Scientific Document Summarization using Citation Context and Multi-objective Optimization

Naveen Saini*, Sushil Kumar, Sriparna Saha, And Pushpak Bhattacharyya
Indian Institute Of Technology Patna, Bihar, India-801106

ABSTRACT

- Rapid growth in the field of scientific literature put challenges for the researchers to gain up-to date knowledge of the current advancements. The summarization of scientific article address this challenge.
- Abstract are not enough as they covers only a broad idea of the article.
- Citation contexts are also not enough to summarize a SD as they may lack the context to support the main content of the RP and sometimes, may be inaccurate or misunderstood by the authors
- We propose a system for scientific document summarization (SDS) having two components: identifying the relevant sentences in the article using citation context; generation of the summary by posing SDS as a binary optimization problem.
- For the purpose of optimization, a meta-heuristic evolutionary algorithm (binary differential evolution) is utilized and various aspects measuring the relevance of sentences are simultaneously optimized using the concept of multi-objective optimization.
- Inspired by the popularity of graph-based algorithms like LexRank which is popularly used in solving summarization problems of different real-life applications, its impact is studied in fusion with our optimization framework.
- An ablation study is also performed to identify the most contributing aspects for the summary generation.

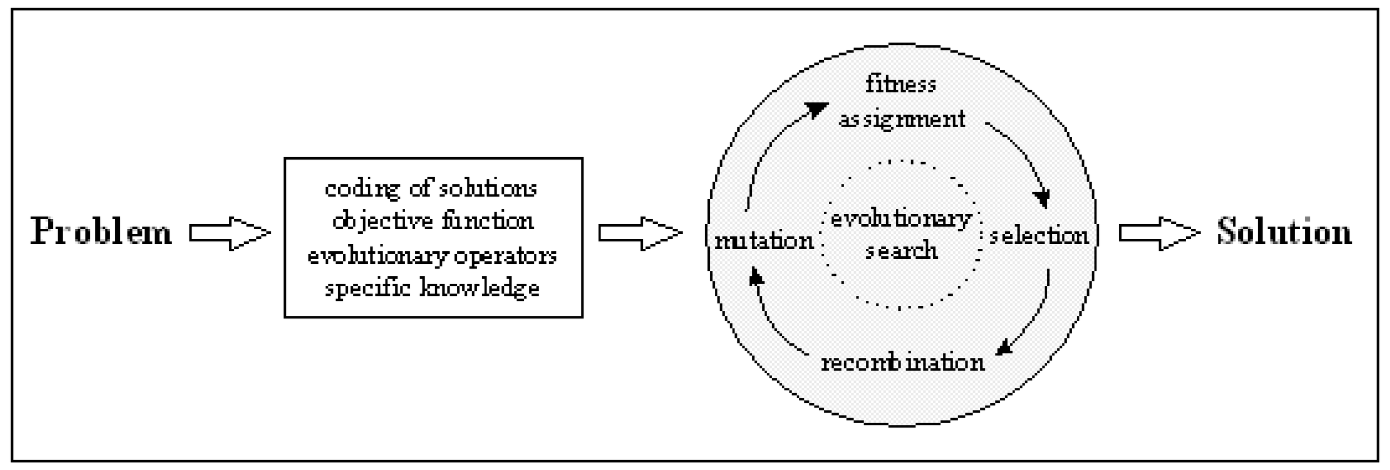
BACKGROUND

Multi-objective optimization

- Optimization of more than one objectives simultaneously
- Provide a set of Pareto optimal solutions in a single run
- Users can choose any one based on his/her choice

Evolutionary algorithms

- Meta-heuristic optimization algorithms to find an optimal solution
- Inspired by biological phenomenon (crossover, mutation and selection) in the natures
- Differential evolution (DE) is one such algorithm.
- It starts from a set of solutions (called as population)
- Each solution is associated with some objective(s) values to measure their quality
- These solutions evolves over the iteration to generate new population.
- Only those solutions proceed to the next generation which are good in terms of "survival of the fittest" principle.



