 35.20% | 10199% | 10.80% |
| Asia | 4,241,972,790 | 54.70% | 2,160,607,318 | 50.90% | 1790% | 50.10% |
| Europe | 8664,33,007 | 11.20% | 7050,64,923 | 81.40% | 571% | 16.30% |
| Latin America / Caribbean | 6583,45,826 | 8.50% | 4382,48,446 | 66.60% | 2325% | 10.20% |
| Middle East | 2583,56,867 | 3.30% | 1700,39,990 | 65.80% | 5076% | 3.90% |
| North America | 3664,96,802 | 4.70% | 3456,60,847 | 94.30% | 219% | 8.00% |
| Oceania / Australia | 418,39,201 | 0.50% | 284,37,577 | 68.00% | 273% | 0.70% |
| WORLD TOTAL | 7,753,483,209 | 100.00% | 4,312,982,270 | 55.60% | 1095% | 100.00% |

### 2.1 Component of an Information Retrieval system

The information retrieval system consists of three main components:

1. Documentary Database: This component stores the representation of the documents and their information content. It is linked to the indexer module, which by extracting the content of the document, generates a representation automatically for each document.
2. The subsystem of queries: It enables the clients to plan their inquiries and presents the applicable archives recovered by the framework to them.
3. The mechanism for matching: It assesses the extent to which the document representations meet the query requirements

Fig. 3. *Basis Structure of Information Retrieval System*




### 2.2 Information Retrieval System Evaluation

There are so many ways to measure the exceptional of an IRS, consisting of the system efficiency and effectiveness, and several subjective factors associated to the user pride. Usually, the retrieval effectiveness most likely founded on the record relevance with recognizes to the consumer's desires is essentially the most considered. There are different criteria for measuring this aspect, the most used being the precision and the recall. Precision is the number of documents recovered and the total number of documents recovered by the IRS in response to a query. Recall is the rate within the information base between the quantity of significant documents retrieved and the full number of relevant records to the current question. Each of them shows the mathematical expression as follows:

$$P = \frac{\sum_d r_d . f_d}{\sum_d f_d} \quad (1)$$

$$P = \frac{\sum_d r_d . f_d}{\sum_d f_d} \quad (2)$$

Since $r_d \in \{0,1\}$ is relevance of document $d$ to the user and $f_d \in \{0,1\}$ is the retrieval of document $d$ in the current query processing. Note that both measurements are defined as the optimal value in { 0,1 }. The evaluation function here is the average non - interpolated accuracy [4,5]. Which is similar to average accuracy but equivalent to the training documents with the cutoff points? The documents are simply ranked in this measurement function. Let a1.a2,...,$a_{|D|}$.Denotes sorted documents as the values of the similarity measurement function decrease. Where the number of training documents is represented by |D|. The r(d) function gives a document d's relevance. If d is relevant, it returns 1 and 0 otherwise.

## III. Evolutionary Algorithm

Evolution computing uses computer models of evolutionary techniques as key computer - based problem elements to solve and implement systems design. Also used to refer to EAs is the term evolutionary computation, but generally includes optimization algorithms motivated by other natural processes, such as particle swarm optimization and artificial immune systems. Over the years, EAs have become an extremely rich and diverse field of study, and the sheer number of publications in this area can create challenges in the field for new people. There are a variety of developed and studied evolutionary computational models, known as EAs. [6]. Genetic algorithms (GAs), evolution strategies (ESs), differential evolution (DE), genetic programming, evolutionary programming, multi-objective evolutionary algorithm, particle swarm optimization, Ant colony Optimization, Binary Bat Algorithm and distribution algorithm estimation (EDAs) are the family members of Evolutionary Algorithm in contemporary use. Although in a number of respects these algorithms differ from one another, All of them are based on the same core process.

