Linguistic Search by C.S. Boosting Optimization Algorithm

[1]B Shashirekha, Assistant Professor, Department of CSE (AI & ML), Institute of Aeronautical Engineering, Hyderabad, b.shashirekha@iare.ac.in

[2]G Akhila, Assistant Professor, Department of CSE (AI & ML), Institute of Aeronautical Engineering, Hyderabad, g.akhila@iare.ac.in

[3]NM.Deepika Assistant Professor, Department of CSE, Institute of Aeronautical Engineering, Hyderabad, nm.deepika@iare.ac.in

[4]T Sahithi, Assistant Professor, Department of CSE, Institute of Aeronautical Engineering, Hyderabad, t.sahithi@iare.ac.in

[5]V.Sudheshna, Assistant Professor, Department of CSE, Institute of Aeronautical Engineering, Hyderabad,v.sudheshna@iare.ac.in

[6]B Padmaja, Associate Professor, Department of CSE (AI & ML), Institute of Aeronautical Engineering, Hyderabad, b.padmaja@iare.ac.in

Article History: Received: 10 December 2020; Revised 12 February 2021 Accepted: 27 February 2021; Published online: 5 May 2021

Abstract: The internet data developed incrementally and gathered as unstructured information like textual documents, pictures, videos and sound content etc. This information material is created daily from several social networking resources such as sites, Twitter news posts, community websites, even Face book, etc. Information sharing is occurring across programs of various businesses like healthcare, education, societal sites, and army forums, community-based Q.A. sessions, client attention, etc. Information retrieval methods are used for getting relevant data from acollection of tools to resolve consumer requirements anywhere at any moment. Users expect precise responses with a brief response time instead of a listing of files during internet searches. Data sources show up in various formats, and thus it's pretty tricky to recover data without processing for simple recovery. The construction and arrangement of information in a predetermined format are required, and it's accomplished by data preprocessing and clustering from its comparable classes. In this paper, a cuckoo searchoptimization algorithmis proposed to reduce the search space and time complexity. An experimental result demonstrates that the algorithm generated greater accuracy in contrast to state-of-art procedures.

Keywords: Information retrieval; cuckoo search optimization, Particle swarm optimization algorithm

I. Introduction

Information retrieval (I.R.) will be the procedure of obtaining necessary information from the group of information resources that hunts in a record. Additionally, it uses metadata describing data in addition to databases of texts, pictures, or sounds. Search engines would be the best possible application section of I.R. (Kolomiyets 2011). The look for advice is the practice of obtaining necessary information with indexation based online articles. On account of heaviness ofir material, the search or recovery of information from the internet database takes additional hours, and the insignificant retrieval is high. To overcome the issues mentioned above, an automated data recovery procedure uses to decrease information overload. Several traditional approaches like machine learning models. Natural Language Processing (NLP), similarity calculations, and also record group techniques are embraced.

As a result of immense development of unstructured data on the World Wide Web, there's a vital requirement to grow systems to bargain with servicing the info needed community. The practice of answer production to get an individual query might be acquired by employing numerous methods; however, also the accuracy of QAS is quantified concerning answer appropriateness using a brief response period. To attain this extraction of applicable documents from the search space, these documents/sentences need to be lessened by employing various optimization methods. This chapter discussed the suggested candidate answer creation utilizing Cuckoo Search Optimization (CSO) algorithm.) Additionally, it explained the beneficiaries of employing CSO in record clustering and employed it in instruction scenarios.

Another level of I.R. system is personal information retrieval. Recently customers in Apple's Mac OS X Spotlight or Windows Vista's Instant Search used consumer operating systems with integrated information retrieval process. For example, Email provides the search according to the text, filters the junk mail in spam, and classifies the mail manually or automatically to place them in appropriate folders. There are various issues available to handle and maintain the wide range of document types and provide maintenance-free search systems. I.R. process starts with entering the query into the system, and these queries are formal statements of information needs such as search strings in web search engines. But the query does not match the single entity in the set of data: At the same time, variousentities are matched with the query, possibly with different levels of relevancy. Entities are denoted as information in a content collection or database. The queries given by the users are matched with the information of database. Each I.R. system computes the score for every entity in the database, which match the query and rank it with the score. The user has to refine the query for every iteration if they need an appropriate result.

