Int. J. Nonlinear Anal. Appl. 10 (2019) No. 1, 185-192
ISSN: 2008-6822 (electronic)
https:dx.doi.org/10.22075/IJNAA.2019.4063

A Novel Approach for Detecting Relationships in Social Networks Using Cellular Automata Based Graph Coloring

M. Kashani, S.V. Shojaedini and S. Gorgin

Department of Electrical Engineering and Information Technology Iranian Research Organization for Science and Technology (IROST), Tehran, Iran

(Communicated by j. Vahidi)

Abstract

All the social networks can be modeled as a graph, where each roles as vertex and each relation roles as an edge. The graph can be show as $G = [V, E]$, where V is the set of vertices and E is the set of edges. All social networks can be segmented to K groups, where there are members in each group with same features. In each group each person knows other individuals and is in touch with them. In this study, the main goal is introducing a new approach for detecting these groups and minimizing the number of these groups using a cellular automat algorithm. There are two types of social networks, containing simulated social network and real social network. The results show that the introduced method has a great potential to significantly reduce the number of colors assigned and running time of the program.

Keywords: Social networks, Graph coloring, Cellular automata.
2010 MSC: 68R10

1. Introduction

A graph $G = [V, E]$, where V is the set of vertices and E is the set of edges, can be considered as a model of social networks, such that each vertex represents the individuals in the respective social network, and each edge represents the relationships between individuals in that social network [1, 2]. A wide range of complex problems can be analyzed using the Theory of Graphs. Recognizing relationships and sub-classes in social networks is one of these complex problem [3].

*Corresponding author
Email address: m.kashani@irost.ir (M. Kashani, S.V. Shojaedini and S. Gorgin)

Received: February 2019 *Revised:* August 2019

Graph coloring is one of the existing methods for labeling the graph components subject to specific constraints, such that the number of graph colors reduces to the minimum. In the proposed algorithm, a graph is colored in such a way that no two adjacent vertices have the same color. This type of graph coloring is called vertex coloring. It can be interpreted that vertices which have no tie between them can have the same color. Similarly, the edges can also be colored in such a way that two adjacent edges fail to have the same color. Graph coloring can be used in a variety of problems, including biochemistry, electrical engineering, network engineering, and resource allocation [2, 4].

To date, various studies have dealt with the field of social networks analysis using graph coloring. Take [5], for example, as one of these studies conducted in 2011. This study presented a heuristic algorithm using a graph G and its complement graph defined as $\bar{G}$. Graph G was defined as the set of vertices and edges that represent the individuals and relationships, respectively. On the other, its complement graph $\bar{G}$ was defined as $\bar{G} = [\bar{V}, \bar{E}]$, where $\bar{E}$ is a set of edges serving as the complement for the set of edges referred to as E. In general, a social network has two subgroups, namely V and $\bar{V}$, where V refers to a set of individuals in the social network that are connected to one another, and $\bar{V}$ refers to a set of individuals in the social network that are not connected to each other. The aforementioned study investigated the number of friendships in a social network via the complementary graph coloring over the social network. It was shown that the number of colors existing in the complement graph denotes to the number of friendships in the initial graph. This study allows one to find a friend fits best with him/her [6-8].

Graph coloring is by no means an easy task because it is intended to minimize the number of colors available. In [9], an algorithm was presented to find the maximum number of independent sets in complex graphs. This algorithm was implemented on graphs of different sizes, e.g., $20, 30, 40, \cdots$ and 2000. In this study, it was claimed that the speed of this algorithm is much higher than its other counterparts used for complex graphs. In another study[6], the weights of the vertices were used to find the maximum number of independent sets, in which the weight of vertex would be proportional to the degree of that vertex. The study also indicated that the complexity of this algorithm partly depends on the number of vertices existing in the graph. Other studies have used graph coloring for clustering properties and data [10], for example, classification of students under appropriate vertices[11, 12].

