Analysis of diversity in an evolutionary algorithm for the graph burning problem

Gabriela Alva Zavala

Advisor: José Alejandro Cornejo Acosta

Instituto Tecnológico Superior de Purísima del Rincón, Guanajuato, México.

Introduction

The *graph burning problem* is an NP-hard optimization problem that helps quantify how vulnerable a graph is to contagion (it simulates an infection spread) and aims to find which vertices burn the whole graph with the minimum possible sequence length.

On March 11th, 2020, COVID-19 was declared a pandemic by the World Health Organization (WHO) general director Dr. Tedros Adhanom Ghebreyesus at a global press conference. At such moments as a pandemic, it is essential to make strategic decisions to prevent the infection from spreading faster. The graph burning problem measures how fast an infection spreads. Therefore, we know how vulnerable a population is to an infection, not just COVID, but also other types of infections such as measles, chicken pox, pneumonia, tuberculosis, etc.

Charles Darwin, the father of evolutionary theory, said in his theory that individuals need to adapt to their environment to survive. In nature, at the moment of reproduction, the features that make an individual more suited to competition are preserved, and weaker elements are eliminated. The features are genes and a set of genes form chromosomes, therefore the fittest individuals that survive transmit their fittest genes to the next generation [1].

Evolutionary Algorithms (EAs) are iterative algorithms that make the individuals of a population evolve over a number of generations [2]. This type of algorithm is an efficient problem-solving method applied for global optimization problems and can be used in real-world applications of high complexity (robotics, operation research, decision making, bioinformatics, machine learning, data mining), where EAs use simulated evolution to find a solution to the complex real-world problems [1]. EAs have been proven effective in solving the graph burning problem. Nevertheless, there is an issue with the loss of diversity in the solutions. This study aims to analyze the diversity of an evolutionary algorithm and proposes a distance metric to measure diversity. Then, some experimentation is performed with instances of the literature.

State-of-the-art and related work

Table 1: State-of-the-art summary of the graph burning problem

Author	**Technique**	**Journal**	**Reference**
Bonato & Kamali. 2019	Approximation Algorithms	arXiv	[3]
Rezai Farokh et al. 2020	Heuristics in connected graphs	arXiv	[4]
Kumar Gautam et al. 2021	Heuristics approaches: BBGH, ICCH, and CBRH	arXiv	[5]
García Díaz et al.2022	A farthest-first traversal-based approximation algorithm BFF	IEEE Access	[6]
García Díaz et al. 2022	Mathematical Formulations	Mathematics	[7]
Nazeri et al. 2023	A genetic algorithm CBAG	Applied Soft Computing	[8]
García Díaz et al. 2024	Deterministic Greedy Heuristics	arXiv	[9]
Pereira et al. 2024	Row Generation Algorithm	arXiv	[10]

Objectives

General objective: To analyze the diversity of an evolutionary algorithm and propose a distance metric to diversity measure.

Specific objectives:

- Study the literature concerning the graph burning problem.
- Propose a distance metric algorithm to study diversity.
- Perform experimentation with instances from the literature.

Methodology

To perform a study of the diversity of the evolutionary algorithm "CBAG" a literature review of the graph burning problem was done. The CBAG evolutionary algorithm was downloaded from a GitHub repository and executed. The algorithm required the desired burning number (square root of the number of nodes of the graph [11]), the algorithm did not return the shortest burning sequence or the optimal burning number needed for the analysis. Therefore, some changes were made to the code. Additionally, the CBAG implementation was modified to receive as parameters the number of generations and the population size. Subsequently, the Levenshtein Distance method (metric distance), and the average and standard deviation calculation of the distances were added to the code. The optimal burning sequence from a specific generation was obtained to perform experiments with instances of the literature.