### 3.1 Genetic Algorithm

Genetic set of rules is a probabilistic set of policies that simulate the natural selection mechanism of residing organisms and is often used to solve problems with highly-priced responses. The search space consists of candidate answers to the query in the genetic algorithm; each depicted by a string is called a chromosome. Each chromosome, called fitness, has an objective function value. A set of chromosomes is called the population together with their associated fitness. This population is called a generation in a given iteration of the genetic algorithm. In the context of continuous non - linear optimization, Holland, De Jong and Goldberg pioneered the genetic algorithm [7-9Genetic algorithms aren't new for facts retrieval [10-11], Gordon proposed, representing a chromosome posting and selecting good indexes with genetic algorithms [12]. user feedback from GAs to select weights in a query for search terms suggested by Yang et al [13]. Morgan and Kilgour proposed a mediator between the client and informational retrieval framework utilizing GAs to pick seek terms from a thesaurus and dictionary [14]. Boughanem et al. [15], Horng and Yeh [16], and Vrajitoru [17] examine GAs for retrieval of information and suggest new operators for crossover and mutation. Vrajitoru examined the impact of population size on mastering capacity, concluding that a large population size is crucial. Given their long history, the implementation of genetic algorithms varies significantly. However, the use of a mutation operator is fairly common with a certain probability to change each decision variable.

### 3.2 Evolution Strategies

Evolution techniques additionally have long records, and this parallels the improvement of GA. even as early ESs had been limited to a single point search factor and used no recombination operator, present day formulations have converged in the direction of the GA norms, and have a tendency to apply both a population of search factors and recombination. Adapting strategy parameters during an ES run is standard practice, with the essential thought that distinctive kinds of developments will be advantageous at various phases of the search. ES utilize distinctive GAs recombination operators and frequently utilize multiple parents to make a solution for each child. For example, Intermediate recombination provides a child solution with average values for each decision variable in each of the parent solutions. Multi - recombination weighted is comparable, however utilizes a weighted normal dependent on each parent's wellness. ESs additionally will in general utilize deterministic instead of probabilistic determination systems not at all like GAs, whereby the best arrangements are constantly utilized as next - generation parents in the population.

### 3.3 Differential Evolution

Differential evolution is a highly latest EA system, using an adaptive search mechanism that does not use distributions of probabilities [18-19].

The basic distinction in structure better arrangement is that genetic algorithm depends on crossover in the meantime as DE depends on mutation process. This foremost operation is based totally on differences inside the population between pairs of solutions randomly sampled. The algorithm uses mutation operation to direct the search to the prospective regions in the search space as a search technique and selection process. An advantage of the usage of simplex-like mutations in DE is that the arrangement of standards is basically self-adjusting; with developments precisely decreasing in each estimation due to the population converges. More generally, the authors of the method claimed that this kind of self - adaptation meant that the size and direction of movements were automatically matched to the search landscape, a phenomenon that they called contour matching.

The DE algorithm is a new heuristic technique specifically having 3 blessings; locating the proper worldwide minimum irrespective of the initial parameter values, fast convergence, and the usage of some manipulate parameters. The DE algorithm appears to be an appalling approach to problems with engineering optimization.

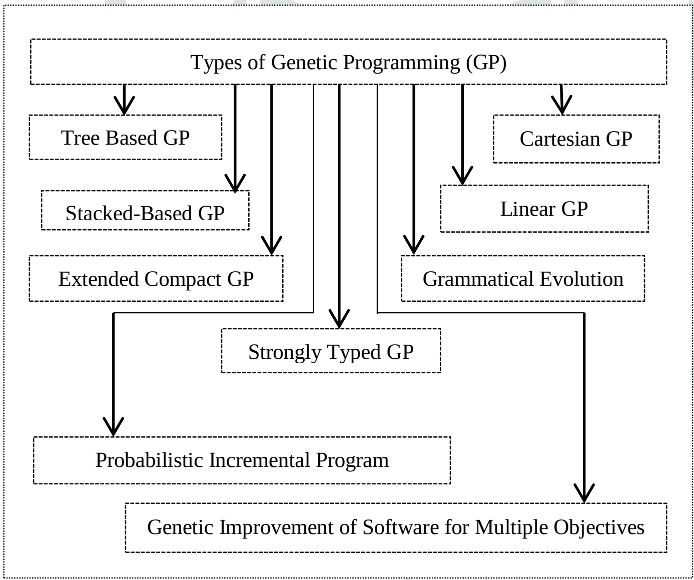
Fig. 4. *The DE algorithm's steps:*

*Initialization*
*Evaluation*
***Repeat***
*Mutation*
*Recombination*
*Evaluation*
*Selection*
***Until*** *(termination criteria are met)*

**3.4 Genetic Programming**

One of the vital significant challenges of Computer science is to get a computer to do what desires to be accomplished, without telling it the best way to do it. Genetic programming addresses this undertaking by means of offering a way for developing a work computer application mechanically from a high - stage problem assertion. Genetic programming achieves this goal of automatic programming by genetically engineering a computer program population using the principles of Darwinian natural selection and biologically inspired operations (sometimes referred to as program synthesis or induction). These operations include reproduction, crossover (sexual recombination), mutation, and architecture - altering operations patterned in nature following gene duplication and gene deletion [20].

