Analysis, Comparison and Evaluation of Robustness Metrics in Road Networks

Learning From Network Project - 26th January 2024

Lorenzo Riccò ✉ Maulik Sompura ✉ Mohammadali Jafari ✉

Abstract

The current work represents a deep analysis of the concept of robustness applied to networks and a simple implementation about possible metrics of interest. In particular, through an exhaustive review of the existing literature, we navigate the landscape of robustness, synthesizing the collective knowledge acquired to date. The vast expanse of robustness-related studies reveals a field rich in exploration but marked by a lack of definitive clarity. This analysis serves as a foundation for comprehending the concept of robustness and its related metrics, establishing a framework for future research and practical applications in domains such as the one analyzed in the study.

Keywords

Robustness, Metrics, Attacks, Road Networks.

1. Introduction

In the field of network science, a pivotal question is: ”given a network, is it robust for a given service?”. This question is, however, inherently challenging because determining the robustness of a network is no straightforward task. The mentioned problem is, in some way, ill-posed: while humans intuitively have a notion of robustness, it remains vague and difficult to quantify from a mathematical perspective. Over the years, numerous definitions and measures of robustness have emerged, reflecting a diverse range of approaches aimed at capturing the resilient properties of networks.

In this survey, we condense significant discoveries granting access to essential information by (2) defining the concept of robustness and analyzing the contextual framework in which we apply it; (3) conducting an analysis of the key metrics based on topological characteristics of the networks; (4) delving into a comprehensive examination of primary attacks, providing detailed insights into their nature and impact and (5) giving the reader a list of possible defensive techniques. Following this, (6) we highlight the methodology employed to implement a comprehensive analysis of these metrics. Through straightforward graphical representations, (7) we also assess and compare the various metrics under consideration, categorizing their behavior in response to different attacks.

2. Theoretical Framework

Robustness represent a critical and multifaceted quality and in an effort to model it is necessary to first provide a definition.

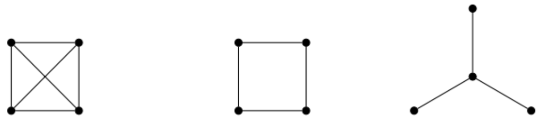
We can interpret robustness as *the capacity of a network to continue performing well when it is subject to failures or attacks.*

This is not the sole definition documented in the literature; however, it is most commonly employed. From this definition, one may intuitively infer that it is all about back-up possibilities, or alternative paths. So in order to decide whether a given network is robust, it is crucial to establish quantitative measures for this purpose. However, the scope extends beyond measurement alone. From a general perspective, there are three key objectives in the exploration of network robustness. The primary task involves developing metrics to effectively quantify network robustness. In the next section, we have summarized the classical robustness measures along with how each measure is linked to the evaluation of graph robustness. The second task centers on identifying mechanisms employed in network attacks, while the third task involves the construction of defensive techniques to resist network failures and recover from attacks.

In this context, the definition fits perfectly with the scenario we have posed our attention on: road networks. Road systems can be modeled as networks and investing the robustness of this complex interconnected systems, purposely designed to operate with minimal redundancy and high capacity in order to minimize costs, can help identify the streets and locations that most strongly affect the efficiency of the entire road network. Is the road system of a city efficient when a street is closed? How much can impact the flow a new street? How does the road network adapt to special events like protests or emergencies? These are some of the main question that can be considered in the moment in which we develop an analysis on the road systems, defined by a proper network.

3. Robustness Metrics

Following the above definition, we can say that networks become more robust when links are added, and a connection between two nodes is more robust when there is more than one path between them.



In the above examples of graphs with four vertices following our intuition the graphs are arranged from left to right in descending order of robustness. Given this assumption, we now perform a comprehensive comparison of the commonly used robustness metrics for graphs. In particular, We reviewed several graph robustness measures, each of which takes as input an undirected, unweighted graph $G = (V, E)$, where $|V| = n$ is the set of vertices, and $|E| = m$ is the set of edges. After describing each measure, we describe its link to the study of network robustness and its computational complexity in order to give the reader a full comprehension of the metrics under observation.