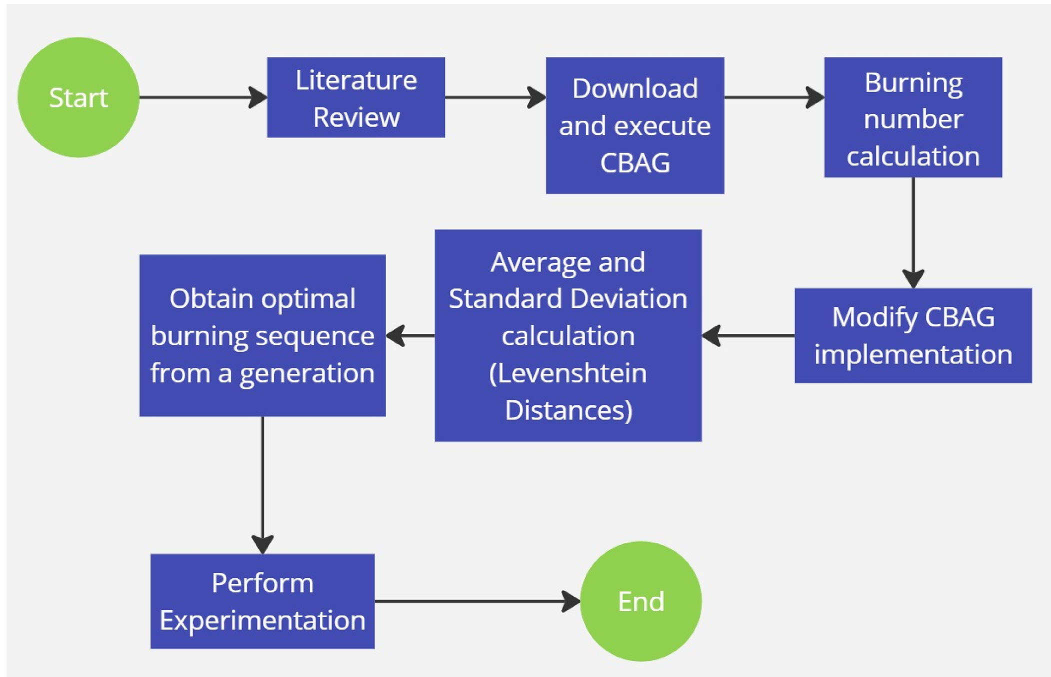

Figure 1: Methodology

Table 2: Dataset used

Instance name	$\lvert V\rvert$	$\lvert E\rvert$
tvshow	3892	17262
politician	5908	41729
government	7057	89455
crocodile	11631	180020

Levenshtein Distance (metric distance)

The *levenshtein distance* algorithm calculates the minimum number of operations (replace, insert, delete) that need to be done to convert a vector of numbers into another vector of numbers. This method has a time complexity of $O(mn)$.

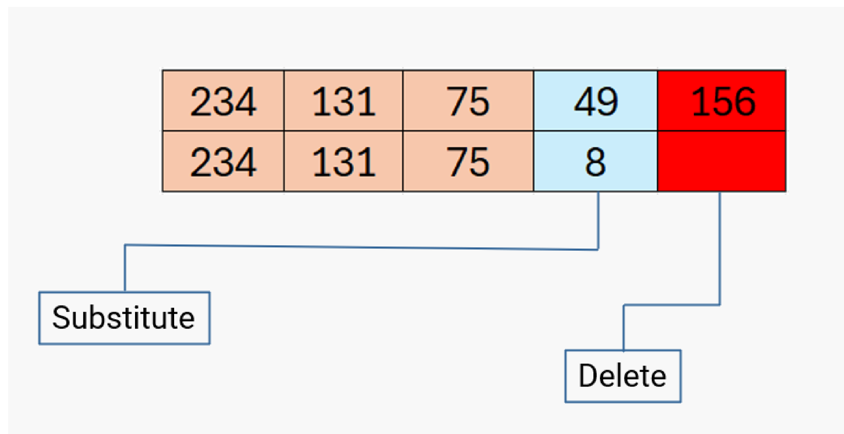

Figure 2: Levenshtein Distance example

Results

A) Random search results

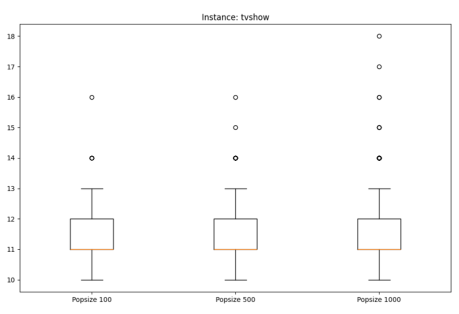

Figure 3: Burning numbers – instance: tvshow – Generation 0 – Population: 100, 500, 1000.