2. RELATED WORK

In 2014, a study was conducted that presented three strategies to identify users based on their data entry process in a social network [13]. The first strategy was a spatiotemporal strategy that depended on the time when the user entered the network. The second strategy was dependent on the repetition and the number of entries and visits from a specific page. The third strategy was a combination of the first two strategies. In another study, it was shown that Artificial Neuro-Fuzzy Logic System could be used to detect the individual emotions and behaviors in social networks [14]. These emotions can also be considered as crucial indicators to match people together for friendship [13].

In another study, it was argued that merely a limited number of friends could be considered for each individual. Therefore, choosing the best friends for each user, i.e., friends feeling the bond of true personality compatibility, would be a crucial and difficult task [15]. Obviously, recognizing the personality of individuals to determine the level of compatibility may take considerable time; however, processing the relationships of users seems to be a key strategy to scrutinize their emotions and spirits and minimize the time required to this end.

There are many other applications for graph coloring, such as the optimization of the examination timetable [16]. In this case, no student should have two exams on the same day. Reducing the time buffers between the exams is another change applied to the examination timetable.

Social networks modeling through a graph may require a large-scale graph processing due to the considerably large number of social networks users. Plenty of studies have provided methods to optimize the graph coloring [17].

In 2018, a study derived to develop a decentralized coloring approach to diversify the nodes in a complex network. The key is the introduction of a local conflict index that measures the color conflicts arising at each node which can be efficiently computed using only local information. We demonstrate via both synthetic and real-world networks that the proposed approach significantly outperforms random coloring as measured by the size of the largest color-induced connected component. Interestingly, for scale-free networks further improvement of diversity can be achieved by tuning a degree-biasing weighting parameter in the local conflict index [18].

In 2017, a study introduced a hybrid approach of graph coloring and ACO based summarization for social networks. The aim of this work was to enable the users to get a powerful brief of comments without reading the entire list. This paper opens up a new field of short text summarization (STS) predicated on a hybrid ant colony optimization coming with a mechanism of local search, called ACO-LS-STS, to produce an optimal or near-optimal summary. Initially, the graph coloring algorithm, called GC-ISTS, was employed before to shrink the solution area of ants to small sets. Evidently, the main purpose of using the GC algorithm is to make the search process more facilitated, faster and prevents the ants from falling into the local optimum. First, the dissimilar comments are assembled together into the same color, at the same time preserving the information ratio as for an original list of comment. Subsequently, activating the ACO-LS-STS algorithm, which is a novel technique concerning the extraction of the most interactive comments from each color in a parallel form. At the end, the best summary is picked from the best color. This problem is formalized as an optimization problem utilizing GC and ACO-LS to generate the optimal solution. Eventually, the proposed algorithm was evaluated and tested over a collection of Facebook messages with their associated comments [19].

The present study also intends to introduce a new graph coloring method based on the cellular automata and analyze the results by taking into account the improvement in recognition of relationships in social networks. Finally, the results obtained are intended to be compared to the results reported by the previous studies.

3. MATERIALS AND METHODS

As mentioned earlier, this study intends to optimize the relationships in social networks based on the cellular automaton algorithm and complementary graphs. To this end, the complement graph is initially extracted from the initial graph, and then the former is colored using cellular automata. The number of colors in the complement graph presents the number of groups in the social network. Therefore, this section initially describes the complement graph concept, and finally, the graph coloring method will be introduced using cellular automata.

3.1. Complement graph

The concept of a graph which is a model of a social network is firstly dealt with to delineate a complement graph. If a social network is modeled by a graph as $G = (V, E)$, two sets are defined within this graph, i.e., V and E. The former, V, refers to the graph vertices representing the individuals in a social network and is expressed as Eq. 3.1; and the latter, i.e., E, refers to the graph edges representing the existence or non-existence of a relationship between two distinct individuals and is defined as Eq. 3.2.

$$V = [V_1, V_2, \cdots, V_n] \tag{3.1}$$

$$E(i,n) = \begin{cases} 1 & \text{if i.j are adjancent} \\ 0 & \text{if i.j are not adjancent} \end{cases} \tag{3.2}$$

More precisely, the combination of the main graph and its complement gives a complete graph. The following figure displays a sample of a graph and its complement.