Literature Review

AbdessamadEchihabi et al. (2006) Suggested a Q.A. system that employs the top layout mechanically adjusts to inspect the question and answer routines. The primary approach uses calculations that depend on rich knowledge-bases benchmark data set TREC 2005 and 2007, complex syntactic and semantic processing. The following approach employs unsupervised learned patterns employing computational biology-inspired recovery algorithms. The following approach utilizes statistical noisy-channel algorithms very similar to those found in machine interpretation. This Boolean information retrieval version's disappointing outcomes have transferred investigators to some other means of solving problems using cognitive algorithms. The candidate answer scores have been calculated by mapping the exact question and answer layout types with all the length and quantity of mutual search phrases and phrases.

Xin-She Yang &Suash Deb (2010) Performed a broader comparative study that was shared with some typical test purposes and stochastic evaluation purposes. Additionally, it implemented the Cuckoo search engine to solve optimization issues. The best solutions developed by the Cuckoo hunt are a lot better compared to the most effective solutions made by particle swarm optimization. Additionally, it analyzed using this genetic algorithm together with user responses to select weights to search terms in a query together side the population size and the achievement of developments when along using search engine optimization.

Sweta P. Lende& M. M. Raghuwanshi (2016) Go over different methodology and execution information of question answering system into shut domain Q A System for tackling documents with the instruction domain utilizing NLP methods to recover more precise responses. TannazAlinaghi&ArdeshirBahreininejad (2013) discussed supporting the material of questions and available tools, including course materials, FAQS, and evaluation from various students applying a recommender system. The most suitable answer(s) regarding several conditionals like student's comprehension, research background, history of earlier questions, and the candidate replies applicable to the question is going to be maintained.

Methods for optimization

Traditional optimization techniques help find the maxima or minima optimum solution from all feasible solutions of continuous or discrete nature. Using optimization aims to achieve the best outcome with relative prioritized criteria such as reducing error, search space rate and response time, etc. The analytical and differential calculus methods are applied to locate the optimum solution. Under given constraints conditions, achievable performance is obtained by maximizing efficiency result outcome and minimizing undesired ones about time consumption.

Several optimization techniques based on evolutionary and natural inspired algorithms have been introduced to obtain the global and local optimum solution. The outcome of optimization algorithms lies in high-quality results, high convergence rates, complex structures and deciding the values for the number of input parameters.

The answers can be extracted from documents in the cluster of knowledgebase to frame high-quality clusters; cluster form is tilt & classified by grouping entities. Optimization algorithm improves the quality of clustering and reduces time complexity, but the success based on the specific controlling parameters such as several iterations and population size. Generally, the distance between a cluster of similar nature is measured by Euclidean distance, city block etc. It is also compared based on purity, the average distance between documents, F-Measure, Entropy, etc. The amount of additional meaningless information that occurred in the document is called noise. Some more noisy data in the unstructured text cannot be understood and interpreted correctly by machines. Statistical learning technique has been used to deal with noisy/incomplete data. By eliminating the noise, the quality of cluster can be improved to handle noisy data, and the following standard techniques can be adopted.

Application of domain-specific cleansing and normalize techniques, such as Image de-noising. Video de-noising, Band-pass ring (when deals with time series), De-noising Autoencoders, Stemming & stops words in NLP) etc.

Cross-validation with k-folds to generalize error, split of data depends on the researcher and scenario. For example, the ratio is given as training set (75%) & test set (25%).

The various optimization techniques are available to optimize the document results for retrieving the answers. Naturally, inspired algorithm possesses more advantage than conventional approaches due to its accuracy level and less complexity. The following sections briefly explain the concept of nature-inspired algorithms.

Cuckoo search optimization

The cuckoo search optimization algorithm has been combined with the SWAG model to reduce the time and space complexity in answer generation.The optimization algorithms play an essential role in obtaining the best accurate results from the retrieved answers. The effect of a variety of techniques to achieve high accuracy of results retrieved, or hybrids of two or more of optimization algorithms are optimized using. The recent optimization algorithm chosen for efficient search method on most problems is natured inspired optimization algorithms. In the proposed research framework, cuckoo search optimization is acclimatized, document clustering, and extract the best sentences from the list of sentences resultant from the SWAG model. The advantage of choosing the cuckoo search optimization is its stochastic nature, non-deterministic, and it balances local and global optima efficiently with the limited number of parameters. The proposed method formulates the Cuckoo Search breeding behaviour incorporated to apply optimization, and their three categories are as follows.