3.1. Connectivity

Binary Connectivity (κ) is a classical graph measure which determines whether or not a graph is connected (κ=1) or unconnected (κ=0) by examining simply whether all pairs of vertices have a connecting path: clearly, a graph is unconnected if at least one pair of vertices does not have a connecting path. In practical terms, binary connectivity is an inadequate metric for assessing robustness since it identify whether a network is disconnected or not. Apart from the classical definition, two more connectivity measures have been proposed: *Vertex Connectivity* (κ_v) and *Edge Connectivity* (κ_e). The first represents the minimal number of vertices that need to be removed to disconnect the graph, the second on the other hand indicates the number of edges that need to be removed to disconnect the graph. For an incomplete graph the relations is that $\kappa_v \leq \kappa_e \leq \delta_{min}$, where δ_{min} is the minimum degree of the vertices. For both this two measures, they naturally ties to graph robustness, since they increase as the graph becomes harder to disconnect. From a computational perspective binary connectivity ca be computed using breadth-first or depth-first search, starting at any vertex with time complexity $O(n+m)$; vertex connectivity can be reduced to a max-flow problem with a time complexity of $O(n^{8/3}m)$; and for edge connectivity the best known algorithm has complexity of order $O(mn)$.

3.2. Distance

Among the metrics based on distance we can consider the *Diameter* (d_{max}) and the *Average Distance* ($\overline{d}$). In order to define both of them, let the distance d_{ij} be the length (number of edges) of the shortest path between vertices i and j. The maximum distance over all these distances is actually called diameter and the average over all pairs, so the average distance, is denoted by

$$\overline{d} = \frac{2}{n(n-1)} \sum_{i=1}^{n} \sum_{j=i+1}^{n} d_{ij}$$

The rationale behind using the diameter and average distance as measures of robustness starts from the principle that shorter paths indicate greater robustness. In particular, for the diameter this occurs because as the network's diameter decreases (by adding edges) the longest shortest paths between distant vertices shorten, resulting in a more tightly interconnected network. The main disadvantage of this two metrics is that adding back-up paths is not considered: the average distance consistently decreases with the addition of edges, whereas the diameter may remain unchanged even with the inclusion of additional edges.

Up to these observations the average distance is commonly modified by computing the average inverse distance, also known as *Efficiency* (E) and defined by

$$E = \frac{2}{n(n-1)} \sum_{i=1}^{n} \sum_{j=i+1}^{n} \frac{1}{d_{ij}}$$

In terms of efficiency, a higher value indicates greater robustness since the reciprocals of path lengths are utilized. This measure offers the advantage of applicability to unconnected network, however it also shares the disadvantage of not considering alternative paths. In computational terms, to calculate the diameter of a network is required a time of complexity of $O(n^3)$ using the Floyd–Warshall algorithm; also, both the average distance and efficiency have a time complexity of $O(n^3)$ due to the all pairs shortest path computation.

3.3. Clustering

The presence of triangles is captured by the *Clustering Coefficient* (C), which compares the number of triangles to the number of connected triples, and provides an indication of how well nodes tend to cluster together. By definition, a triplet is three nodes connected by either two edges (open triplet) or three edges (closed triplet); where a closed triplet is called a triangle.

$$C = \frac{\text{closed tripltes}}{\text{closed triplets + open triplets}}$$

While it was initially tailored for social networks to validate the likelihood that two friends of an individual are also friends with each other, the clustering coefficient can be proposed to assess robustness in various network contexts. As the global clustering coefficient increases, we begin to see the formation of tight knit groups (due to an increased density of triangles) which creates redundant pathways between neighbors, and increases the overall robustness of the network. The time complexity for computing the global clustering coefficient is $O(nd^2_{max})$, where d^2_{max} is the size of the largest adjacency list across all vertices in the graph.