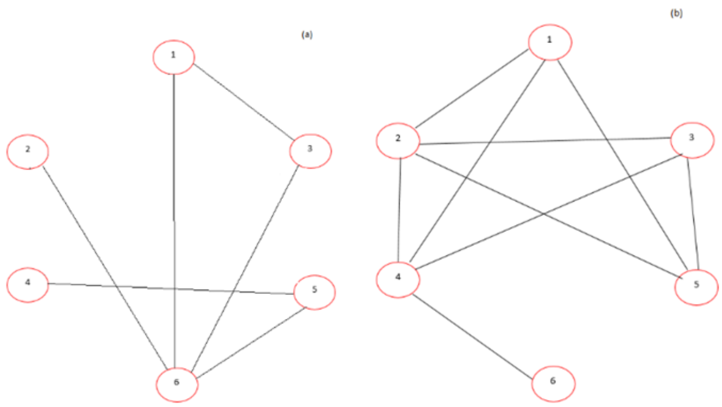

Figure 1: (a) the main graph (b) the complement graph

3.2. Cellular automaton

In this study, graph coloring was conducted using an adaptive cellular automaton algorithm to classify the groups in the social network. In this algorithm, the following three characteristics should be taken into account:

1 All vertices must be colored.

2 Vertices sharing the same color should not be juxtaposed.

3 The minimum number of colors should be applied

4 To this end, the following steps were followed:

5 All vertices were assigned color 1.

6 The value of n was set to 2.

In case vertex n juxtaposed to or shared the same color with vertices 1 to $n-1$, one color was added to the color of vertex n.
$n = n + 1$ was obtained, and step 3 was applied until all graph vertices were checked.
In the following, an example of graph coloring by the given method is presented.