Non-Dominated Sorting

Used to sort the solutions based on multiple objectives. It partition the population into *L* different non-dominated fronts. Top best solutions are selected (considering rank-wise fronts) to proceed for the next generation. In case of a tie, the solutions within a front are further sorted based on crowding distance and the solutions having higher crowding distance, are added as part of next generation.

ALGORITHM

Instead of using all sentences present in the article, we have considered only those sentences which are relevant to the citation contexts. Therefore, as the first step, the relevant sentences in the reference paper (RP) are identified in the semantic space and then, the following algorithm is applied

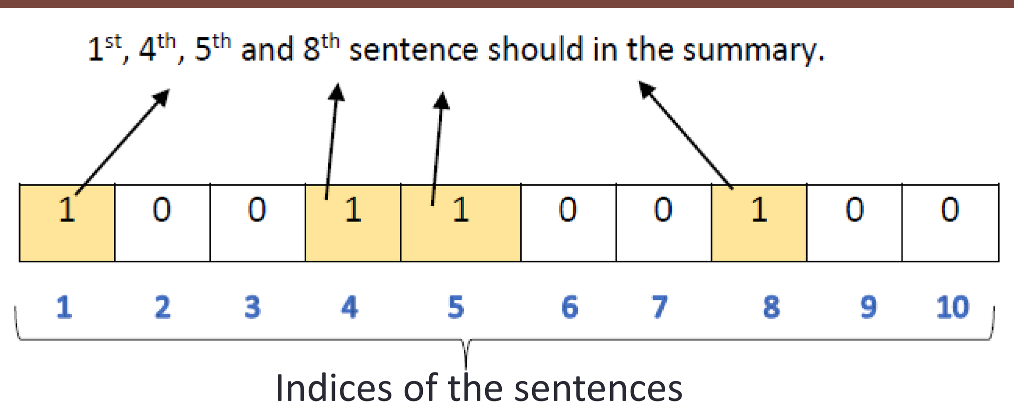
Algorithm 1 Procedure of MOOTweetSumm+

1: $\mathbb{P} \leftarrow$ Initialize Population $< X^1, X^2, X^3, \ldots, X^{|\mathbb{P}|} >$
2: For each solution $\mathcal{X}$, evaluate objective functional values
3: $CGen$=0 ▷ Current generation number
4: Repeat step-5 to 9 until $CGen < MaxGen$
5: $\mathbb{P}'$=[] ▷ Population to store new solutions
6: For each solution $\mathcal{X} \in \mathbb{P}$, generate new solution

(a) Randomly select three solutions $r1, r2$ and $r3$ from $\mathbb{P}$ to form a mating pool
(b) $Prob(\mathcal{X}) \leftarrow$ Perform probability estimation operator using selected random solutions and $\mathcal{X}$
(c) $Y' \leftarrow$ Convert $Prob(\mathcal{X})$ into a binary solution
(d) $Y'' \leftarrow$ Perform crossover between Y' and $\mathcal{X}$
(e) Evaluate objective functions for Y''
(f) Add Y'' into $\mathbb{P}'$

7: Merge Old population ($\mathbb{P}$) and new population ($\mathbb{P}'$)
8: $\mathbb{P} \leftarrow$ Select the best $|\mathbb{P}|$ solutions based on their objective functional values using non-dominating sorting and crowding distance operator
9: $CGen \leftarrow CGen$+1
10: **return** the best summary

ENCODING OF SOLUTION



OBJECTIVE FUNCTIONS

- F1: Sentence position in the article (F1) (↑)
- F2: Maximum similarity with the Title (F2)
  - Representation of sentences using fast-text word2vec model followed by cosine similarity (↑)
  - Reciprocal of Word mover distance utilizing GoogleNews word2vec model (↑)
- F3: Maximum overlap with the top-scoring sentences provided by LexRank algorithm (F3) (↑)

Thus, objective is to maximize these objective functions represented as

$$\max\{F1_{avg}, F2_{avg}, F3_{avg}\}$$

RESULTS

TABLE II
COMPARISON OF OUR PROPOSED APPROACHES WITH TOP SUBMITTED SYSTEMS OF CL-SCISUMM 2016 SHARED TASK.

