LVI SBPO SIMPÓSIO BRASILEIRO DE PESQUISA OPERACIONAL
DE 04 ATÉ 07 NOVEMBRO 2024
FORTALEZA | CEARÁ

Application of biased random-key genetic algorithm and formulations for the Grundy coloring problem and the connected Grundy coloring problem

Mateus C. Silva, Rafael A. Melo
Instituto de Computação, Universidade Federal da Bahia
Salvador, BA, Brasil
mateuscs@ufba.br, rafael.melo@ufba.br

Marcio C. Santos
Departamento de Ciência da Computação, Universidade Federal de Minas Gerais
Belo Horizonte, MG, Brasil
marciocs@dcc.ufmg.br

ABSTRACT

The master thesis summarized in this article addresses the Grundy coloring problem and its connected variant. These problems define the worst-case behavior for the well-known and widely used first-fit greedy coloring heuristic (in the second case, with connectivity restrictions). First, we provided a new combinatorial upper bound that can improve over well-established ones from the literature. Moreover, the bound was used to enhance some of the proposed methods. Second, we provided the first optimization approaches to solve the problems for general graphs. We proposed integer programming formulations and biased random-key genetic algorithms (BRKGAs) for the two problem variants. We also filled the gap in the literature by providing the first computational tests for the problems through an extensive benchmark that allowed the evaluation of the proposed approaches and the verification of their practical applicability as a method for evaluating greedy criteria for graph coloring.

KEYWORDS. Vertex coloring. Integer programming. BRKGA.

Paper Topics (Dissertation Awards, Combinatorial Optimization, Graph Theory and Algorithms)

RESUMO

A dissertação resumida neste artigo aborda o problema de coloração de Grundy e sua versão conexa. Esses problemas definem o comportamento de pior caso para a conhecida e amplamente utilizada heurística de coloração gulosa *first-fit* (no segundo caso, com restrições de conectividade). Inicialmente, fornecemos um novo limite superior combinatório capaz de melhorar limites bem estabelecidos na literatura. Além disso, o limite foi usado para aprimorar alguns dos métodos propostos. Em segundo lugar, fornecemos as primeiras abordagens de otimização para resolver os problemas para grafos gerais. Propusemos formulações de programação inteira e algoritmos genéticos de chave aleatória viesados (BRKGAs) para as duas variantes do problema. Além disso, preenchemos outra lacuna na literatura ao fornecer os primeiros testes computacionais para os problemas, através de um extenso *benchmark* que permitiu avaliar as abordagens utilizadas e verificar a aplicabilidade prática como um método de avaliar critérios gulosos para coloração em grafos.

PALAVRAS CHAVE. Coloração de vértices. Programação inteira. BRKGA.

Tópicos (Prêmio de dissertação, Otimização Combinatória, Teoria e Algoritmos em Grafos)

1. Introduction

Combinatorial optimization problems (COPs) consist of obtaining the best solution (minimum or maximum) in a discrete set of possible solutions. Some COPs can be represented as graph coloring problems and have many practical applications. Some examples include scheduling [Gamache et al., 2007], timetabling [Babaei et al., 2015], communication networks [Pateromichelakis e Samdanis, 2018], and railway station design [Jovanović et al., 2020]. Several of these problems are NP-hard, so it is not known whether there is an algorithm that can optimally solve them efficiently (polynomial time).

When it comes to optimization methods, we can classify them based on their capacity to certify optimality. This classification includes exact methods and heuristics. The main difference is that given enough time, an exact method will find the optimal solution. In contrast, a heuristic may not be able to certify that it has found the optimal solution. However, this extra time for the exact methods to prove they found the optimal solution can be very long. Sometimes, this approach could take a long time even to find an initial solution to a problem. Although the heuristic does not guarantee an optimal solution, it can provide good-quality solutions in a reasonable time. Thus, it would be interesting to evaluate how good a heuristic is to solve a problem and even more to analyze its worst case (if the result is close to the worst possible value).

One of the most famous heuristic approaches to coloring graph problems is the greedy first-fit heuristic or, simply, the greed heuristic. As mentioned earlier, there may be no guarantee regarding the quality of solutions yielded by this approach, and an evaluation of this property would be interesting. To assist in this analysis, we consider the following concepts. Given a simple undirected graph G, its Grundy chromatic number $\Gamma(G)$ (or Grundy number) defines the worst-case behavior for the well-known and widely-used greedy first-fit coloring heuristic. Specifically, $\Gamma(G)$ is the largest k for which a k-coloring can be obtained with the first-fit heuristic. Analogously, the connected Grundy chromatic number $\Gamma_c(G)$ is the largest k for which a connected k-coloring can be obtained with the first-fit heuristic.