Fig 5. *Types of Genetic Programming:*




**3.5 Multi-Objective Evolutionary Algorithm**

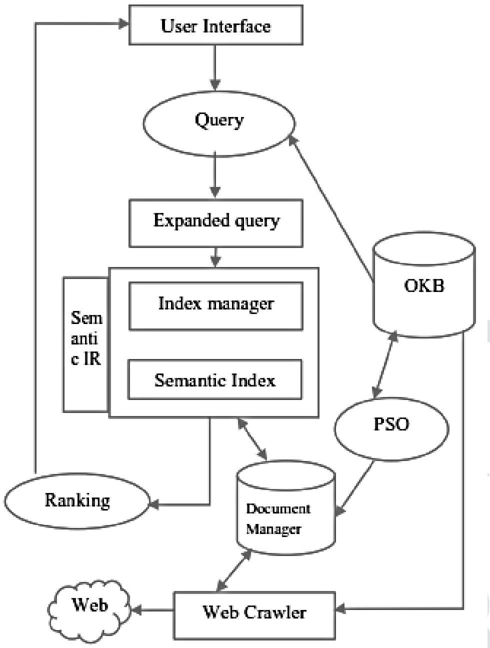
Multi-target EAs (MOEAs) are utilized to take care of issues with various objectives that are often conflicting. A focal idea for MOEAs and, all the more for the most part, multi - target enhancement is a non - ruled arrangement. This is a solution that is no worse than any of the other solutions within the population when all goals are taken into consideration and an MOEA will likely form and keep up a population of non - overwhelmed arrangements that spread all exchange - offs between objectives. The center test looked by multi-target improvement (and missing from single-target streamlining) is: the means by which to rank applicant arrangements such that prompts viable choice weight, particularly when the whole populace (or a large portion of it) might be non-overwhelmed. There are numerous different methodologies, and MOEAs are progressively being utilized as it is acknowledged that real - world issues are in nature almost invariably multi - objective. Further exchange of the last point just as a first prologue to MOEAs can be found in [21], whereas an example of a relatively recent MOEA review can be found in [22].

**3.6 Particle Swarm Optimization**

In 1995, particle swarm optimization (SO) based on the analogy of swarm of bird proposed by Eberhart et al [23]. The guideline points of interest of the PSO calculation are summarized as: simple idea, clean implementation, robustness to govern parameters, and computational efficiency whilst compared with mathematical set of rules and different heuristic optimization strategies. The valid PSO has been completed to a learning inconvenience of neural systems and characteristic optimization problems, and efficiency of the technique has been confirmed.

A. Gomathi et al [24] proposed an information retrieval using PSO. A. Gomathi et al investigated a method for PSO based on clustering of documents for semantic information retrieval. They used PSO as one of the document clustering algorithm to find the nearest optimal solution.

*Fig 6. Information retrieval using PSO [24]:*




### 3.7 Binary Bat Algorithm

The Binary Bat Algorithm is inspired by Yang's 2010 algorithm [15]. The bat set of rules is a brand new technique for optimizing swarm intelligence, in which the hunt set of rules is inspired by means of bats ' social conduct and the echolocation phenomenon for sensing distance. BBA developes A discrete bat set of rules model to resolve selection problems and classifications capabilities. Every artificial bat has a vector of position, speed and frequency in BBA. The position in BBA is either 0 or 1. In BBA each bat is characterized by its position $x_i^t$, velocity $v_i^t$ , frequency $f_i$ , loudness $A_i^t$ and the emission pulse rate $r_i^t$ in a d-dimensional search space. Solution $x_i^t$ and the velocity $v_i^t$ at time step *t* are given by

