

## Biogeography-Based Optimization Algorithm for Automatic Extractive Text Summarization

Seyed Hossein MirShojaee, Behrooz Masoumi\* & Esmail Zeinali

*Seyed Hossein MirShojaee*, Department of Computer and information technology Engineering, Qazvin Branch, Islamic Azad University

*Behrooz Masoumi*, Department of Computer and information technology Engineering, Qazvin Branch, Islamic Azad University

*Esmail Zeinali*, Department of Computer and information technology Engineering, Qazvin Branch, Islamic Azad University

### KEYWORDS

Text Summarization,  
BBO Algorithm,  
Extractive Method,  
Online Sources,  
NP complete.

### ABSTRACT

Given the increasing number of documents, sites, online sources, and the users' desire to quickly access information, automatic textual summarization has caught the attention of many researchers in this field. Researchers have presented different methods for text summarization as well as a useful summary of those texts including relevant document sentences. This study selects extractive method out of different summarizing methods (e.g. abstract method). Extractive method involves summarizing text through objective extraction of some parts of a text like word, sentence, and paragraph. A summarization issue would be unsolvable by exact methods in a reasonable time with considering documents with high amount of information (NP complete). These kinds of issues are usually solved using metaheuristic methods. A biogeography - based optimization algorithm (BBO), which is a new metaheuristic method in the domain of extractive text summarization is used in this article. This method is tested on a set of Doc's standard documents in 2002 and is analyzed, using ROUGE software. The obtained results of these tests show that this kind of method can be used as an effective method for text summarization.

© 2017 IUST Publication, IJIEPR. Vol. 28, No. 1, All Rights Reserved

### 1. Introduction

Given the increasing information and sites and online sources as well as the necessity of people's quick access to the internet, automatic text summarization has caught the attention of many

researchers. Das and Martins [1] presented text summarization is one of the ways which makes it possible for users to overview texts. Therefore this method helps users at the time of decision making. Hovy [2] classified text summarization into extractive and abstract summarization. There are also other classifications for text summarization. Gupta [3] Presented Text summarization can be done based on the number of documents such as a single document or

\* Corresponding author: *Behrooz Masoumi*

Email: [masoumi@qiau.ac.ir](mailto:masoumi@qiau.ac.ir)

Received 6 March 2017; revised 3 May 2017; accepted 11 June 2017

multiple documents or it can be done based on the aim of summarization such as inquiry or general summarization. extractive summarization, some parts of a text such as a sentence or paragraph are extracted objectively. Abstract summarization which is done using linguistic methods is based on a text comprehension and a text rewritten in a few sentences. Different methods were proposed for extractive text summarization. Ledeneva et al. [4], Garcia and Ledeneva[5] presented One of the important methods in the domain of text summarization is TFIDF. In this method summarization is based on the repeated words of sentences and a reverse of repeated sentences. Repeated sentences refer to sentences of a document with the same words. In this method some parts of every sentence may be repeated in other sentences. An advantage of this method is its easy usage and its disadvantage is that some unimportant words are repeated frequently which make summarization awful. Zhang and Li [6] presented Documents are usually written based on different subjects. That is, sentences are written in an explicit or implicit form. Some people summarize texts, using a clustering method. Based on this method, the amount of similarity among the sentences is studied according to some parameters and similar sentences are put in one cluster. Therefore, every cluster shows one subject. Then in each cluster, the sentences with the highest similarity to the subject which have obtained high scores are selected for summarization. An advantage of this method is that it is possible to identify all subjects of each text very well and a disadvantage of this method is that the number of clusters have an important role in the results of the study. There can be a large or small number of clusters. On the other hand, it is difficult to select an appropriate number of sentence clusters from one textual document Chuang and Yang [7]. In another method of extractive text summarization presented a set of data labeled by human beings are used. On the other hand, in this method, input texts and their summaries are at hand. In this method, first a text is divided into different parts based on some parameters and each part is displayed based on some features such as the amount of word repetition, word part of speech, and the number of title words. Following the extraction of the features from each part, a learning method is used to train the people who summarize the texts. Song et al.[8], and Suanmali et al. [9] presented These methods include a