Grundy [1939] first studied the concept of Grundy number in the context of game theory and digraphs [Berge, 1973], where $\Gamma(G)$ was related to the impartiality of a given game state. Christen e Selkow [1979] formally introduced this concept in graph theory. The *connected Grundy number* was only introduced in 2014 [Benevides et al., 2014]. Studies on the Grundy number and its connected variant have focused primarily on the complexity for specific graph classes [Goyal e Vishwanathan, 1997; Zaker, 2005, 2006; Bonnet et al., 2018; Effantin e Kheddouci, 2007]. Despite this, some algorithms have already been proposed to solve the problem, such as a linear time algorithm to find a Grundy coloring in trees [Hedetniemi et al., 1982], this result being extended in [Telle e Proskurowski, 1997] who proposed a dynamic programming algorithm with complexity $k^{O(w)} * 2^{O(wk)} * n = O(n^{3w^2})$, where w is the (*treewidth*). Bonnet et al. [2018] presented an exact exponential time algorithm $O(2.443^n)$ to the Grundy number problem but left open the question whether there is an $O(c^n)$ exact algorithm for the connected problem, with c a constant. However, there are no computational results from applying such proposed algorithms.

1.1. Main contributions and organization

The main contributions of this master thesis were **(a)** the study of the effectiveness of the *biased random-key genetic algorithm* (BRKGA) and **(b)** formulations to determine the Grundy and connected Grundy numbers, using two integer programming formulations approaches: standard and by representatives. Additionally, we analyze the problem's structure to derive an upper bound based on the vertex neighborhood. Thus, **(c)** we provide a new combinatorial upper bound that is valid for both problems and can be calculated in polynomial time using dynamic programming. Then, **(d)** we used this new bound together with others from the literature to reinforce the set of constraints of the formulation and its performance, reducing the total number of variables and constraints generated. Also, **(e)** we expanded the set of common instances for coloring problems used in the literature. Furthermore, **(f)** we provided a method that allows connecting unconnected instances that aim to preserve the connected Grundy number of each component, allowing us to reuse instances from the literature. Moreover, **(g)** we established the first benchmark for the problems through computational tests with a diverse set of instances since there are no tests on instances of the approaches proposed in the literature. Finally, **(h)** we compared greedy criteria that aim to minimize the number of colors, representing the worst-case scenario for them.

We organize the remainder of this article as follows. Section 2 provides the new combinatorial upper bound and a polynomial-time algorithm for its calculation. Section 3 briefly describes the ideas used in the integer programming (IP) formulations and an excerpt of the results. Section 4 explains the idea of applying the BRKGA, the tests to evaluate solutions, the BRKGA's capability to converge to high-quality solutions, and a snippet of the results. Section 5 displays the publications derived from the master's research. Section 6 mentions details of the defense.

2. A new combinatorial upper bound

Consider the values $\psi(v, k)$ for $v \in V$ and $k \in \{1, \dots, \Delta(G)+1\}$ recursively as

$$\psi(v,k) = \begin{cases} \max\{l \mid \exists (u_1, \dots, u_{l-1}) \subseteq N(v) \text{ s.t. } \psi(u_i, k-1) \geq i, \forall i, 1 \leq i \leq l-1\}, \text{if } k \geq 2; \\ 1, \text{otherwise.} \end{cases}$$

Define the *connected degree sequence value* $\Psi(G)$ as:

$$\Psi(G) = \max_{u \in V} \{\psi(u, \Delta(G)+1)\}.$$

Proposition 1. $\Gamma(G) \leq \Psi(G) \leq \Delta(G) + 1$.