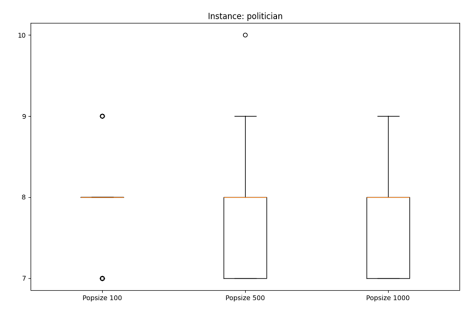

Figure 4: Burning numbers – instance: politician – Generation 0 - Population: 100, 500, 1000.

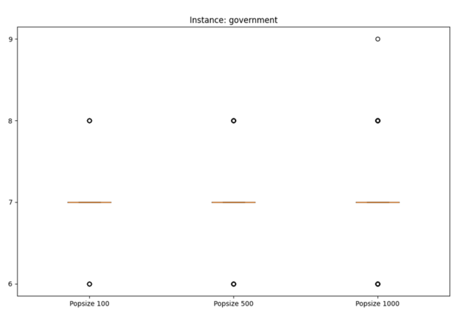

Figure 5: Burning numbers – instance: government – Generation 0 - Population: 100, 500, 1000.

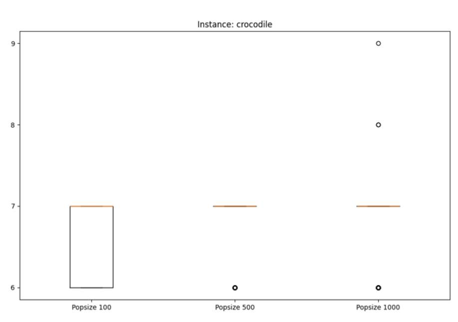

Figure 6: Burning numbers – instance: tvshow – Generation 0 - Population: 100, 500, 1000.

B) Distances between solutions

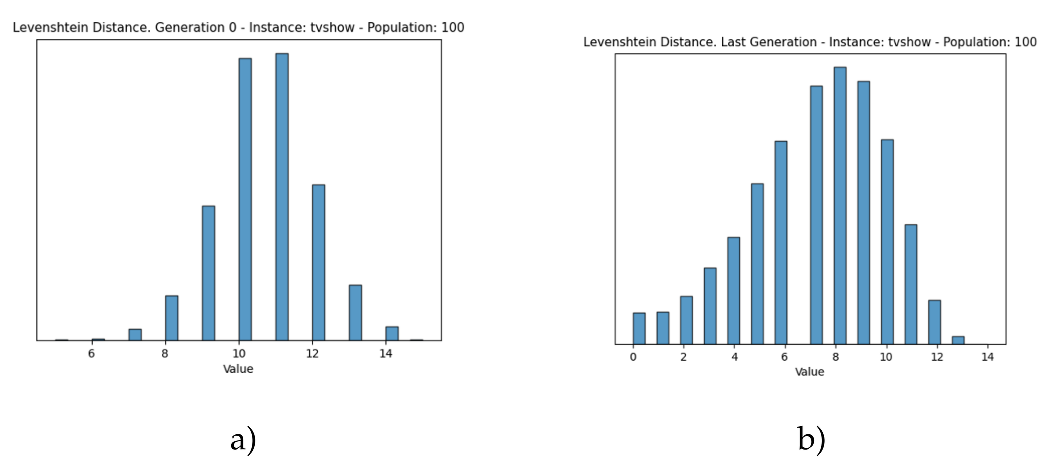

Figure 7: Levenshtein Distance – instance: tvshow – Population: 100. a) Generation 0 and b) Last Generation.