decision tree, Byzh law, neural networks, and fuzzy logic. A decreased precision in the case of larger documents as well as a decrease in speed due to a large number of comparisons are the main disadvantages of this method. One of the main disadvantages of a fuzzy logic method is that the unsuitability of defined rules results in a decrease in precision. Meta-heuristic methods such as genetic algorithm (GA), [10, 11] and particle swarm optimization (PSO)[12] and bacterial foraging optimization algorithm (BFOA)[13] are another set of methods used for extractive text summarization. In meta-heuristic methods, sentences with high scores are selected. A main challenge in extractive summarization is that there is a large amount of information in documents and many studies should be done on them. So, this method cannot solve the issue. In larger texts which consist of more words, scoring and more importantly selecting sentences are very difficult. Therefore, in this method, summarization speed and precision will be decreased. In such a condition, using meta-heuristic optimization methods are helpful. Although optimization methods are not used in local minimums and maximums. a biogeography-based optimization method is used in this article. The results show that this method has better performance than other previous methods. This article is organized into five parts. In the first part of this article, an introduction is presented followed by the definition of the major concepts in the second part. In the third part of this article, the method of the study and its features are described. In the fourth part of this article, the tests are conducted and their results are compared. Finally, the fifth part concludes this study.

## **2. Definitions and Concepts**

### **2-1. The process of extractive text summarization**

In text summarization, important information of one source and different sources are put together for the purpose of users' usage. In an extractive summarization method, important sentences and paragraphs of a text are put together to make a shorter form of the text [14]. This method has a lot of advantages. Some of its advantages include its simplicity, high rate of summarization, and expense reduction. Also, in this method, users spend less time for studying text information. A disadvantage of this method is that the amount of extracted sentences may be more than a normal amount. Moreover, as the

important and relevant information may be distributed among sentences, an extractive method cannot identify them. Extractive summarization is done in two steps including preprocessing and processing steps based on Fig. 1. In the preprocessing step, the end of the sentences will be identified, the words which do not have significant meaning are omitted, and the root of the words is specified. In the processing step, the amount of sentence effects and its relationships to a main subject are identified and a specified score is given to them. Finally, sentences with the highest scores are selected for text summarization.

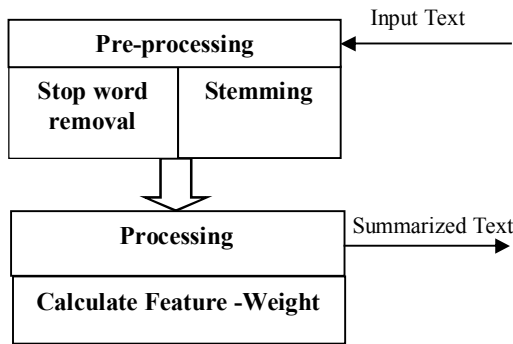


Fig. 1. An Overall Architecture of Extractive Summarization Method [3]

2-2. Biogeography-based optimization method

As new global optimization algorithm a biogeography-based optimization method (BBO) is based on a biogeography theory and studies the geographical distribution of biological organisms. A biogeography-based optimization algorithm like genetic algorithms (GA) and particle swarm optimization (PSO) are inspired by nature. Simon [15] presented in this algorithm, immigration of species which are some parts of biogeography science are used for solving the optimization issues. In biogeography-based optimization, every habitat is identified as an individual member and has its own habitat suitability index (HIS). A biogeography-based optimization algorithm is like a genetic algorithm, with its higher habitat suitability index in one habitat showing a good habitat. On the other hand, habitats with a low habitat suitability index make themselves look like habitats with high habitat suitability index through taking their properties. There are two migration factors in a migration model, namely, emigration and immigration. These two kinds of migrations have their own rate and are identified with the names of input rate ( $\lambda$ ) and output rate ( $\mu$ ). In a biogeography-

based optimization method, two migration factors in a migration operator and a mutation operator are used for predicting the habitat, which maximizes the habitat suitability index. A good habitat has a higher emigration rate and a lower immigration rate as Fig. 2 [15].