- Each cuckoo selects one sentence simultaneously and leaves it in a chosen relevant nest population for the following possible answer sentence.
- The best population with high quality of sentence will carry to the next generations.
- The number of available host nest population is fixed if a current possible sentence possibly identifies the sentence for their relevance with the probability of pa 0.1. Low-quality answers are discarded or passed to the newly built generation.

In this case, the host bird can either take back the egg or ignore it and build an entirely new nest.

The procedure of Cuckoo Search Optimization algorithm is shown in Figure 1.

$pseudo\ code\ for\ cuckoo\ search\ optimization\ algorithm$

$Input: Object\ size\ N, itmax\ is\ the\ maximum\ number\ of\ iterations,$

$Tolerance\ tol, functions\ is\ objfun, worest\ solution\ Pa$

$step\ 1: Population\ size\ is\ N\ which\ is\ generated\ randomly$

$step\ 2: The\ given\ obj\ function\ will\ be\ used\ ti\ analyze\ the\ fitness\ function$

$step\ 3: Get\ the\ max\ fitness\ value\ which\ is\ optimum\ \tan given$

$step\ 4: whle\ itmax$

$step\ 5: Find\ optimal\ new\ populattion\ and\ put\ existing\ best\ nests$

$step\ 6: find\ the\ fitness\ val\ for\ newly\ generated\ nests$

$step\ 7: Delete\ the\ nests\ with\ worest\ calculus\ i.e\ based\ on\ Pa\ value$

$step\ 8: find\ the\ fitness\ for\ the\ remaining\ nests\ using\ given\ existing$

$step\ 9: Get\ themax\ fitness\ values\ as\ the\ optimal\ solution$

$step\ 10: Display\ the\ optimal\ solution\ which\ is\ best\ nest$

Figure 1: CSO algorithm

Cuckoo Search optimization locates the best way to solve an individual, the recovered sentence with different iterations. The locations are made up of initial pair of recovered paragraphs; number ratio should be come outside for improved results. The initial production is d e, along with also the cuckoo hunt is put to the neighborhood look for those phrases. The birds' footprints have a tiny pair of paragraph /paragraphs; the health value can be utilized for the connection between the sentences along with question key word. Cuckoo nest values have been added, updated, and removed during each iteration by applying the cuckoo search algorithm with the SWAG model to obtain better results. All nests of each sentence group and subgroup were updated simultaneously with updating of levy flight value. The nature of cuckoo bird is to eliminate the host bird egg or the nest itself when it identifies another bird's egg in its own nest. By adopting this bird's nature, the candidate solution for the given query is not found to be related to the stimulated fraction pa. The nest containing related sentences are removed along with the nest. The optimum prey solution is substituted randomly with the new related nest on the search space. The best solution is kept unchanged& saved in the nests, and automatically carried out to the next iteration, which is indicated as the following selection for a solution.

The number of clusters with different values on various dynamic nests is used for searching the related one with initialized centroids values. At the time of implementation, active nests are updated after each random walk to detect the next best optimal nest. Random walk in the process of step-lengths with a probability distribution. It is defined as a walk in a search space of dimension greater than one, the steps made in random directions.

For the efficient document cluster, the selection of sentences will be y and time proficient. The levy flight is a more efficient search space scenario that produces a new solution from the current bear solutions. Like memetic algorithms, Cuckoo search capabilities explore values on L cal search with self-improvement ability. The fitness function is calculated for all the retrieved top 'n' sentences. If the fitness function value for the new cuckoo egg (sentence) is better, then the current sentence value will be replaced with the current best one. Similar to that, the cuckoo footprints are calculated with all the physical exercise worth. In the event the physical exercise price of cuckoo 'I' is higher than the exercise price of cuckoo 'j' replace the cuckoo j with the brand newest solutions. The nest using low physical fitness values have been lost; the newest Testament solution is passed and stored into the following creation. The procedure is repeated until it reaches the optimum value from the best worldwide alternative for better outcome with a endurance speed of 0.05 percent.The fitness function is calculated for the average number of mutual keywords between each sentence of paragraph (sentence) and question (Q), which is calculated using the equation below

$$f(x) = \frac{\sum_{i \epsilon N_s} \text{Keywords}(Q) \cap \text{Keywords sentance}_i}{N_s X \text{ Keywords}(Q)}$$