3.4. Betweenness

The betweenness metric quantifies the number of shortest paths between pairs of vertices, passing through a particular vertex or edge. In cases where multiple shortest paths exist between two vertices, each of these paths is equally weighted. To gauge the robustness of a network, we compute the *Average Vertex Betweenness* $(\bar{b}_v)$ and the *Average Edge Betweenness* $(\bar{b}_e)$. It turns out that the average vertex and edge betweenness are linear functions of the average distance $\bar{d}$. In fact, we have that,

$$\bar{b}_v = 1/2(n-1)(\bar{d}+1)$$
$$\bar{b}_e = \frac{n(n-1)}{2m}\bar{d}$$

Due to these linear relationships, when comparing the robustness of two graphs with the same number of vertices, both the average distance and the average vertex betweenness will consistently identify the same graph as the most robust. Average vertex betweenness is inherently linked to graph robustness as it quantifies the average network throughput of vertices. A smaller average suggests a more robust network as it indicates a more even distribution of shortest paths across nodes, rather than relying heavily on a few central nodes. Average edge betweenness has similar robustness properties to average vertex betweenness. The smaller the average the more robust the network, since the shortest paths are more evenly distributed across each edge. Computing the vertex betweenness centrality for a single node has a time complexity of $O(nm)$ using Brandes'algorithm. Therefore, the average vertex betweenness has a time complexity of $O(n^2m)$. The time complexity for average edge betweenness is the same.

3.5. Community Structure Analysis

Community structure analysis involves the identification of cohesive groups of nodes within a network, known as communities or modules. These communities signify clusters of nodes sharing common features or interactions, offering a nuanced understanding of the network's internal organization. A critical aspect of this analysis is *Modularity*, a metric that quantifies the strength of the network's division into distinct communities. The formula for modularity (Q) is given by:

$$Q = \frac{1}{2m}\sum_{ij}\left(A_{ij} - \frac{k_i k_j}{2m}\right)\delta(c_i, c_j)$$

Here, A_{ij} is the element of the adjacency matrix representing the connection between nodes i and j, k_i and k_j are the degrees of nodes i and j, m is the total number of edges in the network, and $\delta(c_i, c_j)$ is the Kronecker delta function, equal to 1 if nodes i and j are in the same community and 0 otherwise.

Higher modularity values indicate a more pronounced separation of nodes into meaningful modules, providing a quantitative measure of the network's structural significance. Modularity plays a crucial role in uncovering the inherent organizational patterns of networks. By assessing how well-defined and distinct the communities are within a network, modularity helps researchers grasp the network's internal structure.

Moreover, modularity is a powerful tool for improving network robustness. Networks with high modularity often exhibit clear community structures, and fortifying connections within communities can enhance the network's ability to withstand targeted attacks or random failures. Identifying and reinforcing the connections that contribute most significantly to modularity can serve as a strategy to improve the network's resilience, maintaining its functionality even in the face of disruptions.

3.6. Other Common Metrics

Several other measures have been proposed in the literature among the past decades. Among the most important we can find *Measures based on Adjacency Matrix Spectrum*, where the adjacency matrix is a common network representation often used when enumerate walks, and *Measures based on Laplacian Matrix Spectrum*, for which the Laplacian matrix is often used when a problem can be related to spanning trees or the incidence of vertices and edges.

4. Common Attacks

The structural robustness of networks subject to various attacks is an important research area in complex network studies and in this chapter we explore common types of attacks that target networks, ranging from traditional techniques to more sophisticated strategies. By comprehensively analyzing these attacks, we aim to provide insights into their mechanisms, potential impacts, and strategies for defense. A common point in literature is that the topological structure diversity is essential to network robustness: in particular, with a higher level of degree diversity, the network displays a higher level of tolerance to random failures but is more vulnerable to attacks targeting high-degree nodes. With random failures we refer to the spectrum of attacks characterized by an indiscriminate targeting of nodes or edges, which represent one of the most common and easily relatable scenario in real-world applications: in city road for example random attacks can be translated into malfunctions can be caused by events such as natural disasters, traffic accidents, congestion or regulations on specific roads. Overall, considering the degree distribution as a preliminary step in implementing graph robustness analysis can be an initial key indicator of significance.