Proof. Consider any feasible Grundy coloring c using k colors. Notice that any vertex can receive color 1. On the other hand, for a vertex v to receive color k, all the other colors in $\{1, \dots, k-1\}$ have to be already used for its neighbors. Recursively, each neighbor receiving color $i \in \{2, \dots, k-1\}$ must have at least one neighbor with each of the colors in $\{1, \dots, i-1\}$. Therefore, $c(v) \leq \psi(v, c(v))$. In addition, as ψ is a nondecreasing function of k, it follows that $\psi(v, c(v)) \leq \psi(v, \Delta(G)+1)$. Hence, $\Gamma(G) \leq \Psi(G)$. □

Proposition 2. *$\Psi(G)$ can be calculated in polynomial time.*

Proof. We provide an $O(|V|\Delta(G)^2)$ dynamic programming-based algorithm for calculating $\Psi(G)$. We remark that $\psi(v, k)$ can be calculated in linear time $O(\Delta(G))$ using dynamic programming for each pair (v, k) if the values for $\psi(u, k-1)$ are known for every $u \in V$. To see how, assume that the vertices in $N(v)$ are in nondecreasing order based on the values $\psi(u, k-1)$, as such an ordering can be performed in $O(\Delta(G))$ using counting sort. Define $M(j)$ to be the largest l such that there is a sequence $(0, 1, \dots, l)$ in the vertices indexed from 0 to j, where index 0 corresponds to a dummy vertex and $0 \leq j \leq |N(v)|$. Besides, let $P(j)$ be the last element in such a sequence. Define $M(0) = P(0) = 0$. Thus, if $d(v_j) \leq M(j-1)$, we set $M(j) = M(j-1)$ and $P(j) = P(j-1)$. Otherwise, we set $M(j) = M(j-1)+1$ and $P(j) = j$. Finally, $\Psi(G) = \max_{u \in V}\{\psi(u, \Delta(G)+1)\}$. Thus, the $O(|V|\Delta(G))$ elements of ψ can be calculated in $O(|\Delta(G)|)$ each, implying a total running time of $O(|V|\Delta(G)^2)$. □

Now, we summarize the results achieved using the new upper bound $\Psi(G)$ when compared to the available bounds $\Delta(G)+1$, $\Delta_2(G)+1$, and $\zeta(G)$. The results in Table 1 indicate that the new upper bound can improve over $\zeta(G)$ for several instances (19.0%). Besides, it could at least match the $\Delta_2(G) + 1$ results for more than half of the instances. Most of the improvements over $\zeta(G)$ (fifth column) are concentrated on the bipartite graphs (63.7%), followed by the DIMACS instances (19.0%) and the random graphs (11.3%). When it comes to the improvements over $\Delta_2(G)+1$ (third column), we can highlight the DIMACS and the bipartite instances, for which strictly better bounds were obtained for, respectively, 40.4% and 17.5% of them.

Table 1: Comparison of upper bounds

class	$\Psi(G) < \Delta(G)+1$	$\Psi(G) < \Delta_2(G)+1$	$\Psi(G) \leq \Delta_2(G)+1$	$\Psi(G) < \zeta(G)$	$\Psi(G) \leq \zeta(G)$
Random	60.0%	8.7%	40.0%	11.3%	33.7%
Geometric	32.5%	17.5%	62.5%	1.2%	15.0%
Bipartite	73.7%	2.5%	35.0%	63.7%	95.0%
Complement	22.5%	2.5%	46.2%	0.0%	5.0%
DIMACS	54.7%	40.4%	90.4%	19.0%	30.0%
All instances	48.0%	11.6%	51.1%	19.0%	36.3%

3. Integer programming formulations

This section discusses the main concepts used in the IP formulations for the problems and presents some of the corresponding results. We proposed the first IP formulations for the problems, with two approaches for each. The first is a standard approach of partitioning the vertices by color class, while the second is a formulation by representatives. The latter attempts to overcome the symmetries in the problem and relies on the idea that a subset of the vertices in the graph can be represented by one of its vertices, denoted as a representative.

We combined the newly proposed bound with other bounds from the literature to enhance the formulations and reduce the total number of variables and constraints generated, making them more effective. With this, it was possible to verify that the formulation by representatives achieves better results in dense graphs while the standard one performs better in sparse graphs.