Type of Summary→	Human	Community	Abstract
Method	ROUGE-SU4		
Ours ($F1$, $F2_w$)	**0.24**	**0.41**	**0.45**
Ours ($F1$, $F2_v$)	**0.25**	**0.39**	**0.43**
Ours ($F2_w$, $F3$)	**0.21**	**0.36**	**0.43**
Ours ($F2_v$, $F3$)	**0.20**	**0.33**	**0.37**
sys8$PARA_7	0.14	0.13	0.42
sys8$PARA_1	0.11	0.13	0.25
sys3$LMKL1_CCS1	0.12	0.09	0.18
sys3$LMEQUAL_CCS2	0.12	0.10	0.21

TABLE III
COMPARISON OF OUR PROPOSED APPROACHES WITH TOP SUBMITTED SYSTEMS OF CL-SCISUMM 2017 SHARED TASK.

Type of Summary→	Human	Community	Abstract
Method	ROUGE-SU4		
Ours ($F1$, $F2_w$)	**0.26**	**0.38**	**0.27**
Ours ($F1$, $F2_v$)	**0.24**	**0.37**	**0.27**
Ours ($F2_w$, $F3$)	**0.25**	**0.35**	**0.22**
Ours ($F2_v$, $F3$)	**0.25**	**0.35**	**0.21**
CIST Run 3 [9]	0.17	0.16	0.17
UniMA Run 4,5,6 [20]	0.16	0.17	0.19
UniMA Run 7,8,9 [20]	0.16	0.16	0.18
Mean Score (all systems) [5]	0.14	0.14	0.15

TABLE IV
COMPARISONS BETWEEN BEST ROUGE SCORES ATTAINED BY PROPOSED METHODS WITH THE EXISTING METHODS ON CL-SCISUMM 2016 DATASET CONSIDERING HUMAN SUMMARY.

Human Summary		
Evaluation Measure→	ROUGE-2	ROUGE-SU4
Method↓	F1-scores	
Ours ($F1$, $F2_w$)	**0.23**	**0.24**
Ours ($F1$, $F2_v$)	**0.23**	**0.25**
Ours ($F2_w$, $F3$)	0.19	**0.21**
Ours ($F2_v$, $F3$)	0.19	**0.29**
LexRank[7]	0.12	0.11
CLexRank [8]	0.06	0.09
SumBasic [10]	0.09	0.12
CIST [9]	0.22	0.14
LMKL [11]	0.19	0.11
LMeq [11]	0.19	0.12
SUMMA [34]	0.13	0.09

TABLE V
COMPARISONS BETWEEN BEST ROUGE-2 SCORES (PRECISION (P), RECA (R) AND F1-SCORE (F1)) ATTAINED BY THE PROPOSED METHOD WITH T EXISTING METHODS ON CL-SCISUMM 2017 DATASET CONSIDERING ABSTRACT SUMMARY.

Abstract Summary			
Evaluation Measure→	ROUGE-2		
Method↓	Avg. P	Avg. R	Avg. F1
Ours ($F1$, $F2_w$)	**0.26**	**0.41**	**0.26**
Ours ($F1$, $F2_v$)	**0.26**	**0.40**	**0.26**
Ours ($F2_w$, $F3$)	0.20	**0.33**	0.20
Ours ($F2_v$, $F3$)	0.19	**0.34**	0.19
Word2vec [19]	0.21	0.27	0.24
tfidf-1:1 [19]	0.21	0.24	0.22
tfidf-1:3 [19]	0.20	0.24	0.21
tfidf-1:2 [19]	0.19	0.24	0.21
Lambdamart [33]	0.19	0.23	0.21
Filter [19]	0.18	0.23	0.20
Lambda [33]	0.26	0.02	0.20