4.1. Targeted Attacks

Adversaries have two main strategies for attacking a network: removing nodes and removing edges. The objective of the attacker is to identify and target nodes and edges crucial to the network's functionality and achieving this goal often involves assessing the centrality of nodes and edges within the network. In fact, degree and betweenness are two most widely used node importance characters in targeted attacks. The degree of a node is the number of edges linked to that specific node and an higher degree illustrates that the node contributes to increased accessibility in the local area. In this attack scenario, nodes in the network are ranked based on their degree. With a given bucket, the attacker systematically removes vertices, starting from the nodes with the highest degree. This approach aims to rapidly decrease the total number of edges in the network. A more strong strategy is to recalculate the degree distribution of nodes after the removal of each vertex. On the other hand, when we take into account betweenness the purpose is to destroy as many paths as possible. This attack simply ranks each node according to its betweenness centrality defined as the proportion of the number of the shortest paths between nodes that pass a particular node and the total number of the shortest paths in the network. As before, simulated successive attacks are adopted to destroy the network until it collapses.

Since this attack considers information from across the network, it is considered a global attack strategy, while the former which only considers local neighborhood when making a decision, is considered as a local attack.

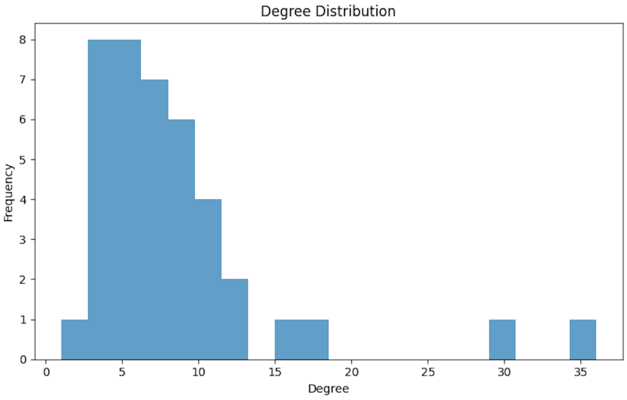
5. Defensive Techniques

The most devastating attacks often result from targeted rather than isolated or cascading failures. Consequently, defense mechanisms typically prioritize protecting networks from targeted attacks. As defenders, we often possess information about the graph topology and underlying degree distribution. This knowledge enables us to more accurately assess our network's robustness and point out potential vulnerabilities.

Without delving into specific strategies, methods to safeguard the network include: (1) adding edges, whether randomly or preferentially, to link nodes with low degrees; (2) rewiring edges by disconnecting a random edge from a high-degree node and reconnecting it to another node at random; and (3) identifying key nodes and edges within the network to monitor for suspicious activity. Given the prevalence of heterogeneous degree distributions in real-world networks, characterized by a small number of nodes with numerous links and a larger number of nodes with few links, we can use a combination of the three strategy in order to increase robustness's network.

6. Implementation

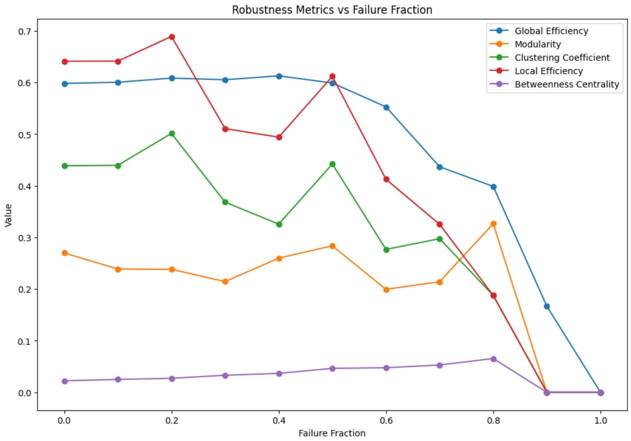
We attempted to develop our own personalized implementation to evaluate several of the different metrics discussed and validate the theoretical analysis cited in the previous chapters through NetworkX. Our analysis was conducted on a small graph to ensure reliable and comparable results. The network can be find at: https://networkrepository.com/road-chesapeake.php. The network considered has 39 nodes and 170 edges. The metrics considered in our implementation are: modularity, clustering coefficient, local and global efficiency, average betweenness centrality. As a preliminary step, we analyzed the degree distribution and continued observing it along with the other metrics considered during multiple iterations of a random attack with varying numbers of nodes and links targeted. The random attack considered in each iteration takes into account the original graph, and for every iteration, it removes a certain percentage of the graph's nodes and link.s