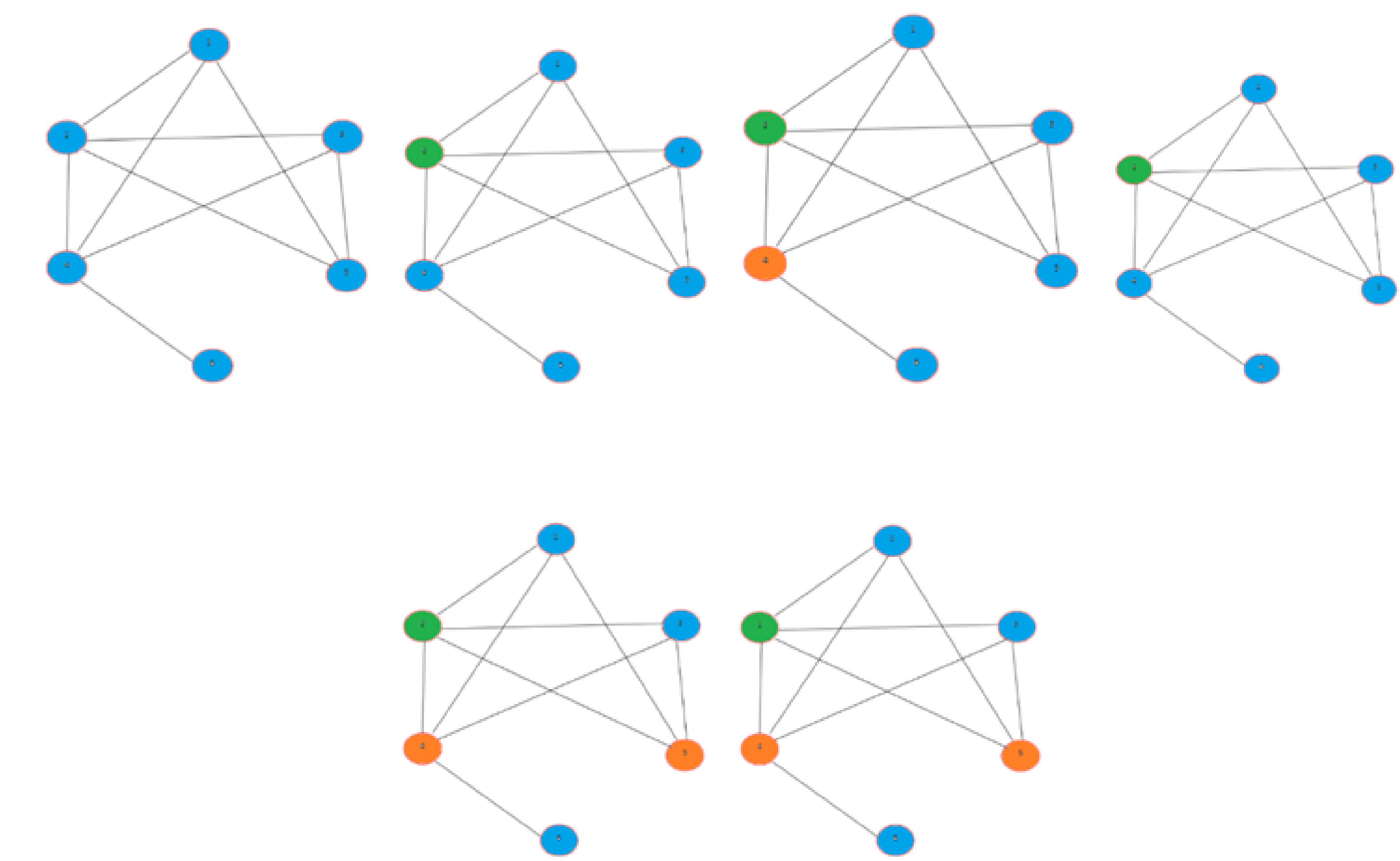

Figure 2: Sample graph coloring scheme

In Figure 2, the graph coloring was conducted in such a way that all vertices were initially assigned a single color, i.e., color 1. Then, the second vertex v2 was checked to understand whether there is a vertex with color 1 in its neighborhood. If there was, one color was added to that vertex thereby the second vertex v2 received color 2. Similarly, the third vertex v3 with color 1, was checked, and as it was neither adjacent to the previous vertices (i.e., v1 and v2) nor of the same color with them, it was allowed to be color 1. The fourth vertex v4 was also checked in a similar way and compared with the previous vertices (i.e., v1, v2, and v3). The comparison revealed that it is adjacent to v3 with color 1. Hence it had to be colored 2. But it was not allowed, because it shared the same color with v2 (i.e., color 2). Therefore, it was assigned color 3, which was distinct from the colors of the previous vertices. The fifth vertex v5 received color 3 according to the respective algorithm. Regarding the above considerations, the color of vertex 6 remained unchanged, i.e., color 1. Accordingly, the whole vertices were colored in such a way that no two adjacent vertices had the same color, and generally, three colors were assigned to the graph at issue.

4. RESULTS

This section presents the results obtained from applying the proposed method to three graphs simulated from social networks. In this coloring method, three graphs simulated in [5], which composed of 100, 500, and 2,000 vertices, were used. In these graphs, the number of colors for graphs with 100, 500, and 2,000 vertices was 20, 100, and 400, respectively.

Table 1: Comparison of the results associated with the proposed method and the method presented in [5]

Graph name	Number of colours in the method presented in [chapula]	Number of colours in the proposed method	Running time in the method presented in [5]	Running time in the proposed method
Artif-100-20	20	20	< 1 sec	< 1 sec
Artif-500-100	100	100	3 sec	2.3 sec
Artif-2000-400	400	400	51 sec	13.6 sec

Table 2: Comparison of the results of the proposed method with Greedy and Random methods

Graph name	Number of colors in Greedy	Number of colors in Random	Number of colors in the proposed method	Running time in Greedy	Running time in Random	Running time in the proposed method
Facebook-like Social Network	11	18	10	1.3 sec	1.08sec	1.01 sec
Facebook-like Forum Network	126	126	123	1.3 sec	1.3sec	1.21 sec

For further comparison of the proposed method with other approaches, e.g., Greedy and Random, data of Facebook social networks, including Facebook-like Social Network and Facebook-like Forum Network, were used. The Facebook-like Social Network graph consisted of 1899 vertices and 20296 edges, and the Facebook-like Forum Network graph included 899 vertices and 142760 edges.

