Spectral Heuristics Applied to Vertex Reliability

Carla Silva Oliveira, Fausto Marques Pinheiro Junior, and José André de Moura Brito

Abstract—The operability of a network concerns its ability to remain operational, despite possible failures in its links or equipment. One may model the network through a graph to evaluate and increase this operability. Its vertices and edges correspond to the users’ equipment and their connections, respectively. In this article, the problem addressed is identifying the topological change in the graph that leads to a greater increase in the operability of the associated network, considering the case in which failure occurs in the network equipment only. More specifically, we propose two spectral heuristics to improve the vertex reliability in graphs through a single edge insertion. The performance these heuristics and others that are usually found in the literature are evaluated by computational experiments with 22000 graphs of orders 10 up to 20, generated using the Models Erdős-Rényi, Barabási-Albert, and Watts-Strogatz. From the experiments, it can be observed through analysis and application of statistical test, that one of the spectral heuristics presented a superior performance in relation to the others.

Index Terms—Graph, Vertex Reliability, Spectral Heuristic.

I. INTRODUCTION

The quality of mass telecommunication services is a key point for companies to prosper as technology dominates a competitive globalized market [1]. Such services are well-known examples of network structures composed of equipment (e.g., routers) and links (e.g., cables) that can connect millions of users.

The capacity of a network to remain operational despite possible (random) failures in some of its equipment or links is called operability. Therefore, it is important to consider the issue of operability to design networks that provide high quality of services. Similar to other problems associated with networks, a new approach was developed to solve the problem of maximizing the operability of a system by modelling its structure through a graph.

Graphs are mathematical objects that allow modelling a wide range of phenomena to identify properties due to their structure [2]. One of such properties is called reliability—the probability of a (connected) graph remaining connected after the (random) removal of an arbitrary number of vertices or edges. In particular, if only the vertices can be removed, then this property is called vertex reliability. Analogously, we have the edge reliability. Thus, the reability of the graph that models the network enables the investigation of the operability of the network.

Theoretically, the graph that models the network can pinpoint which changes to its topology lead to maximal reliability, even when given a limited number of changes. Consequently, these changes can be mapped to adjustments in the network structure that maximize the operability under any desired constraints. However, calculating the reliability of a graph belongs to the class of NP-Hard problems, which makes its exact solution unfeasible in practice for most real networks [3]. Therefore, determining which topological change causes the greatest increase in the reliability of a graph is usually made through heuristics based on its topological properties.

This article specifically addresses the problem of identifying the edge insertion on a graph that causes the greatest increase in its vertex reliability. To achieve this goal, we develop and analyse two spectral heuristics, namely: (a) the heuristic of the greatest increase in algebraic connectivity (α heuristic); and (b) the heuristic of the greatest absolute difference between the components of the Fiedler vector (ϕ heuristic). Both are not entirely original and were previously presented as a single heuristic in the context of edge reliability [4]. However, we propose a critique of this previous heuristic and the investigation of its algorithm and theoretical motivation. Thus, in addition to changing the context, we seek to split this previous heuristic and argue that an unexplored aspect in its algorithm makes it more appropriate than it should be, given the theoretical motivation that underpinned it. In particular, ϕ outperformed α and other commonly used heuristics from the literature in a set of computational experiments.

This paper is organized as follows: Section II presents the main concepts of Graph Theory and Reliability employed in the work; Section III presents a review of the literature focusing on the contextualization of the research problem; Section IV presents the descriptions and algorithms of the proposed heuristics; Section V presents the specification, mode of comparison and analysis of the computational experiments; and, finally, Section VI presents the conclusions and suggestions for future work.

II. PRELIMINARY CONCEPTS

A. Graph Theory

Consider a simple graph \( G = (V(G), E(G)) \), such that \( V(G) = \{v_1, \ldots, v_n\} \) is a finite set whose elements are called vertices, and \( E(G) \subseteq \{\{v_i, v_j\} : v_i, v_j \in V(G), \ i \neq j\} \) is a finite set whose elements are 2-elements subsets of \( V(G) \) and are called edges, such that \( |V(G)| = n \) and \( |E(G)| = m \). If \( V(H) \subseteq V(G) \) and \( E(H) \subseteq E(G) \), then \( H = (V(H), E(H)) \) is a subgraph of \( G \) and \( G \) is a supergraph of \( H \). If \( H \subseteq G \). In particular, if \( E(H) = \{\{v_i, v_j\} : v_i, v_j \in V(H)\} \cup \{\{v_i, v_j\} \in E(G)\} \), then \( H \) is an induced subgraph of \( G \). For non-adjacent
vertices $v_i$ and $v_j$, $Y_{ij} = G + \{v_i, v_j\}$ is a supergraph of $G$ obtained by the insertion of an edge between $v_i$ and $v_j$. The degree of a vertex $v_i \in V$, $d(v_i)$, is the number of edges that are incident to it. If $d(v_i) = 1$, then $v_i$ is called a pendant vertex.