In what follows, we present the definition of one of the variables used in the standard formulation and how the Grundy property can be enforced in such a formulation:

$$x_{kv} = \begin{cases} 1, & \text{if vertex } v \in V \text{ receives color } k \in K_v\text{, i.e., the } k\text{-th color,} \\ 0, & \text{otherwise;} \end{cases}$$

$$x_{k'v} \leq \sum_{u \in N(v) \cap V_k} x_{ku}, \ \forall v \in V, \ k, k' \in K_v, \text{ with } k < k'.$$

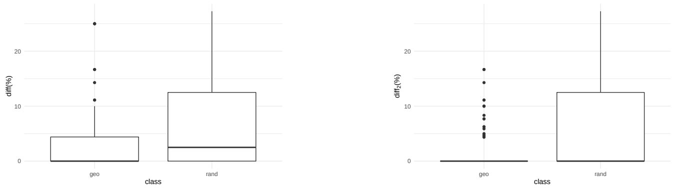
We restricted the sets that are used to generate the variables and constraints using the bounds, as follows: $K_v = \{k' \in K \mid k' \leq \min(\zeta(G), \Delta_2(G) + 1, \psi(v, \Delta(G) + 1))\}$ is the sequence of possible colors for the vertex v and $V_k = \{v \in V \mid k \in K_v\}$ is the set of vertices that can receive the k-th color. Note that in the constraint, bounds were also used to limit the number of constraints that are generated and the total number of variables in a constraint as well. In the representative approach, we will present two variables and the constraints that enforce the Grundy property:

$$X_{vu} = \begin{cases} 1, & \text{if vertex } v \in V \text{ represents the color of vertex } u \in \bar{N}(v)\text{, for } v \leq u; \\ 0, & \text{otherwise;} \end{cases}$$

$$y_{vu} = \begin{cases} 1, \text{if vertices } v, u \in V \text{ are representatives and the color of } v \text{ precedes that of } u\text{, for } v \neq u, \\ 0, \text{otherwise;} \end{cases}$$

$$X_{uv} \leq \sum_{\substack{w \in N(v) \cap \bar{N}[p], \\ w \geq p}} X_{pw} + 1 - y_{pu}, \; \forall\, u, p \in V, \; v \in \bar{N}[u], \; \text{ s.t. } p \neq u \text{ and } \; u \leq v.$$

The Grundy coloring problem and its connected variant proved challenging for the formulations, even for the instances with 50 to 80 vertices. The formulations, despite being able to improve the results from the initial solution provided, failed to prove optimality for most of the instances within the time limit of 3600 seconds. It is worth mentioning that, for the connected variant, the formulations found a solution for only 26% of the instances. This shows that these IP formulations are computationally impractical for the connected version of graphs with more than 50 vertices. However, considering random and geometric classes with up to 30 vertices, the formulation found the optimal solution for 61.25% of cases. Furthermore, in the subset of instances mentioned, it was seen that the median percentage difference for the solutions to the Grundy coloring problem (in which the connectivity requirements are relaxed) is close to 2.5%, Figure 1. It shows that, in these cases, it is reasonable to use the solution found by the models of the Grundy coloring problem to evaluate the solution of its connected version.



(a) Boxplots comparing the deviation between the best solutions of the formulations for $\Gamma(G)$ and $\Gamma_c(G)$.

(b) Boxplots comparing the deviation from the best upper bound with the best solution for $\Gamma_c(G)$.

Figure 1: Boxplots comparing the results of the formulations for $\Gamma(G)$ and $\Gamma_c(G)$

The results in Table 2 show that the formulation by representatives outperformed the standard formulation, proving optimality (*#opt* column) for 304 out of 320 (95%) instances in the set with up to 30 vertices for the Grundy coloring problem. It achieved 77 more optimal solutions (24%) when compared to the standard formulation. Furthermore, the largest mean gap for a specific class happened in the random graphs for both formulations, with a value of 13.2% for the standard formulation against only 2.1% for the one by representatives.

Table 2: Summary of the results using the formulations for smaller instances with up to 30 vertices, separated by classes for the Grundy coloring problem

	std			rep		
class	gap	time	#opt	gap	time	#opt
Random	13.8	1779.2	42	2.1	430.7	73
Geometric	2.6	1287.4	53	0.3	217.8	77
Bipartite	0.2	61.6	79	0.0	79.9	80
Complement	5.7	1295.4	53	0.4	323.1	74
Mean	5.5	1105.9		0.7	262.8	
Total			227			304

4. Biased random-key genetic algorithm

The BRKGA is a metaheuristic that fits well in both problems, as the structure that represents a solution is simple enough that one can generate a valid Grundy coloring for any order of the vertices. The BRKGA uses a decoder to decode a solution for a particular problem represented by a random-key vector, i.e., a vector of real values. This enables generating efficient decoders, and the focus of the process becomes the capacity to converge to good-quality solutions through the *intensification* and *diversification* steps. To evaluate this, we conducted extensive testing on a set of 682 instances, running each one 50 times with a time limit of 300 seconds (5 minutes). In total, testing with the BRKGA for both problems together took more than 5000 hours.