$$f_i = f_{min} + (f_{max} - f_{min})\,\beta \qquad (3)$$

$$v_i^t = v_i^{t-1} + (X_i^t - X_*)\,f_i \qquad (4)$$

$$X_i^t = X_i^{t-1} + X_i^t \qquad (5)$$

where $\beta \in [0, 1]$ is a uniformly distributed random vector. Here $X_*$ is the current best global location (solution) that is loacated by comparing all the solutions between all the n bats. Sigmoid exchange work can refresh the bats ' position as a folllow:

$$S(v_i^t) = \frac{1}{1 + e^{-v_i^t}} \qquad (6)$$

$$x_i^t = \begin{cases} 1 \text{ if sigmoid function} > \sigma \\ 0 \; if \; sigmoid \; function < \sigma \end{cases} \qquad (7)$$

In the event that the sigmoid capacity is more than σ, at that point position of bat is 1; in the event that the sigmoid capacity is not exactly σ, at that point position of bat is 0. Σ is irregular incentive somewhere in the range of 0 and 1.

Moreover, the loudness Ai and the rate ri of pulse outflow refresh in like manner as the cycle continue.

$$A_i^{t+1} = \alpha A_i^t \qquad (8)$$

$$r_i^{t+1} = r_i^0 [\, 1 - \exp(-\gamma t)] \qquad (9)$$

α and γ are constants.

## IV. USE OF THE EVOLUTIONARY ALGORITHM TO RETRIEVE INFORMATION

There has been a growing interest in using AI tools for IR in recent years. EAs usually do not find algorithms, but offer a strong area - unbiased search capability that can be used in many study tasks as learning and self - organization can be considered as problems optimization in many circumstances. For this reason, in the last decade, the application of EAs to IR has grown. Some of the document image retrieval application shown below:

***4.1 Keyword Based Searching***: This application helps users in document images to locate a specified keyword [25].

***4.2 Document similarity measurement***: It approves the user to retrieve archives by specifying whole report Image as query as an alternative than a keyword.

***4.3 Signature based retrieval***: This utility consists of retrieving files from the database by way of specifying the document containing signature as a query [26].

***4.4 Logo based document retrieval***: Logos are usually used as a statement of the supply and ownership of documents in commercial enterprise and government files. Such application uses a logo - containing document as a query to retrieve from the database all documents containing a logo similar to the query document.

*Retrieving imaged documents in digital libraries*: Document image retrieval helps users search a list of articles stored in digital libraries for specific keywords, titles, subtitles and images.

Among other things, EAs were connected to explain the accompanying IR issues:

***4.5 Automatic indexation of documents:*** Gordon [12] proposed the primary methodology to the report ordering by methods for Genetic algorithm in 1988. The creator proposes to connect more than one portrayal with each record and to adjust it after some time as a decent answer for the issue of the different structures that different client inquiries may introduce while scanning for similar archives. Fan et al [27] proposes a GP indexing algorithm - based function learning to obtain an indexing function for the key term weighting of a documentary collection to enhance the IR process.

***4.6 Document clustering and terms:*** The central idea of Robertson and Willet's [28] is to look for get-togethers of terms appearing with similar frequencies in the reports of a gathering. To do in that capacity, the makers consider a GA assembling the terms without keeping up their underlying solicitation. With five different data sets, they run their proposals. The outcomes got for a similar issue that has a lower run time are great, yet like those of another algorithm.

***4.7 Definition of the query:*** This is the largest group of EAs to IR applications. Every strategy in this gathering utilizes EAs either as an input method for significance or as an algorithm for Inductive Query by Example (IQBE). As an IQBE technique, Chen et al [29] utilize a GA to get familiar with the terms of inquiry that better speak to an applicable record set by a client. They have an IRS based on the model of vector space. Robertson and Willet [30] support a GA to look at an upper destined for significance input procedures in vector space IRSs and inspect its results with Robertson and Spark Jones' [31] review pertinence loads method. Sanchez et al. [32] are proposing a GA in a relevant feedback process to gain proficiency with the term loads of expanded Boolean questions for fuzzy IRS. The system's behavior is studied on a patent collection of 479 documents. Unfortunately, the results obtained in the paper are not shown. Horng and Yeh’s [33] used a GA to adjust the weights of the query term to optimize the nearest query vector. To begin with, they consolidate the model of the Bigram and a Pat - tree way to deal with acquire Chinese information data. Lo pez - Pujalte et al.[34-36] assess the efficiency in the vector space model of a GA with various fitness functions and compare the outcomes with the classic Ide dec - hi method.