A path is a graph $P$ such that $V(P) = \{v_0, v_1, \ldots, v_n\}$ and $E(P) = \{\{v_0, v_1\}, \{v_1, v_2\}, \ldots, \{v_{n-1}, v_n\}\}$. In particular, two paths $P$ and $Q$ are vertex- or edge-independent if $V(P) \cap V(Q) = \emptyset$ or $E(P) \cap E(Q) = \emptyset$, respectively. The length of a path is its number of edges. The distance between two vertices $v_i$ and $v_j$, $d_G(v_i, v_j)$, is the number of edges in a shortest path between them, if such path exists. Otherwise, $d_G(v_i, v_j) = \infty$. The diameter of $G$, $\text{diam}(G)$, is the maximum distance between all pairs of vertices in $G$. If $|V| = 1$ or if there is at least one path for every pair of vertices in $V$, then $G$ is connected. Otherwise, $G$ is disconnected.

The vertex connectivity of $G$, $\kappa(G)$, is the smallest number of vertex-independent paths between all pairs of vertices in $V$. If $\kappa(G) = k$, then $G$ is $k$-connected. A $r$-vertex cut set of $G$ is a set of $r$ vertices whose removal makes $G$ disconnected. If there is no $k$-vertex cut set for $0 < k < r$, then the $r$-vertex cut set is the minimum vertex cut set.

The betweenness centrality of a vertex $v_i \in V$, $g(v_i)$, is the ratio between the number of shortest paths between all pairs of vertices other than $v_i$ that contains $v_i$ and the total number of shortest paths between all pairs of distinct vertices.

The Laplacian matrix of $G$, $L(G) = [l_{ij}]$, defined as follows:

$$l_{ij} = \begin{cases} d(v_i), & \text{if } i = j \\ -1, & \text{if } \{v_i, v_j\} \in E(G) \\ 0, & \text{otherwise} \end{cases}$$

The second smallest eigenvalue of $L(G)$, $\alpha(G)$, is called algebraic connectivity of $G$. An eigenvector $v_2(G)$ of $L(G)$ associated to $\alpha(G)$ is a Fiedler vector. The Fiedler distance between two vertices $v_i$ and $v_j$ of $G$ is the absolute difference between the $i$-th and $j$-th coordinate of the Fiedler vector, $d_{v_2(G)}(v_i, v_j) = |(v_2(G))_i - (v_2(G))_j|$.

### B. Reliability

Let $G$ be a connected graph such that each vertex or edge is subject to removal with independent probability $1 - p$, $p \in [0,1]$. The reliability of $G$ is a polynomial function that maps $(G,p)$ to the probability that $G$ remains connected after the removal of a random number of its vertices or edges. Since Moore & Shannon, the study of the reliability of a graph has focussed on the edge reliability, $R_E(G,p)$, which assumes reliable vertices and unreliable edges [5]. Thus, a minor part of the literature turned to the study of the vertex reliability, $R_V(G,p)$, which assumes reliable edges and unreliable vertices [6].

In both cases, determining the reliability of $G$ belongs to the class of NP-Complete problems; because the problem of counting the number of induced connected subgraphs of a given graph is NP-Complete [3]. Specifically, the vertex reliability polynomial of $G$ is given by

$$R_N(G,p) = \sum_{r=1}^{n} S_r(G)p^r (1-p)^{n-r},$$

with

$$S_r(G) + C_{n-r}(G) = \binom{n}{r},$$

where $S_r(G)$ is the number of connected induced subgraphs of $G$ with exactly $r$ vertices, and $C_{n-r}(G)$ is the number of $(n-r)$-vertex cut sets of $G$ [7]. The Fig. 1 shows the graph $C_5$, one of its supergraphs obtained by inserting the edge $\{v_3, v_4\}$, $Y_{34}$, and the vertex reliability of $C_5$ and $Y_{34}$ given $p = 0.9$.

![Fig. 1. Graph $C_5$, supergraph $Y_{34}$. $R_N(G,0.9) = 0.97488$.

Suppose $G \in \Omega(n,m)$, the set of all connected graphs with $n$ vertices and $m$ edges. Then, $G$ is: (a) uniformly best if, and only if, $R_N(G,p) \geq R_N(H,p)$, $\forall p \in [0,1], \forall H \in \Omega(n,m)$; (b) locally best if, and only if, $\exists p \in [0,1]: R_N(G,p) \geq R_N(H,p)$, $\forall H \in \Omega(n,m)$; (c) $\kappa$-optimal if, and only if, $\kappa(G) \geq \kappa(H)$, $\forall H \in \Omega(n,m)$; and (d) $S_r$-optimal, for $1 \leq r \leq n$, if and only if, $S_r(G) \geq S_r(H)$, $\forall H \in \Omega(n,m)$ [6]. Additionally, if $G$ is uniformly best, then it is $\kappa$-optimal; and if $G$ is $S_r$-optimal for all $r \in \{1, \ldots, n\}$, then it is uniformly best [7].

However, a uniformly best graph does not always exist. There are at least two locally best graphs in those cases, one of which is necessarily kappa-optimal [6]. It is conjectured that a graph is at least locally best if it is $S_r$-optimal for a sufficient unknown number of $r \in \{1, \ldots, n\}$ [6]. Trivially, every $G \in \Omega(n,m)$ is $S_1$-optimal, $S_2$-optimal, and $S_n$-optimal. Moreover, by Eq. (2) and Menger's Theorem, a $\kappa$-optimal graph $G$ such that $\kappa(G) = k$ is $S_{n-k}$-optimal [8]. It is reasonable to suggest that if $H \in \Omega(n,m)$ is $S_r$-optimal in more values of $r$ than a $\kappa$-optimal graph $G$, then $H$ is a locally best graph.