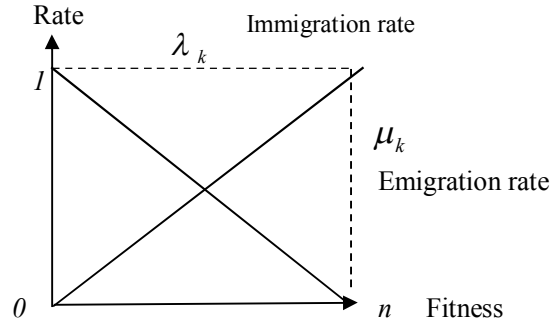


Fig. 2. A Curve of Linear Migration of BBO [15]

Based on Fig.2[15],  $n$  which is the best fitness and maximum amount of emigration and immigration is considered  $I$  for an emigration rate and immigration rate. Savsani et al. [16] presented an emigration rate and immigration rate equal the size of  $n$  and is calculated, using formulas (1), (2), (3).

$$y_k(a) = \delta y_k(a) + (1 - \delta)y_j(a) \tag{1}$$

$$Prob(immigration\ to\ y_k) = \lambda_k \tag{2}$$

$$Prob(emigration\ to\ y_k) = \frac{\mu_k}{\sum_{j=1}^n \mu_j} \tag{3}$$

Ma and Simon [17] presented In these formulas  $a$  represents an index of decision making variable and  $\delta$  represents a real number which ranges from  $0$  to  $1$  and can occur randomly, definitely, and in accordance with a relative fitness  $y_k$  and  $y_j$ . Emigration probability is done through roulette-wheel selection as is shown in Fig.3 [16]. Conducted studies on biogeography-based optimization showed that this method is suitable for most of the issues.

BBO Algorithm

- 1: **for** each candidate solution  $y_k$  **do**
- 2: **for** each candidate solution decision variable index  $a$  **do**
- 3: **use**  $\lambda_k$  to probabilistically decide whether to immigrate to  $y_k$  (see Eq.(2))
- 4: **if** immigrating **then**

```

5:      use {μ} to probabilistically select
the emigrating candidate solution yj(see Eq.(1))
6:      yk(a) = δyk(a) + (1 - δ)yj(a)
7:      end if
8:  end for
9:  probabilistically decide whether to mutate yk
10: end for
    
```

Fig. 3. BBO Algorithm [16]

### 3.The Proposed Method

This part proposes a method for text summarization using the biogeography-based optimization method. In the proposed method, first pre-processing should be done on the input text. This function is done as follows:

1. Reading documents to ensure coherence among them.
2. Omitting additional characters such as {}, [], ...
3. Omitting unessential words such as at, in, on, and of, ...
4. Finding the roots of the words and verbs used in the main text.
5. Separating words from each other and omitting repeated words.
6. Identifying the end of the sentences and separating the sentences from each other.
7. Calculating the number of words and sentences used in the main text.

Then the sentences extracted from the pre-processing step should be weighed, using a weighing method of TFIDF which is equivalent to term-frequency inverse document-frequency. Garsia and Ledeneva [4] presented, In the weighing method in Eq. (4) and Eq. (5),  $freq_{ij}$  represents the repetition number of  $i$  word in  $j$  sentence,  $maxfreq_{ij}$  represents a frequency maximum of  $l$  word in  $j$  sentence,  $maxfreq_{ij}$  represents the number maximum of  $i$  word repetition in  $j$  sentence,  $N$  and  $n_i$  represent the number of sentences in an input text and the number of sentences, respectively. Then a weight of each word in sentence  $w_{ij}$  is calculated, using Eq. (6).

$$tf_{ij} = \frac{freq_{ij}}{\max_i freq_{ij}} \quad (4)$$

$$idf_{ij} = \log \frac{N}{n_i} \quad (5)$$

$$w_{ij} = tf_{ij} * idf_{ij} \quad (6)$$

After the sentences are weighed, a similarity matrix should be weighed, using Eq. (7). In a similarity matrix, sentences are compared based

on their keywords and important words.