***4.8 Query Learning:*** Smith and Smith [37] are proposing a GP - based learning calculation for Boolean IRSs. Despite the fact that they present it as an applicable criticism algorithm, the paper's experimentation is in reality nearer to the IQBE structure. Cordon [38] investigated an augmentation of Smith and Smith's calculation, in view of joining the multi-objective way to deal with the last mentioned. Kraft et al. [39] are proposing an IQBE method for getting to know the entire composition of prolonged Boolean Fuzzy IRS queries.

***4.9 Matching learning function:*** The objective is to utilize an EA to create a likeness degree for a vector space IRS to improve its recovery proficiency for a specific client. This is another input logic of significance since coordinating capacities are adjusted as opposed to inquiries. Pathak et al. [40] propose another weighted coordinating capacity, which is the direct combine of the more than a few comparable functions that exist.

***4.10 Retrieval of images:*** Cho and Lee [41] build up an image recuperation framework dependent on human inclinations and feelings utilizing an interactive genetic algorithm (IGA) to enhance the absence of capacity articulation of the user. The framework extricates the attributes from pictures by changing wavelets and utilizations the IGA to look through the picture that the client has as a top priority when the fitness function cannot be characterized expressly. Kato and Iisaku [42] depict a photograph rebuilding device furthermore basically dependent on a GA with an intelligent component that progressively shows the character client's subjectivity in the results of the recuperation and proposes another strategy for upgrading fortuitous reliance. Another methodology for registering photo closeness is proposed, called local likeness test. It is essentially founded on the idea that recognizing different things in the picture requires different comparability criteria for each thing. Furthermore, a GA - based method for finding a finest mission of similarity criteria to photography areas is proposed and included within the relevant feedback mechanism [43].

***4.11 User profile design for online information retrieval:*** In representing the needs of the user, the lack of customization limits IRSs. In this circumstance, an essential issue is the development of client profiles that keep up recently recovered data identified with the requirements of past clients. In [44], a specialist is proposed to display the client's web seek data needs through a versatile procedure dependent on a GA with fluffy qualities. The GA keeps up the learning of the inclinations of the client and recovers the feed from the client. Larsen et al. [45] present in a person profile a scheme to keep experience - based knowledge about consumer choices. Instead of a recuperation process, the person profile is generated from a filtering system wherein no user statistics is considered. In this manner, the author’s filter the document series the use of the statistics retrieved within the first query. The proposed device uses a GA to find the most discriminatory terms, which allows you to generate the profile, i.e., the ones allowing the device to figure among important and non-significant reports, which may be decided on and saved as a part of the person’s data for use in future queries to the machine. To enhance the exactness over person profiles, Chen or Shahabi [46] endorse an adaptive gentle query (ASQ) system. ASQ consists about joining fundamental components: an off - block tender question system and a lesson mechanism based totally on off - row GA.

## V. CONCLUSION

In this investigation, we have looked into the different uses of EC to IR, breaking down the different sorts of IR issues that have been understood by EAs. EAs have been connected, among different issues: Keyword Based Searching, Document similarity measurement, Signature based retrieval, Logo based document retrieval, Retrieving imaged archives in advanced libraries, Automatic indexation of documents, Clustering of documents and terms, Definition of the query, Query Learning, Matching learning function, Retrieval of images, User profile design for online information retrieval and we have investigated various Evolutionary Family algorithm. In a few cases, the obtained results are extremely encouraging.

## VI. ACKNOWLEDGEMNT

I offer my earnest thanks to my guide Dr. Sandeep Gupta, Associate Professor (CSE Department) for his constant help, worth full guidance and encouragement during the work. I would like to thanks to SVU, Uttar Pradesh for giving me such platform for taking my research work to some heights.