### III. Literature Review

Shier identifies two main areas of research related to reliability in graphs [9]. The first one is analysis, which involves calculating the reliability of a graph, analysing its polynomial function, and finding bounds for the values it can assume. The second one is design, which involves constructing or changing a graph to achieve high reliability, usually through topological optimization and heuristics for insertion or rearrangement. This separation is not rigid despite the distinction, and many works deal with analysis and design. In any case,
Colbourn emphasizes that both are primarily concerned with connectivity and its limitations due to topology. \[10\].

From the analysis perspective, part of the literature turned to the simulation of states of a graph to estimate its main topological properties \([11]\). The primary tool within this subarea is the Monte Carlo simulation, for which several specific techniques have been developed \([12], [13]\). Although these simulation techniques can produce satisfactory approximates to various reliability-related problems, the accuracy limitation inherent to simulations makes them an unfit tool to compare the effect of a single edge insertion on the reliability of an arbitrary graph \([14]\). The difference in reliability between single edge insertion may be smaller than what the simulations can identify.

From the design perspective, part of the literature tries to mitigate the difficulty of reliability by restricting its research to graphs with specific structures (e.g., almost regular complete multipartite graphs), using specific types of reliability (e.g., s-t reliability), or limiting the types and quantities of allowed changes to a graph (e.g., rearranging at most 1 edge). In effect, design is well-placed to develop heuristics for constructing or changing graphs.

These heuristics produce solutions that don’t guarantee optimality in arbitrary graphs but are sufficiently good under the assumed constraints. That is, it is possible to stipulate the criteria of what constitutes a satisfactory solution, such as always being above average in comparison to arbitrary feasible solutions or by the proximity of the optimal in cases where it is possible to compare them with the optimum.

Although there are different heuristics for distinct problems related to connectivity in general, most of them use similar algorithms based on the topological properties of graphs \([15], [16]\). In those cases, the difference usually involves adapting which graph property is considered (e.g., vertex connectivity instead of edge connectivity) and weighting the quality of the solutions by their computational cost.

However, there are some peculiarities about vertex reliability that add difficulty to the investigation when compared to edge reliability. Notably among these: \(a\) \(R_N(G, p)\) is non-monotonic while \(R_E(G, p)\) is monotonic; \(b\) vertex connectivity and the enumeration of connected induced subgraphs are used instead of edge connectivity and the enumeration of spanning trees; and \(c\) most reduction techniques in the literature assume perfectly reliable nodes, such as deletion–contraction by the factoring theorem \([17]–[20]\). Thus, it is necessary to have different heuristics or ensure that adaptations take into account these peculiarities and test their performance in the context of vertex reliability.

The article addresses the problem of deciding which edge insertion leads to the greatest increase in vertex reliability on an arbitrary connected graph without pendant vertices. More specifically, we adapted and evaluated for vertex reliability six heuristics in the literature on edge reliability and other robustness measures. Then, based on the performance results and theoretical considerations, we focus our attention on the proposed \(\alpha\) and \(\varphi\) heuristics and how their behaviours suggest renewed approaches to spectral heuristics and reliability \([4]\).

IV. Method

The heuristics evaluated in this work are divided between basic heuristics, named by lowercase letters, and derivative heuristics, named by uppercase letters. In all cases, a heuristic applied to a graph produces a solution in the form of single edge insertions to increase the vertex reliability of this graph.

The basic heuristics don’t have an explicit tiebreak criterion with regard to the multiplicity of single edge insertions produced as a solution for a graph. That is, a basic heuristic applied to \(G\) can produce more than one single edge insertion, and such insertions are considered equivalent to each other. In turn, each derivative heuristic is a refinement of a basic heuristic by adding a tiebreak criterion to guarantee the unicity of the output.

The \(\alpha\) and \(\varphi\) heuristics we propose in this work are described by their: \(a\) theoretical motivation; and \(b\) algorithm, that is, the steps and calculations prescribed by the heuristic. All the other heuristics are briefly addressed in Section IV-B.

A. Proposed Heuristics

1) \(\alpha\) Heuristic: The theoretical motivation for the \(\alpha\) heuristic is the maximization of the algebraic connectivity of \(G\). Consider \(Y = \{Y_{ij} = G + \{i, j\} : G \in \Omega(n, m), v_i \neq v_j\}\), the set of all supergraphs of \(G\) generated by a single edge insertion. Since \(Y \subset \Omega(n, m + 1)\), it is necessary that if \(Y \in Y\) is locally best in \(\Omega(n, m + 1)\), then it must have the maximum vertex connectivity of all graphs in \(Y\). Thus, the optimal increase in vertex connectivity is a necessary but insufficient criterion for a supergraph to be locally best. Moreover, given that \(\alpha(G) \leq \kappa(G)\) for \(G\) with \(n > 1\) vertices, then the algebraic connectivity is positively correlated with the vertex connectivity \([21]\). Therefore, it is reasonable to ponder if a heuristic that increases the algebraic connectivity might also consistently increase the vertex reliability.