$$sim(s_i, q) = \frac{\sum_{i=1}^t w_{ij} * w_{iq}}{\sqrt{\sum_{i=1}^t w_{ij}^2 * \sum_{i=1}^t w_{iq}^2}} \quad (7)$$

Now, important and summarized sentences should be extracted from the main text, using a biogeography-based optimization method. The proposed method consists of a set of initial parameters based on Fig. 4 which should be valued. In this method, the number of habitats and the number of operation repetitions are considered to be 50 and 100, respectively. Then, the overall number of sentences, the number of summarized sentences, and a similarity matrix are considered as input parameters of the proposed method. Following the implementation of the method, sentences with high scores are selected and displayed as a summarized text.

the cost of each habitat in based optimization algorithm calculated based on its sentences. Following the calculation of the similarities of sentences, a similarity factor is obtained from Eq. (8) and Eq. (9).

$$TRF_s = \frac{TR}{\max_{summary} (TR)} \quad (8)$$

$$TR_s = \frac{\sum_{sj \in summary} sim(s_j, q)}{s} \quad (9)$$

In this case,  $TR$  represents a similarity mean in  $s$  summarization.  $TRF$  represents similarity factor, which is calculated based on  $TR$ . Also, a maximum is calculated from all possible  $S$  summaries.  $TRF$  shows the extent to which a document summarization is similar to a document itself. In the summaries in which the sentences are related to the document,  $TRF$  is approximately 1, but in the summaries in which the sentences are not related to the document,  $TRF$  is approximately zero.  $Rs$  represents a readability factor of a summary with an amount of  $s$ . Therefore, a readability factor  $RF$  can be obtained, using Eq. (10) and Eq. (11) which introduced by Qazvinian et al. [18].

- 
- Step 1: Initial amount of biogeography-based algorithm parameters
  - Step 2: Random assignment of sentences in habitats
  - Step 3: Calculation of immigration rate and emigration rate from Eq. (1) and Eq. (2)
  - Step 4: Investigating habitats based on cost function

Step 5: Updating the situation of habitats based on the best conditions

Step 6: If a final condition is maintained (finding the best sentences), a method would come to an end and important sentences would be shown. Otherwise, that person should return to step 4.

**Fig. 4. Proposed method with the use of BBO**

$$RF_s = \frac{R_s}{\max_{\forall i} R_i} \quad (10)$$

$$R_s = \sum_{0 \leq i \leq s} W(s_i, s_{i+1}) \quad (11)$$

An expense function can be calculated using Eq. (12) and Eq. (13).

$$C_s = \frac{\sum_{\forall s_i, s_j \in \text{summarysubgraph}} W(s_i, s_j)}{N_s} \quad (12)$$

$$CF_s = \frac{\log(C \times 9 + 1)}{\log(M \times 9 + 1)} \quad (13)$$

In Eq. (12) and Eq. (13),  $CF_s$  calculates the amount of sentence cohesion. A cohesion factor will make the existent sentences in a summary to show similar information. A readability factor of sentences shows that the first, second, and final summarized sentences are related to each other with high similarity. Also, in a cohesion factor of sentences,  $C$  represents a similarity mean among sentences in  $S$  summary,  $M$  represents the highest amount of sentence similarity, and  $N_s$  shows an overall amount of summarization.

**4. Simulation and Results of Investigation**

In an extractive method, the results of the study are investigated against four criteria. These four criteria are as follows:

1. Investigation based on a text quality  
This kind of investigation is done by a human being. A summarized text which is based on some defined indicators is given a score.
2. Investigation based on a Selection  
This kind of investigation is done based on sentence correspondence.
3. Investigation based on a Content  
This kind of investigation is done based on word correspondence.
4. Investigation based on a Duty

In this kind of investigation, a quality of a summarized text is measured based on people’s demands. In the proposed method, simulation is done based on the second method (Sentence or response). In this method, an extracted summary is compared with an ideal summary and an extracted summary is investigated against criteria like precision, recall, and score which can be seen respectively in Eq. (14), (15) and (16), which introduced by Abuobieda, Salim et al. [19].