## REFRENCES

[1] G. Salton, M.H. McGill, "Introduction to Modern Information Retrieval", McGraw-Hill, 1983.
[2] R. Baeza-Yates, B. Ribeiro-Neto, "Modern Information Retrieval", Adisson, 1999.
[3] Internet World Stats. "Worldwide internet users", (March 20, 2019). [Online]. Available: http://www.internetworldstats.com
[4] C.H Chang and C.C Hsu. "The design of an information system for hypertext retrieval and automatic discovery on WWW. Ph.D thesis", department of CSE, National Taian University, 1999.
[5] K.L. Kwok. "Comparing representation in chinese information retrieval", ACM SIGIR'97, Philadelphia, McGraw-Hill, 1997.
[6] T. B€aack, D.B. Fogel, Z. Michalewicz, "Handbook of Evolutionary Computation", IOP Publishing and Oxford University Press, 1997.
[7] J. H. Holland, "Adaptation in Natural and Artificial Systems", University of Michigan Press, Ann Arbor, 1975.
[8] K. A. DeJong, "An Analysis of the Behavior of a Class of Genetic Adaptive Systems", Ph.D. Thesis, University of Michigan, 1975.
[9] D. E. Goldberg, "Genetic Algorithms in Search, Optimization", and Machine Learning, Addison-Wesley, Reading, MA., 1989.
[10] H. Chen, "Machine learning for information retrieval: neural networks, symbolic learning, and genetic algorithms". Journal of the American Society for Information Science, 46(3), 1995, pp. 194–216.
[11] J. Savoy and D. Vrajitoru, "Evaluation of learning schemes used in information retrieval (CR-I-95-02)". Universite de Neuchatel, Faculte de droit et des Sciences Economiques, 1996.
[12] M. Gordon, "Probabilistic and genetic algorithms in document retrieval". Communications of the ACM, 31(10), 1988, pp. 1208–1218.
[13] D. Harman. "Overview of the first TREC conference". In Proceedings of the 16th ACM SIGIR conference on information retrieval, 1993, pp. 36–47.
[14] J. Morgan and A. Kilgour. "Personalising on-line information retrieval support with a genetic algorithm". In A. Moscardini, & P. Smith (Eds.), PolyModel 16: Applications of artificial intelligence, 1996, pp. 142–149.
[15] M. Boughanem, C. Chrisment, and L. Tamine. "On using genetic algorithms for multimodal relevance optimization in information retrieval". Journal of the American Society for Information Science and Technology, 53(11), 2002, pp. 934–942.
[16] J. T. Horng and C. C. Yeh. "Applying genetic algorithms to query optimization in document retrieval". Information Processing & Management, 36(5), 2000, pp. 737–759
[17] D. Vrajitoru. "Crossover improvement for the genetic algorithm in information retrieval". Information Processing & Management, 34(4), 1998, pp. 405–415.
[18] S. Das and P. N. Suganthan. "Differential evolution: A survey of the state-of-the-art", IEEE Transactions on Evolutionary Computation 15(1), pp. 4-31. 2011.
[19] R. Storn and K. Price. "Differential evolution – A simple and efficient heuristic for global optimization over continuous spaces", Journal of Global Optimization 11(4), pp. 341-359. 1997
[20] Genetic Programming. Source: http://geneticprogramming.com/
[21] Corne, D.W., Deb, K., Fleming, P.J. and Knowles, J.D., 2003. "The good of the many outweighs the good of the one: evolutionary multi-objective optimization", IEEE Connections Newsletter, 1(1), pp.9-13.
[22] A. Zhou, B. Qu, H. Li, S. Zhao, P. N. Suganthan and Q. Zhang. "Multi-objective evolutionary algorithms: A survey of the state of the art", Swarm and Evolutionary Computation 1(1), pp. 32-49.2011.
[23] J. Kennedy and R. Eberhart, "Particle swarm optimization," Proceedings of ICNN'95 - International Conference on Neural Networks, Perth, WA, Australia, 1995, pp. 1942-1948 vol.4. doi: 10.1109/ICNN.1995.488968
[24] A. Gomathi, J. Jayapriya, G. Nishanthi, K. Pranav, and G. Praveen Kumar, "Ontology based semantic information retrieval using particle swarm optimization," International Journal of Applied Information Communication Engineering, vol. 1, no. 4, pp. 5–8, 2015.
[25] D.Niyogi and S. Srihari 1997, "The use of Document Structure Analysis to retrieve information from Documents in Digital Libraries", In Proc. SPIE, Document Recognition IV, Vol 3027, pp.207-218.
[26] S.Srihari et al, "Document image Retrieval using signature as Queries", In Proc. Of Second International Conference of Document Image Analysis for Linraries.
[27] W. Fan, M.D. Gordon, P. Pathak, "Personalization of search engine services for effective retrieval and knowledge management", in: Proc. 2000 International Conference on Information Systems (ICIS), Brisbane, Australia, 2000
[28] A.M. Robertson, P. Willet, Generation of equifrequent groups of words using a genetic algorithm, Journal of Documentation 50 (3) (1994) 213–232
[29] H. Chen et al., "A machine learning approach to inductive query by examples: an experiment using relevance feedback", ID3, genetic algorithms, and simulated annealing, Journal of the American Society for Information Science 49 (8) (1998) 693–705