The algorithm for the \(\alpha\) heuristic involves calculating \(\alpha(Y_{ij})\) for every possible edge insertion \(\{v_i, v_j\}\) of \(G\), sorting these values, and choosing the edge insertions associated with the maximum values. Note that if \(G\) is a connected graph with \(n\) vertices, it necessarily has at least \((n - 1)\) edges, and at most \(\binom{n}{2}\) edges. Thus, the minimum number of edge insertions is equal to 0, and the maximum number is \(\binom{n}{2} - n + 1\). The calculation of the eigenvalues of a single \(n \times n\) matrix belongs to \(O(n^2)\) \([22]\). Therefore, calculating the algebraic connectivity for every supergraph of a sparse graph is computationally expensive.

2) \(\varphi\) Heuristic: Given that the \(\alpha\) heuristic could be too costly if it calculated the algebraic connectivity for every supergraph of \(G\), the literature suggests the use of a heuristic that maximizes the algebraic connectivity instead of the exact method described in Section IV-A1 \([4]\). The Fiedler distance of an edge insertion is a theoretically sound and reasonable proxy for increases in algebraic connectivity, due to the generally positive correlation between them in single edge insertions and \(\alpha(G) \leq \alpha(Y_{ij}) \leq (d_{\varphi}(G)(v_i, v_j))^2 + \alpha(G)\) \([23]\). It is in this context that the \(\varphi\) heuristic first appeared \([24]\).

It must be noted that too many steps were made while searching for a heuristic for vertex reliability. Initially, a good
candidate was identified – vertex connectivity. However, it is computationally expensive. The algebraic connectivity is a good proxy for it, though, but also too computationally expensive. Finally, Fiedler distance is a good proxy for the algebraic connectivity, so it is possible to wonder if a proxy (Fiedler distance) for a proxy (algebraic connectivity) of a proxy (vertex connectivity) is, in fact, a good heuristic for vertex reliability. Surprisingly, it is.

The correlation between Fiedler distance in single edge insertions and increase in algebraic connectivity in an arbitrary graph depends on its topology. Although the correlation is generally positive, it shows considerable variability when different classes of random graphs with distinct topological properties are considered. In the worst-case scenario described in the literature, the correlation drops to 0 in highly regular $D$-lattices \[24\]. If $\varphi$ is a good heuristic due to the fact that it is an excellent proxy for algebraic connectivity, then in those cases it would perform poorly.

We propose a theoretical motivation for $\varphi$ heuristic based on Fiedler’s Theorem, by which the number of connected induced subgraphs of $G$ increases as the maximum Fiedler distance between its pairs of vertices decreases \[21\]. In Section I-B we postulated that a graph that is $S_\gamma$-optimal in more values of $r$ than a $\kappa$-optimal graph is locally best. Assuming that this is true, having the greatest increase of connected induced subgraphs is a sufficient but unnecessary criterion for a supergraph to be locally best. Assuming that this is true, having the greatest increase of connected induced subgraphs is a sufficient but unnecessary criterion for a supergraph to be locally best.

In addition, Fiedler distance has two important properties:
(a) if $\nu_2(G)$ is a unit Fiedler vector, then $d_{\nu_2(G)}(v_i,v_j) \leq \sqrt{2}$; and (b) $d_{\nu_2(v_{ij})}(v_i,v_j) \leq d_{\nu_2(G)}(v_i,v_j)$. The first property ensures that Fiedler distance is limited; it is bounded between 0 (lower bound) and $\sqrt{2}$ (upper bound) for an arbitrary $G$ of order $n$. The second ensures that the Fiedler distance between two non-adjacent vertices is always greater than or equal to the Fiedler distance between these same vertices on any supergraph that contain an edge incident to these two vertices.

The algorithm for the $\varphi$ heuristic involves calculating the eigenvalues and eigenvectors of $L(G)$, selecting the Fiedler vectors associated with $\alpha(G)$, and then finding the absolute difference between every pair of non-adjacent vertices. Algorithms for calculating the eigenvalues and eigenvectors of a square matrix are $O(n^3)$ \[22\]. If the geometric multiplicity of $\alpha(G)$ is greater than 1, then each pair of vertices has its Fiedler distance evaluated in all Fiedler vectors.

As a derivative heuristic of $\varphi$, we also propose the heuristic $\Phi$ which presents the following tiebreak criterion: if there are multiple single edge insertions with the same Fiedler distance, then the chosen insertion will be the one that maximizes the algebraic connectivity of its supergraph.

B. Other Heuristics

1) $\beta$ Heuristic: The $\beta$ heuristic inserts a single edge between non-adjacent vertices with minimum betweenness centrality. Its theoretical motivation is to reduce the mean betweenness centrality of the vertices of $G$ \[25\]–\[27\]. Given that a relatively high betweenness centrality of a vertex $v_i$ is positively correlated with the probability that $v_i$ belongs to a minimal vertex cut set, it is reasonable to ponder whether a heuristic that minimizes the mean betweenness centrality may also increase the vertex reliability. As a derivative heuristic of $\beta$, we postulate a $\beta$ heuristic. This heuristic is ideal, as it has no operational aspect, and is constructed post hoc as the heuristic that indicates the insertion with the best performance among the results of $\beta$ heuristic.