In the following, the proposed method is compared to the method presented in [5] in terms of the running time and the number of colors assigned. Comparison of the results associated with these two methods is presented in Table 1. As observed in Table 1, two methods detected the same number of colors, while running time in the proposed method is significantly lower compared to the method presented in [5].

In the following, Greedy and Random approaches are compared with the proposed method in terms of the number of colors and running time. The results of the comparison are presented in Table 2.

As shown in Table 2, the proposed method considerably outperforms the other two methods in terms of the number of colors and the running time.

5. CONCLUSION

The present study aimed to provide a new method for minimizing the running time of the graph coloring process as well as the number of colors to determine the number of sub-groups in a social network. To this end, the social network was initially simulated as a graph in which the vertices denote the individuals present in the social network, and the edges denote the relationships between individuals. Afterward, the complement graph was established, and finally, the number of complement graph colors which represents the number of subgroups in a social network was measured using the proposed method.

The results obtained were analyzed using two sets of data, the first one consisted of 3 simulated graphs, and the second one contained two graphs of data associated with social networks. For the first analysis, it had to be determined whether the proposed method detects an optimal number of colors or not. To this end, data from three simulated graphs were used. The respective data revealed that the proposed method has the potential to not only identify the optimal number of colors but also considerably reduce the running time. The result of this section is illustrated as figures 3 and 4. As it is obvious from figure 3, chapulas method and our method are completely for coloring but from figure4, in timing, for small graphs chapula's method and our method are relatively same but for huge graphs, our method is significantly better. In the second stage, the results of the proposed method were analyzed using two graphs of actual social network data and compared to the results of Greedy and Random methods. The results are compared in figures 5 and 6. As the results show, our method is better than other methods, both in coloring and timing.

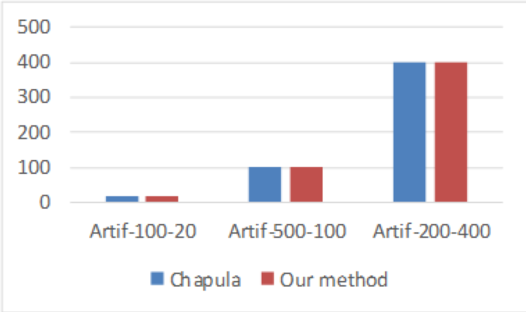

Figure 3: Comparing our method and chapula for colors in artificial social networks

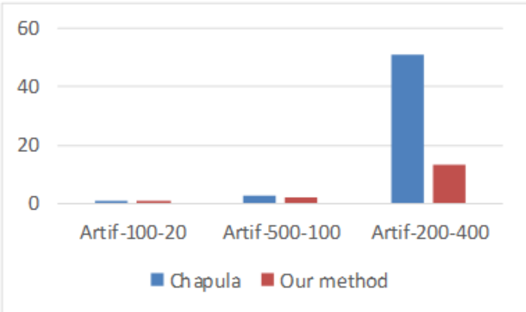

Figure 4: Comparing our method with chapula for timing in artificial social networks

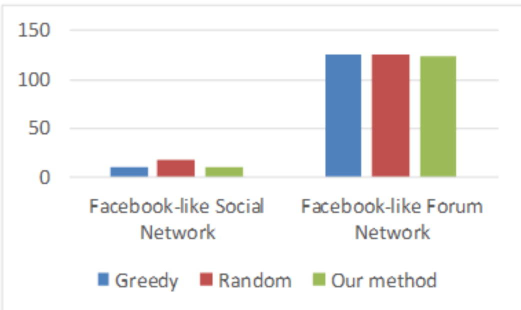

Figure 5: Comparing our method with Greedy and Random for colors in real social networks

Investigations shed light on the fact that the proposed method considerably outperforms the other two methods in terms of the number of colors and the running time.