$$\text{precision} = \frac{\text{RelevantSentences} \cap \text{RetrievedSentences}}{\text{RetrievedSentences}} \quad (14)$$

$$\text{Recall} = \frac{\text{RelevantSentences} \cap \text{RetrievedSentences}}{\text{RelevantSentences}} \quad (15)$$

$$F\text{-Score} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (16)$$

In Eq. (14), Precision is an intersection of extracted summarized sentences and an ideal summary of sentences divided by all the extracted sentences. In Eq. (15), Recall is an intersection of relevant sentences and retrieved sentences divided by all the relevant documents. In Eq. (16), Score is a statistical criterion which is a combination of precision and recall criteria and shows the score of final selected sentences in a summary of a text since calculation of these criteria is a little hard and time consuming, ROUGE automatic software is used for investigating these criteria. A language of this software is Perl, which includes different packages for the purpose of investigating a summarized text. This software studies the mentioned criteria and shows the results. Nine Doc standard documents of 2002 were used for the purpose of investigating the proposed method. The test results for documents with 400, 200, and 100 words were compared with GA [11], PSO [12], BFOA [13] algorithms. The results are displayed in Table 1, Table 2 and Table 3.

**4-1. Experiment**

In this experiment, Table 1 and the graph in Fig 5 show the comparisons between a genetic algorithm method and a biogeography-based optimization method on similar samples of 100 word documents. As it can be seen, precision, recall, and score in the proposed method show more improvement than those in the genetic algorithm method.

Tab. 1. Comparison of 100- word documents

DOC	BBO			GA		
	Precision	Recall	F-score	Precision	Recall	F-score
d061j	0.19048	0.56604	0.28504	0.18619	0.58491	0.28247
d065j	0.8883	0.62000	0.15540	0.10387	0.59000	0.17664
d067f	0.14014	0.53774	0.20956	0.13971	0.53774	0.22180
d070f	0.10383	0.76000	0.18270	0.09504	0.69000	0.16707
d073b	0.19291	0.50000	0.27841	0.20745	0.39796	0.27273
d075b	0.11154	0.53211	0.18442	0.09760	0.55963	0.16621
d079a	0.09792	0.60550	0.16858	0.12990	0.57798	0.21213
d085d	0.10508	0.58252	0.17804	0.10846	0.57282	0.18239
d105g	0.10435	0.60000	0.17778	0.08604	0.61000	0.15081

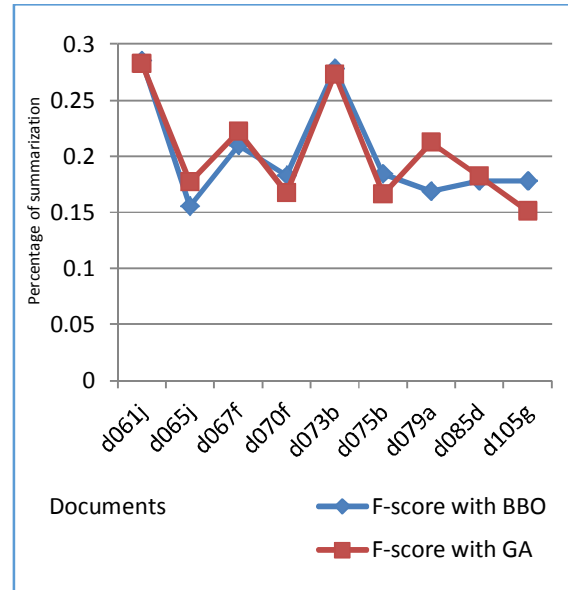


Fig. 5. Comparison Graph of BBO and GA Based on F-score in a Model of 100 – Words Documents

#### 4-2. Experiment 2

This experiment is conducted to compare the proposed algorithm with genetic algorithm and biogeography-based optimization method on similar samples of 200 word documents.