The exact similarity has been evolved to investigate that the similarity among recovered phrases. Sometimes, the sentencing bit can overlap with each other on account of identical words in just two paragraphs; however they disagree in their significance. To handle this particular specific arrangement of phrases, facets must be profoundly studied with opinion investigation. Sentiment analysis classifies the sentiments of a sentence from negative to positive (value ranges from 0 to l) for the input sentence containing n-words. In the given example sentence I: Cat ate the mouse as food and sentence 2 Mouse ate the cat food to explain this overlapping concept. In the below example, sentences consisting of words such as cat, food, mouse, and ate arranged in different order provide different meanings.

Ambiguous words in the question are considered, i.e. different words which provide the same meaning. It is resolved by analyzing the text ofser question and behaviour considered for answer extraction with machine learning models. The concept is explained with the following example sentence 1: Samsung J7 phone is good, fast and has incredible features and sentence 2: Samsung J7 phone is not fit into my pocket (it's about the money/ size feature). The above sentence words defining the mobile phone context as 'good, fast and awesome features' are equal to money features that do not fit into my pocket. The two sample questions for the same word gives different meaning such as 'How much an Apple laptop costs?" and What is the right time to eat an Apple in day time or night time?'. The word 'Apple" has occurred in both sentences. Still, context meaning will be changed. It should be resolved by considering the contextual meaning of question for answering.

Technical challenges can arise for the same question asked in different ways, such as i) What was the date Telanga became a state?, ii) When was the state Telangana created?, iii) When Telangana's entered into the state level? .The meaning of all the questions is 'When Telangana state is formed?".It deals with semantic matching, which needs to map fragments of questions to entities and predicates. In return, it reduces the ample search space, and several semantic parses grow epidemic. From a pair of recovered answer, candidates are ranked based on the odds of accuracy and value. By way of instance, the paragraphs have been recovered answers for its "Where's Mahatma Gandhi comes into the world?" Id socialized together with the proper answer in line with the circumstance. Sentence l: "Mahatma Gandhi (person) was born on 14 March 1879 (Date)"

Sentence2: "Mahatma Gandhi (Person) was born in Germany (Country)"

Sentence3: "Mahatma Gandhi (Person) was born in a Hindu (Religion) family"

The sentences are displayed in the user interface output for the best score on the question of sentiment. Fitness value of the minimum sentences for the first 5 answers retrieved candidate.The advantage of cuckoo search is its uncomplicated nature in providing a solution with limited parameters compared to other algorithms. Performance evaluation of CSO results is compared with Particle swarm optimization algorithm (PSO). PSO, which is a much more equivalent procedure than CSO, has distributed parameters in search attributes.

Results and Discussion

The proposed SWAG model with CSO experiments on user query and benchmark dataset 20newsgroup as the data source. Compared with existing, the test results are evaluated with the standard metrics as F1-score, Missrate, Fallout and Mean Average Precision (MAP) with empirical equation are discussed. The number of 10 nests has been chosen approximately for the cluster formation. A maximum of 8 iterations is fixed because, after this 8th iteration, there is no improvement in the answers convergence. The parameters applied for CSO is as shown in Table 4.1 below.

Parameters of cuckoo search optimization

Parameter value	Value
Number of nests	10
Dimensions	KD+1
Levy exponent	2
Max iterations	8
Lower and upper bound values for centroids	(-1,-1) and (1,1)

Mean Average Precision (MAP) is just one among popular performance climbs in information retrieval to appraise the position of retrieved relevant records with the normal accuracy values. The suggested SWAG algorithm increases MAP for sentence recovery by the phrase clusters created with all the 20newsgroup data set. It enriches the recovery rate, including 0.39 to 0.42, using baseline, ACQUAINT-bigram along with Google bi-gram. It's likewise recovery based on word events from the phrases that will be raised in 0.35 to 0.40 at a DKS-KBC algorithm for forming the expression clusters.The comparisons F1-score values for SWAG, SWAG+PSO and SWAG + CSO algorithm for a minimum of 8 iterations for answer convergence is shown in Figure-4.3. Comparison of Missrate and Fallout measure values for SWAG model. SWAG + PSO model and SWAG +CSO model are shown in Figure 4.4 & Figure-4.5.