2) $\gamma$ Heuristic: The $\gamma$ heuristic inserts a single edge between non-adjacent vertices with the minimum degree. Its theoretical motivation is to reduce the average path length of $G$, which is a common measure of the efficiency of a network \[25\]–\[28\]. Given that the average path length is negatively correlated with connectivity, it is reasonable to suppose whether a heuristic that minimizes the average path length may also increase vertex reliability. As a derivative heuristic of $\gamma$, we postulate $\Gamma$ heuristic. This heuristic is ideal, and is constructed post hoc as the heuristic that indicates the insertion with the best performance among the results of $\gamma$ heuristic.

3) $\delta$ Heuristic: The $\delta$ heuristic inserts a single edge between non-adjacent vertices, one of which has the maximum degree and the other with whom the first has the greatest geodesic distance. Its theoretical motivation is to reduce the diameter of the graph. The diameter is an inverse measure of the efficiency of a network, and it is related with the probability of a graph getting disconnected after the targeted removal of some of its vertices, especially in cases of graphs with non-homogeneous degree distributions \[29\]. Given the similarity of this measure and reliability, it is reasonable to question whether a heuristic that minimizes the diameter may also increase vertex reliability.

4) $r$ Heuristic: The $r$ heuristic inserts a single edge between non-adjacent vertices randomly. In the literature, it is referred to as Random Choice. It is generally used as a reference or basis of comparison for the other heuristics \[15\].

V. Computational Experiments

A. Dataset

The generation, storage and analysis of the graphs were developed in Julia v1.7.1 with the following packages and their respective versions: Combinatorics.jl v1.0.2; DataFrames,jl v1.2.2; GraphIO.jl v0.6.0; Graphs.jl v1.4.1; Polynomials.jl v2.0.17; and SortingAlgorithms.jl v1.0.1. All of them are Free and Open Source Software. The source code for the computational experiments is available at GitHub \[30\]. The hardware used was an Intel i7-7700K CPU with 32GB of RAM running a GNU/Linux 64-bit operating system.

In order to ensure a robust evaluation of the heuristics, a dataset was created with random connected graphs without pendant vertices of order $n \in \{10, \ldots, 20\}$. For each order $n$, 2000 graphs were generated: 1000 Erdős-Rényi (ER) graphs, 500 Barabási-Albert (BA) graphs, and 500 Watts-Strogatz (WS) graphs. In total, the dataset has 22000 random graphs. Regardless of the model used to generate the random graph, each graph was tested to verify that it met the topological requirements and ensure no isomorphic graphs in the dataset.
B. Comparison Criterion

Consider \( H = \{h_1, \ldots, h_c\} \), a set of \( c \) heuristics, \( G = \{G_1, \ldots, G_N\} \), a set of \( N \) connected graphs, and \( 1 \leq k \leq c \).

The relative deviation index (RDI) is used to evaluate the performance of the heuristics \([31], [32]\). The definite integral of the vertex reliability polynomial \( R_N(G, p) \) in the interval \([0,1]\) is selected as its score function, denoted by \( F \). Thus, for each supergraph \( Y_{ij} \) produced by an edge insertion of the \( k \)-th heuristic, there is an associated score value \( F(Y_{ij}) \). The higher the score, the more reliable a supergraph is when \( p \in [0,1] \).

However, as some heuristics can produce more than one insertion of edge as a solution, the score is not exactly from the heuristic, but of an insertion and its associated supergraph. For each \( G \), we define \( F_B \) and \( F_W \), respectively, as the maximum and minimum scores obtained by the supergraphs produced from the edge insertions of all the heuristics in \( H \). Let the \( t \)-th insertion of the \( k \)-th heuristic be denoted by \( h_k^{(t)} \), where \( l_k \) is the number of insertions produced by \( h_k \) such that \( 1 \leq t \leq l_k \). Thus, the RDI of \( h_k^{(t)} \) in \( G \) is given by

\[
RDI_{G,h_k^{(t)}} = \begin{cases} 
\frac{F_B - F(G + h_k^{(t)})}{F_B - F_W}, & \text{if } F_B - F_W \neq 0 \\
0, & \text{if } F_B - F_W = 0
\end{cases},
\]

and the RDI of \( h_k \) in \( G \) is calculated as the arithmetic mean of the RDI of its associated edge insertions:

\[
RDI_G,h_k = \frac{\sum_{t=1}^{l_k} RDI_{G,h_k^{(t)}}}{l_k}.
\]

The mean relative deviation index (MRDI) compares heuristics between graphs in \( G \). The MRDI of \( h_k \) is defined as

\[
MRDI_G,h_k = \frac{\sum_{k=1}^{N} RDI_{G,h_k}}{N},
\]

where the heuristic with minimum MRDI is, on average, closer to the highest scores.