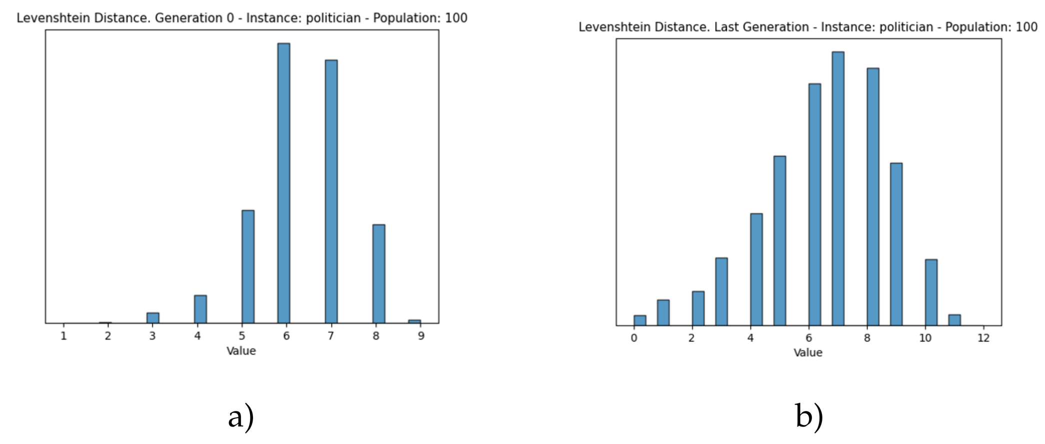

Figure 8: Levenshtein Distance – instance: politician – Population: 100. a) Generation 0 and b) Last Generation.

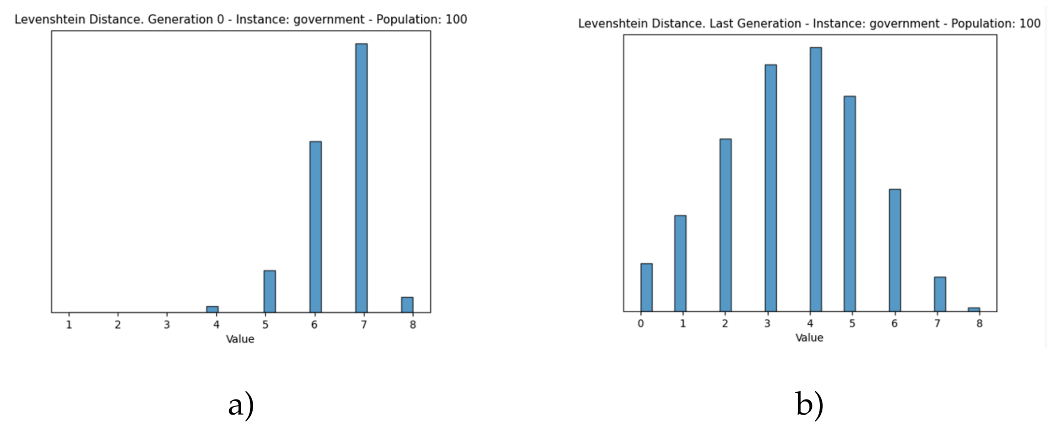

Figure 9: Levenshtein Distance – instance: government – Population: 100. a) Generation 0 and b) Last Generation.