Having a heterogeneous degree distribution serves as a starting point for assessing the robustness of a graph. Two additional metrics used as preliminary steps in the analysis were vertex connectivity and edge connectivity. A high value relative to the initial number of nodes and edges indicates that it is more difficult to disconnect the graph, implying greater robustness of the network.

7. Results

The main challenge encountered during the implementation stemmed from the randomness of the attack. Each transition from a connected to a disconnected graph led to a significant drop in most of the considered metrics due to limitations in the implementation in NetworkX. In any case, by conducting multiple random attacks with different fractions considered, we were able to identify and recognize the trends followed by the various metrics under consideration.



As evident from the graph, all metrics exhibit a decreasing trend as the number of nodes and links removed from the graph increases, before eventually dropping to zero when the graph becomes disconnected. The sudden increase observed in some metrics is attributed to the fact that each random attack iteration starts from the original graph.

8. References

[1] W. Ellensa, R.E. Kooija"Graph measures and network robustness", Nov.2013.

[2] Scott Freitas, Diyi Yang, Srijan Kumar, Hanghang Tong, and Duen Horng Chau, "Graph vulnerability and Robustness: A Survey" in IEEE, March 2022.

[3] Yingying Duan, Feng Lu, "Robustness of city road networks at different granularities", ScienceDirect, Oct 2014.

[4] Swami Iyer,Timothy Killingback ,Bala Sundaram,Zhen Wang, "Attack Robustness and Centrality of Complex Networks," Apr. 2013.

[5] Jing LIU , Mingxing ZHOU, Shuai WANG, Penghui LIU, "A comparative study of network robustness measures," Springer, July 2017.

Background

Lastly, before mentioning the individual contributions, we see fit to first discuss briefly how the idea of the project came to us and how we reached to the work at hand. Before delving into the specifics of individual contributions, it is pertinent to provide a brief overview of the genesis of our project and the trajectory that led us to its current state. The impetus behind this work emerged from our profound interest in optimizing teleportation, specifically within the context of road networks. Recognizing the critical role of road transportation, particularly during crises, where a significant portion of our time is spent on roads, we embarked on this endeavor. Motivated by this perspective, we immersed ourselves in an extensive review of existing literature, with the ambitious goal of implementing robustness metrics for rigorous testing, execution, and comparison. Additionally, we aspired to contribute to the field by potentially refining or creating a novel metric. However, the realization quickly dawned that our initial objectives surpassed the time constraints of our allocated timeframe. Consequently, a strategic pivot was made, and our revised focus narrowed down to the development of a comprehensive survey on metric robustness. This survey serves as an instructive entry point, particularly tailored for individuals who are new to the field, providing foundational insights and understanding.

Individual Contributions

Mohammadali Jafari

Studied the literature and chose the subject - wrote the metrics for degree distribution, and the community structure analysis - wrote parts of the final report - helped in editing the project proposal, mid-term report, and the final report - studied different methods for analyzing graphs and consulted the group on what methods we should use - helped in studying how to get over the connectivity issue for networks after attack

Lorenzo Riccò

Studied the literature - wrote the majority of the project proposal, mid-term report, and the final report - studied different types of attacks and consulted the group on what types of attack we should use for the project - studied how to get over the connectivity issue for networks after attack - helped in writing the codes of attacks

Maulik Sompura

Studied the literature - wrote the sampling of the input graph - wrote the codes for visualizing the results - wrote the attacks and visualized them - studied different sampling methods and how we can improve the performance - helped in studying how to get over the connectivity issue for networks after attack.

Link GitHub to the Project Implementation

https://github.com/maulik225/Robustness-of-network/tree/main