Table 1: F1-score values comparison table for different algorithms

No of iterations	**swag**	**Swag and particle swam optimization algorithm**	**Swam and cuckoo search optimization algorithm**
0	0.44	0.56	0.62
0.1	0.45	0.58	0.63
0.2	0.47	0.57	0.65
0.3	0.45	0.64	0.64
0.4	0.49	0.57	0.58
0.5	0.54	0.68	0.70
0.6	0.54	0.70	0.71
0.7	0.55	0.68	0.72

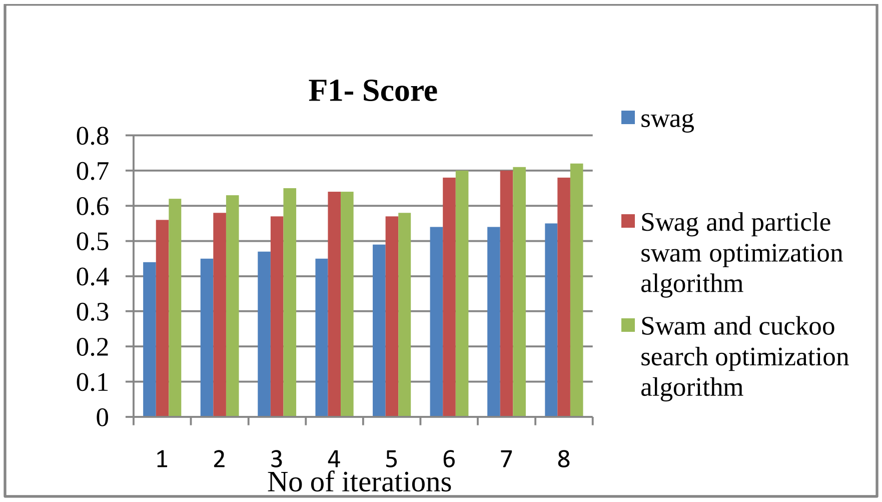

Figure 2: F1-score values comparison

Table 2: Miss Rate values comparison table for different algorithms

No of iterations	swag	Swag and particle swam optimization algorithm	Swam and cuckoo search optimization algorithm
0	0.60	0.51	0.45
0.2	0.61	0.50	0.46
0.4	0.63	0.52	0.48
0.6	0.64	0.44	0.41
0.8	0.58	0.54	0.51
1.0	0.56	0.38	0.35
1.2	0.54	0.42	0.4
1.4	0.50	0.40	0.34

Figure 3: Miss Rate comparisons

Table 3: Fallout Rate values comparison table for different algorithms

No. of iterations	swag	Swag and particle swarm optimization algorithm	Swam and cuckoo search optimization algorithm
0	0.23	0.15	0.12

0.2	0.23	0.13	0.1
0.4	0.22	0.12	0.04
0.6	0.24	0.12	0.09
0.8	0.18	0.13	0.10
1.0	0.17	0.07	0.07
1.2	0.18	0.06	0.06
1.4	0.18	0.08	0.07