C. Procedures

Therefore, to evaluate the heuristics, the following steps are considered:

1) Let \( G \) and \( H \), sets of \( N \) connected graphs and \( c \) heuristics. Then, for each \( G \in G \):
   a) Apply \( h_k \) on \( G \) and tag as \( h_k^{(1)}, \ldots, h_k^{(l_k)} \) the \( l_k \) edge insertions produced by \( h_k \), for \( 1 \leq k \leq c \).
   b) Store all supergraphs \( G + h_k^{(t)} \), for \( 1 \leq k \leq c \) and \( 1 \leq t \leq l_k \), produced by all the heuristics in \( H \).
   c) Enumerate all the connected induced subgraphs of \( G, G + h_1^{(1)}, \ldots, G + h_c^{(l_c)} \).
   d) Calculate the vertex reliability polynomial \( R_N(G, p), R_N(G + h_1^{(1)}, p), \ldots, R_N(G + h_c^{(l_c)}, p) \) as a function of \( p \).
   e) Calculate the RDI for each heuristic in \( H \) for \( G \).
2) Calculate the MRDI for each heuristic in \( H \) for \( G \) and set as best the heuristic with the lowest MRDI.

The power set of \( V, \mathcal{P}(V) \), lists all possible \( 2^n - 1 \) induced subgraphs of \( G \). Let \( S \neq \emptyset \) be an element of \( \mathcal{P}(V) \). Given that \( G \) is connected, then any supergraph \( Y_{ij} \) of \( G \) is necessarily connected. Therefore, for any \( S, a \) connected subgraph induced by \( S \) in \( G \) is also connected in \( Y_{ij} \), regardless of the edge insertion. Thus, none of these subgraphs need to be tested more than once. The number of connected induced subgraphs can only increase for the supergraphs of \( G \) if the subgraph induced by \( S \) is disconnected in \( G \) and \( \{v_i, v_j\} \not\subset S \), then it is necessarily disconnected in \( Y_{ij} \). That is, an edge insertion between vertices that are not in \( S \) does not change the connectivity of the induced subgraph. Therefore, it is only necessary to test the disconnected induced subgraphs that contain both \( v_i \) and \( v_j \) for each \( Y_{ij} \). Once all disconnected induced subgraphs of \( G \) that become connected induced subgraphs in \( Y_{ij} \) are counted, we obtain the vertex reliability polynomial for \( Y_{ij} \).

The RDI of each basic heuristic in \( H \) for each \( G \) is known at the end of the procedure. With this data, one can construct the derivative heuristics, calculate the MRDI, and then analyse the performance of all the heuristics.

For the derivative heuristics, we applied an approximate one-sided Wilcoxon signed-rank test with Bonferroni correction for multiple comparisons and point-biserial correlation coefficient \( r \) for the effect size. The null hypothesis is that the median of pairwise differences is less than or equal to zero; the alternative hypothesis is that the median of pairwise differences is greater than zero. It is possible to identify whether the distributions of the scores between heuristics are significantly distinct and if there is any stochastically dominant heuristic through these tests \([33]\). Two of the derivative heuristics are post hoc, so their comparison with \( \Phi \) is quite conservative, favoring the null hypothesis.

D. Results

For conciseness, the analysis presented here is limited to the case in which all graphs are aggregated by their model – ER, BA, or WS – and order. The analysis of the disaggregated data exhibited similar results in terms of performance.

The quality of a heuristic depends not only on its performance but also on its computational cost. Table I presents an empirical evaluation of the computational time (in milliseconds) of heuristics in \( H \) for two sets of graphs with orders 15 and 20, with each 1000 ER random graphs.

| \# | Heuristic | Mean | Median | Mean ± SD |
|---|-----------|------|--------|-----------|
| 15 | \( \alpha \) | 2.809 | 0.885 | 0.301 ± 0.291 |
| \( \beta \) | 3.434 | 0.077 | 0.079 ± 0.001 |
| \( \gamma \) | 0.015 | 0.047 | 0.019 ± 0.001 |
| \( \delta \) | 0.039 | 0.929 | 0.070 ± 0.226 |
| \( \varphi \) | 2.816 | 0.107 | 0.110 ± 0.087 |
| \( r \) | 0.010 | 0.020 | 0.012 ± 0.001 |
| \( \phi \) | 2.778 | 0.125 | 0.133 ± 0.120 |
| 20 | \( \alpha \) | 4.770 | 2.274 | 2.398 ± 0.509 |
| \( \beta \) | 3.397 | 0.120 | 0.137 ± 0.176 |
| \( \gamma \) | 0.022 | 0.053 | 0.034 ± 0.006 |
| \( \delta \) | 2.346 | 0.056 | 0.083 ± 0.143 |
| \( r \) | 0.142 | 0.175 | 0.189 ± 0.123 |
| \( \phi \) | 2.556 | 0.201 | 0.218 ± 0.146 |

Table II shows the following summary measures for each heuristic: (a) Insertions is the total number of insertions produced by the heuristic; (b) Best (Single) is a pair of values,
where the first is the number of insertions produced by the heuristics that have reached the highest vertex reliability value among all insertions indicated by the set $H$ for the same graph, and the second (in parentheses) is the number of graphs in which the heuristic produced at least one insertion that obtained one of these higher values; (c) MRDI is the average relative deviation index of the heuristic; and (d) SD RDI is the standard deviation of the RDI of the insertions produced by the heuristic. In addition, cells are coloured using a sequence of colours ordered from red to blue ($\bullet$), where the values correspond to a rank classification for each column. A heuristic with the best performance has most of its values in the darkest blue ($\bullet$), especially those related to MRDI.