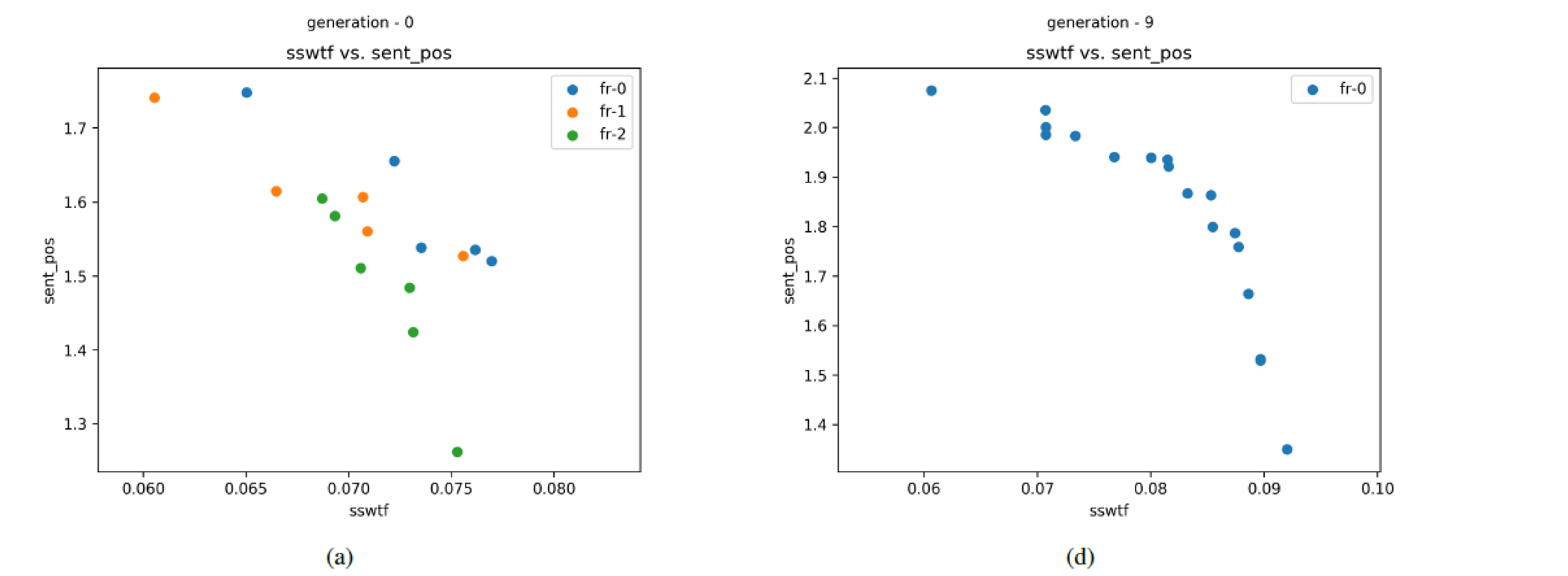


Fig. 3. Pareto Fronts generated by our proposed *MOOSciSumm* approach at the end of $0th$ and $9th$ generations after optimizing two objective functions (i) $SentPos$; (ii) $MaxSimWithTitle$ utilizing FastText vector. Here, sent_pos and $sswtf$ refer to the mentioned objective function in (a) and (b), respectively. Here, fr-0 in the legend indicates the solutions of rank-1 and so on.

FUTURE WORK

Our developed framework is generalized in nature and can be adopted for developing any other summarization systems, including single document summarization, multi-document summarization, microblog summarization, among others. We will be working for developing these systems in future. We will also like to extend this work to write a related wok section on a given topic in an automatic manner.

References:

1. Yasunaga et al., "Scisummnet: A large annotated corpus and content-impact models for scientific paper summarization with citation networks," in Proceedings of the AAAI Conference on Artificial Intelligence, vol. 33, 2019, pp. 7386–7393.
2. Jaidka et al., "The cl-scisumm shared task 2017: Results and key insights," 11 2017.
3. Cohan and N. Goharian, "Scientific document summarization via citation contextualization and scientific discourse," International Journal on Digital Libraries, vol. 19, no. 2-3, pp. 287–303, 2018.
4. Jaidka et al., "Insights from cl-scisumm 2016: the faceted scientific document summarization shared task," International Journal on Digital Libraries, vol. 19, no. 2-3, pp. 163–171, 2018
5. Saini et al., "Extractive single document summarization using binary differential evolution: Optimization of different sentence quality measures," PloS one, vol. 14, no. 11, 2019.