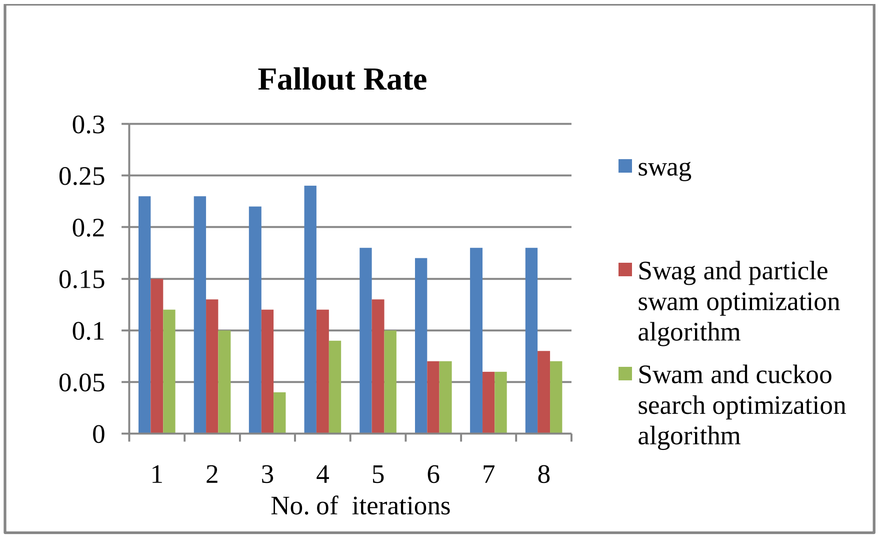

Figure 4: Fallout rate comparisons

Experimental results of proposed approach are discussed and verified for the answer validation task to review the answer returned by QAS for accurateness and relativity to the user query. The evaluation of system is carried with 200 user input questions, in which the system outputs answers for all input questions 200/200. But it rightly answered 172 questions out of 200 questions reaches 86% accuracy. System performance shows with the precision value 0.86, recall 0.78, and F1-Score value is 0.82. The proposed approaches such as SWAG, SWAG + PSO and SWAG + CSO are experimented on the dataset with various lists of various user question types and analyze their performance.

Optimization techniques are used to optimize the SWAG model for reducing response time in retrieving resultant sentences. In this scenario, PSO and CSO are considered for optimizing sentence retrieval. The Fl-score results are obtained for sentence retrieval from benchmark datasets through various iterations.

Based on the accuracy analysis of F1- score for SWAG +CSO & SWAG + PSO, the SWAG+ CSO classifies the answers with a high actual positive rate from the knowledgebase. The aspect of probability in omitting irrelevant answers was also found to be high in SWAG + CSO. The proposed CSO algorithm outperforms with 91% of accuracy.

Based on the Miss Rate analysis of SWAG +CSO & SWAG + PSO, the SWAG + CSO model improves the classification on true positive, which yields a better accurate answer by reducing the false negatives. Based on the Fallout analysis of SWAG +CSO & SWAG + PSO, The SWAG + CSO model improves the classification on false-positive, which yields better accuracy answers by decreasing true negatives. Cuckoo Search Optimization (CSO) yields the best results with limited parameters on easy answer convergence of global and local optima among various existing optimization techniques.

Conclusion and future work

From the suggested approach, optimization methods have been utilized to maximize the SWAG version's outcome to decrease response time. Cuckoo Search optimization (CSO) was employed in the consequent recovered paragraphs from the SWAG version. It supplies a listing of best solutions on local and global optima working with a fitness function. The outcomes obtained by the suggested CSO have demonstrated it capitulates better results in comparison with Particle Swarm Optimization (PSO). In numerous iteration runs completed, the results obtained by the SWAG algorithm are 6-9 %, SWAG + PSO is currently 73 per cent, and SWAG +CSO is currently 78 per cent. This proposed work could lead to success and employ into instruction domain that helps to master community. From the 3rd suggested strategy, the detailed questions can be answered since machine-generated summaries using the reinforcement learning procedure for machine learning. It outlines replies for user inquiries employing the benefit function; the educational version's data are trained throughout the human-generated summaries.

The upcoming enhancement would produce a dialogue agent-based QAS that uses 3D facial expressions and speech that disseminates real artistic encounter into the contested users. Additionally, it specializes in replying to multi-sentence questions since I'm travelling in Delhi summertime. Which exactly is tourist spots offered in Delhi? It creates portable personal assistant applications for real-world situations with multi-faceted & trending requirements, including tourism, healthcare and education program regions, etc. It had progressed in natural language knowledge, question priority, and time stamping automatic creation of similar questions together using profound learning techniques based on the preceding scenario.

Conflict of interest statement: All author declare that they do not have any conflict of interest.

References

[1]. AbdessamadEchihabi, Ulf Hermjakob , Eduard Hovy, Daniel Marcu, Eric Melz& Deepak Ravichandran 2006, “How to Select Answer Sring', Springer Netherlands”.

[2]. KristjanArumae&Fei Liu (2018). “Reinforced Extractive Summarization with Question-Focused Rewards", Proceedings of ACL 2018, Student Research Workshop”. copy right to Association for Computational Linguistics. pp. 105–111, July.