In the following, we present the BFS-based decoder for the connected version. The chromosome has size $|V|$ where each key defines the priority of the vertex to be colored. The original problem is solved with a simplified version of this decoder. In Algorithm 1, the line 1 initializes the color of all vertices to -1. The vertex with the highest key is added to the priority queue Q (lines 2-3). The loop from lines 4-10 repeats the process of selecting the vertex with the highest key from the priority queue (line 5) and coloring it with the lowest possible color following the first-fit algorithm with the auxiliary procedure Color-Vertex (line 6). Color-Vertex sets the color of a vertex as the lowest-index color that has not yet been used for any of its neighbors. Subsequently, its neighbors that have not yet been added to the priority queue are inserted into Q (lines 7-10). Line 11 returns the total number of used colors (fitness).

Algorithm 1: Decoder-connected-Grundy (G, c)

1 $colors \leftarrow \{-1, \ldots, -1\}$;
2 $v \leftarrow$ vertex with highest key value c_v;
3 Enqueue(Q, v);
4 **while** $Q \neq \varnothing$ **do**
5     $v \leftarrow$ Dequeue(Q);
6     $colors[v] \leftarrow$ Color-Vertex(G, v, $colors$);
7   **foreach** u *in* $N(v)$ **do**
8   **if** $colors[u] = -1$ **then**
9             $colors[u] = 0$;
10   Enqueue(Q, v);
11 **return** $max_{v \in V} colors[v]$;

Figure 2 shows that changing the density of the graph does not significantly affect the BRKGA's ability to find better solutions than the formulations for the random and complement of bipartite groups for the Grundy coloring problem. For the geometric graphs, it was close to or above 75% for all the density ranges. This is an interesting behavior, considering the specificity of the proposed formulations that may perform better depending on the density of the graph (as it was observed for the formulation by representatives). It is noteworthy that the BRKGA faced more difficulty in reaching the best solution for the bipartite class, This is the class where the bounds had better contributions and, consequently, the standard formulation performed well.

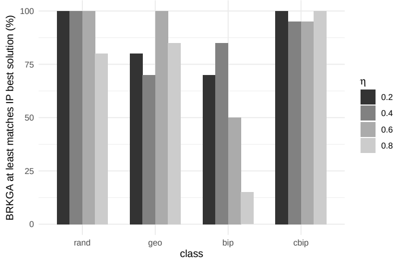


Figure 2: Barplot showing the percentage of the 320 small instances where BRKGA at least matched the best IP solution for $\Gamma(G)$, by graph class and grouped by graph density (η).

The experiments performed with the BRKGA indicate that this metaheuristic can find good-quality solutions within short computational times, achieving robustness regarding the variations between the qualities of the solutions obtained in multiple executions for both problems. Furthermore, the BRKGA found systematically better solutions than a proposed greedy algorithm to maximize the number of colors in a *first-fit* coloring, and in the great majority of the cases, solutions that at least match those obtained by the formulations. The results also showed that the BRKGA is an efficient method for worst-case analysis for the greedy criteria analyzed. Considering instances with at least 50 vertices, for more than 50% of them it took less than 100 seconds to establish a percentage difference above 50% to the greedy criteria. This percentage is not greater due to the denser instances that commonly have a small percentage difference between $\Gamma(G)$, $\Gamma_c(G)$, and $\chi(G)$.

5. Publications and Nominations

The candidate's master research was nominated for the 2024 UFBA Dissertation and Thesis Award and also led to the following publications.

1. **Journal article.** Obtaining the Grundy chromatic number: How bad can my greedy heuristic coloring be?. Computers & Operations Research (2024). Volume 168, 106703.

2. **Conference paper 1.** Algoritmos genéticos de chaves aleatórias enviesadas para o problema da coloração de Grundy. In Anais do LV Simpósio Brasileiro de Pesquisa Operacional (SBPO 2023), São José dos Campos. SOBRAPO.

3. **Conference paper 2.** Formulações de programação inteira para o problema da coloração de Grundy. In Anais do LV Simpósio Brasileiro de Pesquisa Operacional (SBPO 2023), São José dos Campos. SOBRAPO.