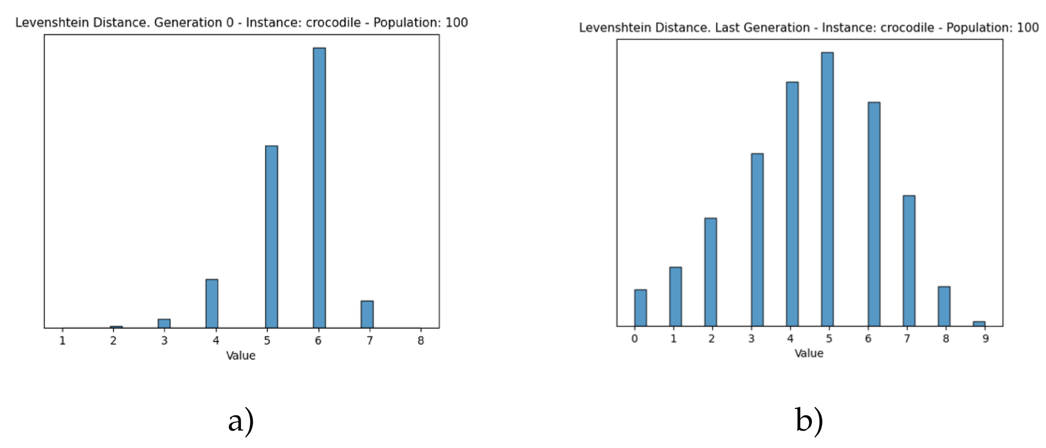

Figure 10: Levenshtein Distance – instance: crocodile – Population: 100. a) Generation 0 and b) Last Generation.

Result analysis

As shown in the boxplots (Figures 3-6), despite population size growth, the median burning number stays the same. Therefore, based on the results, it can be argued that random search is not a good strategy for the graph burning problem. Figures 7-10 show the distribution of distances between the solutions in the EA. From these figures, we observe that in the initial generation, we have just a few distances with a value close to 0. Nonetheless, in the last generation, we started having more distances closer to 0. This means a loss of diversity since there are some equal burning sequences.

Conclusion and future work

- Random search is not the best approach to solve the graph burning problem.
- It is not necessary to make use of a large population size to solve the graph burning problem.
- Diversity loss exists in the evolutionary algorithm.
- Possible future work consists of finding a technique that handles with diversity loss in EAs.

References

[1] P. A. Vikhar, "Evolutionary algorithms: A critical review and its future prospects," in *2016 International conference on global trends in signal processing, information computing and communication (ICGTSPICC)*, pp. 261–265, IEEE, 2016.

[2] A. Pétrowski and S. Ben-Hamida, *Evolutionary algorithms*. John Wiley & Sons, 2017.

[3] A. Bonato and S. Kamali, "Approximation algorithms for graph burning," in *Theory and Applications of Models of Computation: 15th Annual Conference, TAMC 2019, Kitakyushu, Japan, April 13–16, 2019, Proceedings 15*, pp. 74–92, Springer, 2019.

[4] Z. R. Farokh, M. Tahmasbi, Z. H. R. A. Tehrani, and Y. Buali, "New heuristics for burning graphs," *arXiv preprint arXiv:2003.09314*, 2020.

[5] R. K. Gautam and A. S. Kare, "Faster heuristics for graph burning," *Applied Intelligence*, vol. 52, no. 2, pp. 1351–1361, 2022.

[6] J. García-Díaz, J. C. Pérez-Sansalvador, L. M. X. Rodríguez-Henríquez, and J. A. Cornejo-Acosta, "Burning graphs through farthest-first traversal," *IEEE Access*, vol. 10, pp. 30395–30404, 2022.

[7] J. García-Díaz, L. M. X. Rodríguez-Henríquez, J. C. Pérez-Sansalvador, and S. E. Pomares-Hernández, "Graph burning: Mathematical formulations and optimal solutions," *Mathematics*, vol. 10, no. 15, p. 2777, 2022.

[8] M. Nazeri, A. Mollahosseini, and I. Izadi, "Cbag: An efficient genetic algorithm for the graph burning problem," *arXiv preprint arXiv:2208.01008*, 2022.

[9] J. García-Díaz and J. A. Cornejo-Acosta, "A greedy heuristic for graph burning," *arXiv preprint arXiv:2401.07577*, 2024.

[10] F. d. C. Pereira, P. J. de Rezende, T. Yunes, and L. F. B. Morato, "Solving the graph burning problem for large graphs," *arXiv preprint arXiv:2404.17080*, 2024.

[11] A. Bonato, J. Janssen, and E. Roshanbin, "Burning a graph as a model of social contagion," in *Algorithms and Models for the Web Graph: 11th International Workshop, WAW 2014, Beijing, China, December 17-18, 2014, Proceedings 11*, pp. 13–22, Springer, 2014.