[30] A. Robertson, P. Willet, "An upperbound to the performance for ranked-output searching: optimal weighting of query terms using a genetic algorithm", Journal of Documentation 52 (4) (1996) 405–420

[31] S.E. Robertson, K. Spark Jones, "Relevance weighting of search terms", Journal of the American Society for Information Science 27 (1976) 129–145

[32] E. Sanchez, H. Miyano, J. Brachet, "Optimization of fuzzy queries with genetic algorithms.Applications to a data base of patents in biomedical engineering", in: Proc. VI IFSA Congress,Sao-Paulo, Brazil, 1995, pp. 293–296.

[33] J. Horng, C. Yeh, "Applying genetic algorithms to query optimization in document retrieval", Information Processing and Management 36 (2000) 737–759

[34] C. Lo pez-Pujalte, V. Guerrero, F. Moya, "A test of genetic algorithms in relevance feedback", Information Processing & Management 38 (2002) 793–805.

[35] C. Lo pez-Pujalte, V. Guerrero, F. Moya, "Genetic algorithms in relevance feedback: a second test and new contributions", Information Processing & Management 39 (5) (2003) 669–807.

[36] C. Lo pez-Pujalte, V. Guerrero, F. Moya, Order-based fitness functions for genetic algorithms applied to relevance feedback, Journal of the American Society for Information Science and Technology 54 (2) (2003) 152–160.

[37] M.P. Smith, M. Smith, "The use of genetic programming to build Boolean queries for text retrieval through relevance feedback", Journal of Information Science 23 (6) (1997) 423–431

[38] O. Cordo n, E. Herrera-Viedma, M. Luque, "Evolutionary learning of Boolean queries by multi-objective genetic programming", in: Proc. PPSN-VII, Granada, Spain, 2002, pp. 710–719, LNCS 2439.

[39] D.H. Kraft, F.E. Petry, B.P. Buckes, T. Sadasivan, "Genetic algorithm for query optimization in information retrieval: relevance feedback", in: E. Sanchez, T. Shibata, L.A. Zadeh (Eds.), Genetic Algorithms and Fuzzy Logic Systems, 1997, pp. 155–173.

[40] P. Pathak, M. Gordon, W. Fan, "Effective information retrieval using genetic algorithms based matching functions adaption", in: Proc. 33rd Hawaii International Conference on Science (HICS), Hawaii, USA, 2000.

[41] S. Cho, J. Lee, "A human-oriented image retrieval system using interactive genetic algorithm", IEEE Transactions on System, Man and Cybernetics. Part A: Systems and Humans 32 (3) (2002) 452–458.

[42] S. Kato, S. Iisaku, "An image retrieval method based on a genetic algorithm", in: Proc. Twelfth International Conference on Information Networking (ICOIN-12), 1998, pp. 333–336.

[43] S. Kato, S. Iisaku, "An image retrieval method based on a genetic algorithm, in: Proc. Twelfth International Conference on Information Networking (ICOIN-12), 1998, pp. 333–336

[44] M.J. Martin-Bautista, H. Larsen, M.A. Vila, "A fuzzy genetic algorithm approach to an adaptive information retrieval agent", Journal of the American Society for Information Science 50 (9) (1999) 760–771.

[45] H. Larsen, N. Mar ın, M.J. Mart ın-Bautista, M.A. Vila, "Using genetic feature selection for optimizing user profile", Mathware & Soft Computing 7 (2000) 275–286.

[46] Y. Chen, C. Shahabi, "Automatically improving the accuracy of user profile with genetic algorithm", in: Proc. International Conference on Artificial Intelligence and Soft Computing, Cancun, Mexico, 2001.