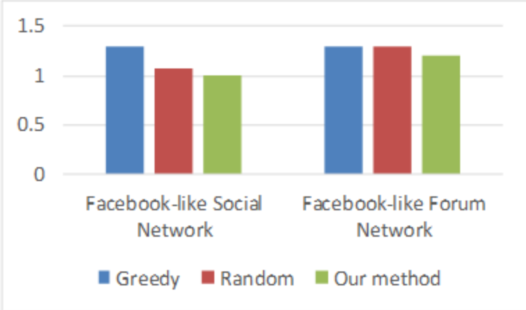

Figure 6: Comparing our method with Greedy and Random for timing in real social networks

References

[1] A. Hertz, and D. de Werra, Using tabu search techniques for graph coloring. Computing., Vol39, 1987, 345-351.

[2] T. Cormen, and C. Leiserson,RL. Rivest , and C. Stein, Introduction to algorithms. Cambridge, MA: The MIT Press, 2009.

[3] O. A. Murad, Sleit and A. Sharaiah, Improving Friends Matching in Social Networks Using Graph Coloring. International Journal, Vol15(8), 2016.

[4] R.A. Rossi, and N.K. Ahmed, Coloring large complex networks. Social Network Analysis and Mining, Vol4(1), 2014, 228-229.

[5] D. Chalupa, On the ability of graph coloring heuristics to find substructures in social networks. Information Sciences and Technologies, Bulletin of ACM Slovakia, Vol3(2), 2011, 51-54.

[6] A. al-Jabir, Algorithms Based on Weight Factors for Maximum Independent Set 1999.

[7] D. Brlaz, New methods to color the vertices of a graph. Communications of the ACM, Vol22(4)1979 251-256.

[8] I. Guy, et al, Do you want to know?: recommending strangers in the enterprise. in Proceedings of the ACM 2011 conference on Computer supported cooperative work. 2011. ACM.

[9] A. Sharieh, et al. , An algorithm for finding maximum independent set in a graph. European Journal of Scientific Research, Vol23(4), 2008, 586-596.

[10] A. Sleit, S. Abusharkh, and W. AlMobaideen, Enhancing Modularity-Based Graph Clustering. World Applied Sciences Journal, Vol9(9), 2010, 984-996.

[11] I. Salah, et al., Efficient Method for Assigning Students to Proper Groups. Editorial Advisory Board e, 21(2), 2005, 249-358.

[12] G. Karypis, and V. Kumar. Analysis of multilevel graph partitioning. in Supercomputing, 1995. Proceedings of the IEEE/ACM SC95 Conference. 1995. IEEE.

[13] L. Rossi, and M. Musolesi. It's the way you check-in: identifying users in location-based social networks. in Proceedings of the second ACM conference on Online social networks. 2014. ACM.

[14] M. Malkawi, and O. Murad, Artificial neuro fuzzy logic system for detecting human emotions. Human-Centric Computing and Information Sciences, 3(1),2013, 3-5.

[15] R. Dunbar, How many friends does one person need?: Dunbar's number and other evolutionary quirks. 2010: Faber and Faber.

[16] M. Malkawi, M.A.-H. Hassan, and O.A.-H. Hassan, A New Exam Scheduling Algorithm Using Graph Coloring. International Arab Journal of Information Technology (IAJIT), Vol5(1), 2008.

[17] V. Donderia, and P. K. Jana, A novel scheme for graph coloring. Procedia Technology, Vol4, 2012, 261-266.

[18] R. Garcia-Lebron, et al. "Node diversification in complex networks by decentralized coloring." arXiv preprint arXiv:1811.11197 2018.

[19] M. Mohamed Atef, A. Hamouda, and M. Marei. "Graph coloring and ACO based summarization for social networks." Expert Systems with Applications, Vol74 2017, 115-126.