[3]. LiangZhenqiu2012, “Design of Automatic Question Answering System Base of CBR”, Proceeding Engineering 29, pp.981-985.

[4]. LotfiZadeh 2006, “From Search Engines to Question Answering Systems - The Problems of World Knowledge”. Relevance. Deduction &Precisiation', Elsevier- FUZZY LOGIC AND THE SEMANTIC WEB, volume 1. pp. 163-209.

[5]. Manoranjitham G &Veeraselvi SJ. 2013. “Mobile Question AnsweringSystem based on Social Network”. International Journal of Advanced Research in Computer & Communication Engineering September. pp 3620-3624

[6]. Man-Ching Yuen, Irwin King &Kwong-Sak Leung 2011. “A Survey of Crowdsourcing Systems”. 2011. IEEE International Conference onPrivacy, Security. Risk, & Trust, & IEEE International Conference on Social Computing. IEEE. pp. 766-773.

[7]. FarshadKiyoumarsi 2014, “Evaluation of Automatic Text Summarizations Based on Human Summaries”, 2nd GLOBAL CONFERENCE on LINGUISTICS & FOREIGN LANGUAGE TEACHING E LT-2014. Procedia - Social & Behavioral Sciences, E SEVIER.. 83 - 91.

[8]. Fernando Zacarias F2009. MOA: “Question Answering in Mobile devices”, Benemérita Universidad Autónoma de Puebla, pp.44-55.

[9]. Jade Goldstein, Vibhu Mittal, Jaime Carbonell& Mark Kantrowitz2000. “Multi-document Summarization by Sentence Extraction”, Proceedings of 2000 NAACLANLP Workshop on Automatic Summarization, ACL Volume 4, pp.40-48

[10]. George A. Miller 1995. “WordNet: A Lexical Database for English Communications of ACM”, Volume 38. Issue 11, pp- 39-4 1.

[11]. Gang Liu &TianyongHao 2012. “User-based Question Recommendation for Question Answering System”. International Journal of Information & Education Technology volume 2, Issue 3, pp. 243-246

[12]. Abhishek Sharma & Kenneth D.Forbus 2010, "Graph-Based Reasoning & Reinforcement Learning for Improving Q/A Performance in Large Knowledge-Based Systems". Association for the Advancement of Artificial Intelligence, pp. 96 -101

[13]. Carlos Cobos, Henry Muñoz-Collazos, RicharUrbano-Muñoz, Martha Mendo., Elizabeth León & Enrique Herrera-Viedma d 2014, "Clustering of Web Search Results based on Cuckoo Search Algorithm & Balanced Bayesian Information Criterion", Information Sciences, pp.248- 264.

[14]. Chin-Yew Lin 2004, ROUGE: "A Package for Automatic Evaluation of Summaries', In Proceedings of Workshop on Text Summarization branches out Post-Conference Workshop of Association for Computational Linguistics", Barcelona, Spain, pp. 74-81.

[15]. Alexander Clark, Chris Fox & Shalom Lappin, (Eds.) 2010. "The Handbook of Computational Linguistics & Natural Language Processing"

[16]. BabuKartik. R.G. &Dhavachelvan. P. 2012, "Hybrid Algorithm by the advantage of ACO & Cuckoo Search for Job Scheduling", International Journal of Information Technology Convergence & Services, Vol. 2, No. 4, p. 25-34.

[17]. Bahadir Ismail Aydin, YavuzSelim Yilmaz, Yaliang Li, Qi Li, Jing Gao& Murat Demirbas 2014, "Crowdsourcing for Multiple-Choice Question Answering Association for the Advancement of Artificial Intelligence" (www.aaai.org), pp.2946-2953.

[18]. BalaSubramanian.N, James Allan & Brue Croft. W 2007. "A Comparison of Sentence Retrieval Techniques". In Proceedings of Annual International ACM SIGIR Conference on Research &Development in Information Retrieval ACM, pp. 813-814.

[19]. Ali GhobadiTapeh&MaseudRahgozar 2008, "A knowledge-based question answering system for B2C e-Commerce", Elsevier Knowledge-Based Systems 21 (2008), pp.946-950.

[20]. Amit Mishra & Sanjay Kumar Jain 2016. "A survey on question answering systems with classification". Journal of King Saud University - Computer &Information Sciences (2016) 28, pp. 345 361