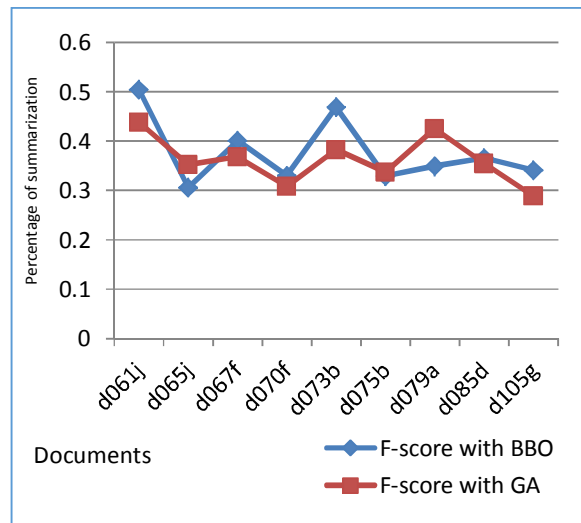
Table 2 and the graph in Fig. 6 show the results obtained for this experiment. As it is shown, precision, recall, and score in the proposed method performs better than the others algorithms.

#### 4-3. Experiment 3

In this experiment Table 3 and the graph in Fig. 7 show the comparisons between a genetic algorithm method and a biogeography-based optimization method on similar samples of 400 word documents. As it can be seen, precision, recall, and score criteria in the proposed method show relatively more improvement than those in the genetic algorithm method.

**Tab. 2. Comparisons of 200 -word documents**

DOC	BBO			GA		
	Precision	Recall	F-score	Precision	Recall	F-score
d061j	0.41587	0.63902	0.50384	0.35435	0.57561	0.43866
d065j	0.20201	0.64239	0.30619	0.24472	0.62332	0.35146
d067f	0.39224	0.64000	0.40126	0.27451	0.56000	0.36842
d070f	0.22131	0.65060	0.33027	0.20661	0.60241	0.30769
d073b	0.43307	0.51163	0.46908	0.40957	0.35814	0.38213
d075b	0.23269	0.56019	0.32880	0.22720	0.65741	0.33769
d079a	0.22849	0.74757	0.35000	0.30309	0.71359	0.42547
d085d	0.24694	0.70500	0.36576	0.24265	0.66000	0.35484
d105g	0.23130	0.64878	0.34102	0.18618	0.64390	0.28884



**Fig. 6. Comparison Graph of BBO and GA Based on F-score in a Model of 200- words Documents**

**4-4. Experiment 4**

In this experiment, Table 4 and the graph in Fig. 8 show the comparisons between genetic algorithm method Particle Swarm Optimization, Bacterial Foraging Optimization Algorithm and

biogeography-based optimization method on similar samples of 400 word documents. As it can be seen, F-score criteria in the proposed method show the relative improvement compared to those in other methods.

**Tab. 3. Comparisons of 400 -word documents**

DOC	BBO			GA		
	Precision	Recall	F-score	Precision	Recall	F-score
d061j	0.64127	0.51007	0.55571	0.57.57	0.46117	0.51007
d065j	0.32951	0.59432	0.42396	0.36972	0.52264	0.43979
d067f	0.32951	0.59432	0.42396	0.42647	0.41527	0.42080
d070f	0.37568	0.67402	0.48245	0.39532	0.70343	0.50618
d073b	0.59055	0.36232	0.44910	0.62234	0.28261	0.38871

<i>d075b</i>	0.45192	0.55035	0.49630	0.43040	0.62998	0.51141
<i>d079a</i>	0.41242	0.65258	0.50545	0.50928	0.57981	0.54226
<i>d085d</i>	0.38704	0.55388	0.45567	0.36581	0.49875	0.42206
<i>d105g</i>	0.43130	0.66488	0.52320	0.33427	0.63539	0.43807