4. **Conference paper.** Formulações de programação inteira para o problema da coloração de Grundy conexa. Submitted to LVI Simpósio Brasileiro de Pesquisa Operacional (2024).

5. **Working paper.** The connected Grundy coloring problem: Formulations and a local-search enhanced biased random-key genetic algorithm. To be submitted to a top-tier international journal until September.

The source codes and benchmark instances for the Grundy coloring problem are available at `https://github.com/mateuscsilva/grundycoloring`.

6. About the Defense

The candidate Mateus Carvalho da Silva began his master's degree in March 2023 through the postgraduate program in computer science at the Federal University of Bahia. Less than one year later, his master thesis was defended on 13th December 2023 to meet the criteria to obtain the master's degree in Computer Science at the Federal University of Bahia (UFBA). The advisor was Dr. Rafael Augusto de Melo and the co-advisor was Dr. Márcio Costa Santos. The work was evaluated by a committee composed by: Dr. Rafael Augusto de Melo (UFBA, Brazil), Dr. Mauricio Guilherme de Carvalho Resende (University of Washington, USA), Dr. Pedro Henrique González Silva (UFRJ, Brazil), Dr. Frederico Araújo Durão (UFBA, Brazil). All the committee members unanimously approved the candidate and highlighted the high quality of the work. Link to full dissertation file `https://repositorio.ufba.br/handle/ri/39322`.

Acknowledgments

This study was funded in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001 and by CNPq, grant 314662/2020-0. This study was funded in part by the FAPESB INCITE PIE0002/2022 project. We are also grateful for the collaboration of Dr. Rodrigo F. Toso and Dr. Mauricio G. C. Resende in the development of this master's thesis.

References

Babaei, H., Karimpour, J., e Hadidi, A. (2015). A survey of approaches for university course timetabling problem. *Computers & Industrial Engineering*, 86:43–59.

Benevides, F., Campos, V., Dourado, M., Griffiths, S., Morris, R., Sampaio, L., e Silva, A. (2014). Connected greedy colourings. In Pardo, A. e Viola, A., editors, *LATIN 2014: Theoretical Informatics*, p. 433–441, Berlin, Heidelberg. Springer Berlin Heidelberg.

Berge, C. (1973). *Graphs and hypergraphs*. North-Holland Publishing Co.

Bonnet, É., Foucaud, F., Kim, E. J., e Sikora, F. (2018). Complexity of Grundy coloring and its variants. *Discrete Applied Mathematics*, 243:99–114.

Christen, C. A. e Selkow, S. M. (1979). Some perfect coloring properties of graphs. *Journal of Combinatorial Theory, Series B*, 27(1):49–59.

Effantin, B. e Kheddouci, H. (2007). Grundy number of graphs. *Discussiones Mathematicae Graph Theory*, 27(1):5–18.

Gamache, M., Hertz, A., e Ouellet, J. O. (2007). A graph coloring model for a feasibility problem in monthly crew scheduling with preferential bidding. *Computers & Operations Research*, 34(8): 2384–2395.

Goyal, N. e Vishwanathan, S. (1997). NP-completeness of undirected Grundy numbering and related problems. *Manuscript, Bombay*.

Grundy, P. M. (1939). Mathematics and games. *Eureka*, 2:6–9.

Hedetniemi, S. M., Hedetniemi, S. T., e Beyer, T. (1982). A linear algorithm for the grundy (coloring) number of a tree. *Congressus Numerantium*, 36:351–363.

Jovanović, P., Pavlović, N., Belošević, I., e Milinković, S. (2020). Graph coloring-based approach for railway station design analysis and capacity determination. *European Journal of Operational Research*, 287(1):348–360.

Pateromichelakis, E. e Samdanis, K. (2018). A graph coloring based inter-slice resource management for 5G dynamic TDD RANs. In *2018 IEEE International Conference on Communications (ICC)*, p. 1–6.

Telle, J. A. e Proskurowski, A. (1997). Algorithms for vertex partitioning problems on partial k-trees. *SIAM Journal on Discrete Mathematics*, 10(4):529–550.

Zaker, M. (2005). Grundy chromatic number of the complement of bipartite graphs. *Australasian Journal of Combinatorics*, 31:325–330.

Zaker, M. (2006). Results on the Grundy chromatic number of graphs. *Discrete Mathematics*, 306 (23):3166–3173.