### Table II

**Performance of the Edge-Insertion Heuristics**

| Heuristic | Insertions | Best (Unique) | MRDI | SD RDI |
|-----------|------------|---------------|------|--------|
| $\alpha$  | 23615      | 11959 (11225) | 0.115049 | 0.194493 |
| $\beta$   | 53815      | 6569 (6216)  | 0.515709 | 0.359231 |
| $\gamma$  | 74466      | 14062 (13447)| 0.311627 | 0.287609 |
| $\delta$  | 66868      | 5545 (5198)  | 0.649908 | 0.297609 |
| $\varphi$ | 23955      | 14870 (14299)| 0.052739 | 0.105383 |
| $\phi$    | 22000      | 734 (743)    | 0.070052 | 0.309186 |
| $B$       | 22000      | 6216 (6216)  | 0.200597 | 0.299210 |
| $\Gamma$  | 22000      | 13447 (13447)| 0.093376 | 0.194638 |
| $\Phi$    | 22000      | 14299 (14299)| 0.042533 | 0.058820 |

In addition, the distribution of the RDI of the insertions produced by each heuristic can be seen in Fig. 2. Some elements of the boxplots are not visible in some heuristics because they are flattened around 0.

![Fig. 2. Boxplot of the edge-insertion heuristics’ RDI](image)

The $\varphi$ heuristic presented the best performance regarding MRDI among all basic heuristics in the dataset of graphs. It also has the lowest standard deviation of the RDI, which is a desirable property. Accordingly, $\varphi$ is comparatively better than the other basic heuristics tested by the criterion we have established.

In the case of the derivative heuristics, it can be seen that $B$ and $\Gamma$ present lower values of MRDI than their respective basic heuristics, as expected. In turn, $\Phi$ exhibits a modest improvement over MRDI and SD RDI. Nevertheless, the number of graphs in which $\Phi$ produced the best result had a slight reduction compared to its basic heuristic. The results of the Wilcoxon signed-rank tests for each pair of derivative heuristics are shown in Table III with the following columns: (a) $r$ is the effect size given by the point-biserial correlation coefficient; (b) $T^*$ is the standardized test statistic; and (c) $p$-value is the probability of a value as extreme as the observed given the null hypothesis. These results provide sufficient evidence for $\Phi$ to be ranked as the best heuristic among the derivative heuristics.

### Table III (Approximated) Wilcoxon Signed-Rank Tests for the Heuristics

| Heuristics | $r$ | $T^*$ | $p$-value |
|------------|-----|-------|-----------|
| $(\Phi, B)$| 0.9618 | 94.3794 | $< 0.0001$ |
| $(\Phi, \Gamma)$ | 0.6760 | 28.4584 | $< 0.0001$ |
| $(\Gamma, B)$ | 0.8855 | 76.1101 | $< 0.0001$ |

### VI. Conclusion

In this paper, we proposed an adaptation of a spectral heuristic for the problem of vertex reliability of a single edge insertion. We named the basic adaptation as $\varphi$ and its refinement as $\Phi$. Both were compared in a set of random graphs with other heuristics, some from the literature, while others are purely idealized.

Following the defined comparison criterion, the heuristics proposed have had the best performance. The heuristic $\Phi$ has a higher computational cost than its basic form. However, its most costly computation is the tiebreak criterion. In practice, the computational time is expected to be, on average, very close. Therefore, it is reasonable to conclude that $\Phi$ provides a good enough answer to the research problem presented here and should be used in instances where it is feasible.

For future works, the cases where $\varphi$ outperforms $\Phi$ could be analysed. If this behavior is related to some topological properties of the graphs, then $\Phi$ can be further improved with minor adjustments. There are also good reasons to believe that it is possible to improve the performance of $\Phi$ by identifying other tiebreak criteria for the multiplicity of edge insertions. Alternatives may involve, for example, calculating the Fiedler vectors of the supergraphs resulting from the insertions of edges to identify with more accurately, via Fiedler’s Theorem, the supergraph with the largest number of connected induced subgraphs. Considering the theoretical motivation developed here for the $\varphi$ heuristic, is desirable to explore more properties of the Fiedler vector in the context of reliability in general.

### References

[1] S. G. Ratings, “Industry top trends 2020: Telecommunications,” 2019.
[2] M. Newman, *Networks*. Oxford Univeristy Press, 1st edição ed., 2010.
[3] K. Sutter, A. Satyanarayana, and C. Sulfet, “The complexity of the residual node connectedness reliability problem,” SIAM J. Comput., vol. 20, pp. 149–155, 02 1991.

[4] B. Barreto and C. Oliveira, “Relações entre o aumento da conectividade algébrica e da confiabilidade de redes utilizando o vetor de fiedler como estratégia de inserção de arestas,” in Anais do XLIX - Simpósio Brasileiro de Pesquisa Operacional, pp. 3338–3349, 2017.

[5] E. F. Moore and C. E. Shannon, “Reliable circuits using less reliable relays,” Journal of the Franklin Institute, vol. 262, no. 3, pp. 191–208, 1956.