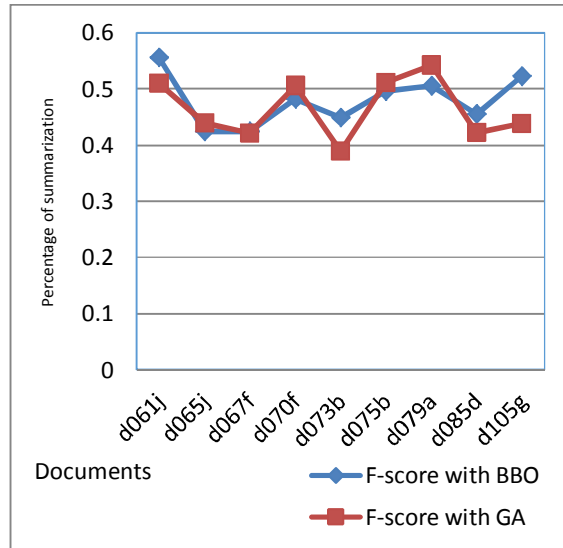


Fig. 7. Comparison of BBO and GA based on F-score in a model of 400 –words Documents

### 5. Conclusion

In this article, different methods of text extractive summarization along with their advantages and disadvantages were discussed. Then, a biogeography-based optimization method was proposed (Fig. 4) and used in a text extractive summarization. Next, the ways in which the short sentences are selected and scored were presented. Finally, a proposed

method was tested with 9 sets of doc standard documents of 2002. These documents were those of 100 (Table. 1 and Fig. 5), 200 (Table .2 and Fig. 6) and 400 (Table. 3 and Fig. 7) words. Then, the test results were analyzed using ROUGE software. The results of four tests showed the superiority of the proposed method compared to other methods in terms of efficiency (Table. 4 and Fig.8).

Tab. 4. Comparisons of 400- word documents

	BBO	GA	PSO	BFOA
<i>DOC</i>	F-score	F-score	F-score	F-score
<i>d061j</i>	0.55571	0.51007	0.42869	0.43543
<i>d065j</i>	0.42396	0.43979	0.37992	0.37997
<i>d067f</i>	0.42396	0.42080	0.44637	0.44126
<i>d070f</i>	0.48245	0.50618	0.40616	0.41765
<i>d073b</i>	0.44910	0.38871	0.58658	0.54654
<i>d075b</i>	0.49630	0.51141	0.48764	0.48965
<i>d079a</i>	0.50545	0.54226	0.41177	0.41186
<i>d085d</i>	0.45567	0.42206	0.39094	0.41653
<i>d105g</i>	0.52320	0.43807	0.39517	0.39121



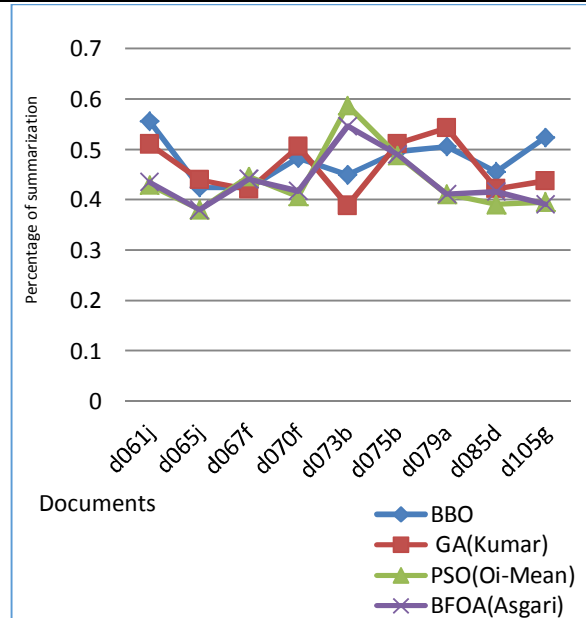


Fig. 8. Comparison Graph of BBO , PSO, BFOA and GA based on F-score in a model of 400-word Doc

References

[1] Das, D., and Martins, A., "A Survey on Automatic Text Summarization", Language Technologies Institute Carnegie Mellon University, No.( 2007).

[2] Hovy, E., "Text summarization",chapter The Oxford Handbook of Computational Linguistics, (2005).

[3] Gupta, V., "A Survey of Text Summarization Extractive Techniques", Journal of Emerging Technologies in web Intelligence, vol. 2, ( 2010).

[4] Ledeneva, Y., and Gelbukh, A., and Hernández, R., "Terms derived from frequent sequences for extractive text summarization", in Proceedings of the 9th international conference on Computational linguistics and intelligent text processing, (2008), pp. 593- 604.

[5] Garcia-Hernandez, R., and Ledeneva, Y., "Word Sequence Models for Single Text Summarization", in Advances in Computer-Human Interactions, (2009), pp. 44- 48.

[6] Zhang, P., and Li,C., "Automatic text summarization based on sentences clustering and extraction", in Computer Science and Information Technology 2nd IEEE International Conference on, (2009), pp. 167-170.

[7] Chuang, W., and Yang, J., "Text Summarization by Sentence Segment Extraction Using Machine Learning Algorithms", in Knowledge Discovery and Data Mining Current Issues and New Applications.vol.1805,( 2000), pp. 454-457.

[8] Song, W., and Cheon Choi, L., and Cheol Park, S., and Feng Ding, X., "Fuzzy evolutionary optimization modeling and its applications to unsupervised categorization and extractive summarization", Expert Systems with Applications,vol.38, (2011), pp. 9112-9121.

[9] Suanmali, L., and Salim, N., and Salem Binwahlan, M., "Fuzzy Logic Based Method for Improving Text Summarization", International Journal of Computer Science and Information Security (IJCSIS), vol. 2, (2009).

[10] Fattah, A., and Ren, F., "GA, MR, FFNN, PNN and GMM based models for automatic text summarization", Computer Speech & Language, vol. 23, (2009), pp. 126-144.

- [11] Kumar Meena, Y., Gopalani, D., "Evolutionary Algorithm for Extr Text Summarization ", in International Conference on Intelligent Compting, (2015).
- [12] Oi-Mean, F., and Oxley, A., "A hybrid PSO model in Extractive Text Summarizer", in Computers & Informatics (ISCI), 2011 IEEE Symposium on, (2011), pp.130-134.
- [13] Asgari, H., and Masoumi, B., "Provide a method to improve the performance of text summarization using bacterial foraging optimization algorithm", the seventh iran dataminig conference,( 2013).
- [14] Uddin, N, and Khan S., "A study on text summarization techniques and implement few of them for Bangla language", in Computer and information technology, (2007), pp. 1-4.
- [15] Simon,D, "A Probabilistic Analysis of a Simplified Biogeography-Based Optimization Algorithm. "Evolutionary Computation", vol. 19, No. 2. (2011), pp. 167-188.
- [16] Savsani, P , Jhala, R.L. and Savsani,V., "Effect of Hybridizing Biogeography-Based Optimization technique with Artificial Immune Algorithm and Antcolony Optimization", in Applied Soft Computing 21 International Journal, (2014), pp. 542-553.
- [17] Ma, H., Simon D., MinruFei ,Xinzhan Shu,"Hybrid Biogeography-Based Evolutionary algorithm", in Engineering Application Of Artificial International 30 International Journal (2014), pp. 213-224.
- [18] Qazvinian V., Hassanabadi, L.Sh., Halavati, R. "Summarising text with a genetic algorithm based sentence extraction", International Journal of Knowledge Management Studies, Vol. 2, No. 4, (2008).
- [19] Abuobieda, A., Salim, N., and Kumar, Y., and Osman, A., "An Improved Evolutionary Algorithm for Extractive Text Summarization", in Intelligent Information and Database Systems. vol. 7803, (2013), pp. 78-89.

Follow This Article at The Following Site

MirShojaee H, Masoumi B, Zeinali E. Biogeography-Based Optimization Algorithm for Automatic Extractive Text Summarization . IJIEPR. 2017; 28 (1) :75-84  
 URL: <http://ijiepr.iust.ac.ir/article-1-722-en.html>  
 DOI: 10.22068/ijiepr.28.1.75