[6] O. Goldschmidt, P. Jaillet, and R. LaSota, “On reliability of graphs with node failures,” Networks, vol. 24, pp. 251–259, 1994.

[7] S. Liu, K.-H. Cheng, and X. Liu, “Network reliability with node failures,” Networks, vol. 35, pp. 109–117, 03 2000.

[8] R. Diestel, Graph Theory, Graduate Texts in Mathematics 173, Springer-Verlag Berlin Heidelberg, 5ª edição ed., 2017.

[9] D. R. Shier, Graph Theory, pp. 661–666. Boston, MA: Springer US, 2013.

[10] C. J. Colbourn, “Combinatorial aspects of network reliability,” Annals of Operations Research, vol. 33, pp. 1–15, 1991.

[11] Y. Shpungin and I. B. Gertsbakh, Models of Network Reliability: Analysis, Combinatorics, and Monte Carlo, CRC Press, 1ª edição ed., 2012.

[12] M. Lomonosov, “On monte carlo estimates in network reliability,” Probability in the Engineering and Informational Sciences, vol. 8, no. 2, p. 245–264, 1994.

[13] I. B. Gertsbakh and Y. Shpungin, Using D-Spectra in Network Monte Carlo: Estimation of System Reliability and Component Importance, pp. 23–31. London: Springer London, 2012.

[14] G. Fishman, Monte Carlo: Concepts, Algorithms, and Applications, Springer Series in Operations Research and Financial Engineering, Springer New York, 2013.

[15] A. Beygelzimer, G. Grinstein, R. Linsker, and I. Rish, “Improving network robustness by edge modification,” Physica A: Statistical Mechanics and its Applications, vol. 357, no. 3, pp. 593–612, 2005.

[16] A. Sydney, C. Scoglio, and D. Gruenbacher, “Optimizing algebraic connectivity by edge rewiring,” Applied Mathematics and Computation, vol. 219, no. 10, pp. 5465–5479, 2013.

[17] O. Theologou and J. Carlier, “Factoring and reductions for networks with imperfect vertices,” IEEE Transactions on Reliability, vol. 40, no. 2, pp. 210–217, 1991.

[18] D. Torrieri, “Calculation of node-pair reliability in large networks with unreliable nodes,” IEEE Transactions on Reliability, vol. 43, no. 3, pp. 375–377, 1994.

[19] S.-Y. Kuo, F.-M. Yeh, and H.-Y. Lin, “Efficient and exact reliability evaluation for networks with imperfect vertices,” IEEE Transactions on Reliability, vol. 56, no. 2, pp. 288–300, 2007.

[20] J. L. Brown and L. Mol, “The shape of node reliability,” Discrete Applied Mathematics, vol. 238, pp. 41–55, 2018.

[21] M. Fiedler, “Algebraic connectivity of graphs,” Czechoslovak Mathematical Journal, vol. 23, no. 2, pp. 298–305, 1973.

[22] V. Y. Pan, Z. Q. Chen, and A. Zheng, “The complexity of the algebraic eigenproblem,” 1998.

[23] C. Godsil and G. Royle, Algebraic Graph Theory. Graduate Texts in Mathematics 207, Springer-Verlag New York, 1ª edição ed., 2001.

[24] H. Wang and P. Van Mieghem, “Algebraic connectivity optimization via link addition,” in Proceedings of the 3rd International Conference on Bio-Inspired Models of Network, Information and Computing Systems, BIONETICS ’08, (Brussels, BEL), ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2008.

[25] P. Holme, B. J. Kim, C. N. Yoon, and S. K. Han, “Attack vulnerability of complex networks,” Phys. Rev. E, vol. 65, p. 056109, May 2002.

[26] Z. Jiang, M. Liang, and D. Guo, “Enhancing network performance by edge addition,” International Journal of Modern Physics C, vol. 22, no. 11, pp. 1321–1326, 2011.

[27] X. Ji, B. Wang, D. Liu, G. Chen, F. Tang, D. Wei, and L. Tu, “Improving interdependent networks robustness by adding connectivity links,” Physica A: Statistical Mechanics and its Applications, vol. 444, pp. 9–19, 2016.

[28] R. Albert and A.-L. Barabási, “Statistical mechanics of complex networks,” Rev. Mod. Phys., vol. 74, pp. 47–97, Jan 2002.

[29] R. Albert, H. Jeong, and A.-L. Barabási, “Error and attack tolerance of complex networks,” Nature, vol. 406, p. 378–382, Jul 2000.

[30] F. M. Pinheiro Jr., “Github repository: Monografia,” https://github.com/fausto-mpj/monografia_grafos 2022. (Accessed: 02 March 2022).

[31] Y.-D. Kim, “Heuristics for flowshop scheduling problems minimizing mean tardiness,” The Journal of the Operational Research Society, vol. 44, no. 1, pp. 19–28, 1993.

[32] M. A. Moro, “Meta-heurísticas grasp e brkga aplicadas ao problema da diversidade máxima,” dissertação de mestrado, Universidade Estadual de Campinas, 2017.

[33] F. Wilcoxon, Individual Comparisons by Ranking Methods, pp. 196–202. New York, NY: Springer New York, 1992.