Identifiers for structural warnings of malfunction in power grid networks

by
Miroslav Andjelković, Slobodan Maletić, Ivan Tomanović
VINČA Institute of Nuclear Sciences - National Institute of the Republic of Serbia
University of Belgrade, Belgrade, Mike Alasa 12-14, 11351 Vinča, Serbia

Abstract

Although its uninterrupted supply is essential for everyday life, the electricity occasionally experiences disruptions and outages. The work presented in the current paper aims to initiate the research to design a strategy based on advanced approaches of algebraic topology to prevent such malfunctions in a power grid network. Simplicial complexes are constructed to identify higher-order structures embedded in a network and, alongside a new algorithm for identifying delegates of the simplicial complex, are intended to pinpoint each element of the power grid network to its natural layer. Results of this methodology for analysis of a power grid network can single out its elements that are at risk to cause cascade problems which can result in unintentional islanding and blackouts. Further development of the outcomes of research can find implementation in the algorithms of the energy informatics research applications.

Key words: algebraic topology, simplicial complexes, complex systems, power grid network, power grid redundancy

1 Introduction

Prevention of damages and blackouts they cause put a challenge and efforts on the development of reliable methods in energy informatics research which applies thinking and skills of information systems to increase energy efficiency, stability, and sustainability, with the aim to provide reliable power grids with an uninterrupted supply, transitioning at the same time toward the smart power grids. In order to make electricity widely and reliably accessible, it is necessary to have an infrastructure resistant to severe disruptions or unintentional islanding - an event

*Corresponding author, e-mail: mandjelkovic@vinca.rs
where part completely separates from the rest of the power grid network and continue running. Nonetheless, power interruptions and power outages mostly occur due to equipment failures, hence causing the blackouts in most extreme scenarios. Furthermore, the malfunction and low reliability of the power grid indirectly leads to the fragility of other technological networks, which are highly dependent on it. Thus, disturbances in the functioning of the power grid network can have serious implications on everyday life, making all attempts to develop strategies to prevent, or when a problem happens to overcome, are of vital importance.

The work presented in current paper is motivated by the tendency to prevent a vulnerability of the power grid network and hence contribute the efficient energy management. In this context, the aim is to introduce a novel approach, based on tools originating from the topological analysis of data, to the study of failure of the power grid network and contribute to the development of prevention strategies. So far research in complex networks was mostly based on the methods originating from the graph theory [1], with some recent applications of the topological data analysis [2]. Topological analysis of data, as a set of tools, is drawing an attention of the wide variety of practitioners, from research scientists to industry. Although there are attempts in applications of algebraic topology, in particular persistent homology [3] on power grid networks [4], in this paper we are extending an interest in less studied structures which are characterized by the connectivity of mesostructures. Namely, unlike common topological approaches to datasets and complex networks, which are mostly based on the calculation of Betti numbers [3], that is higher-dimensional holes within the structure, here the objects of interest are nontrivial higher-dimensional chains of connectivity that are responsible for the system’s functioning.

Properties of power grids are well studied by applying complex network quantities, in particular with an aim of considering the effects caused by the power failures. For example, in order to prevent the worst scenarios that can lead to cascading failures in a power grid, one must understand the nature of grid vulnerability. A good insight into the current state of the grid and potential weaknesses can be identified through the use of different static and dynamic models to predict potential cascading failure scenarios in grids [5]. Similarly, Nesti et al. [6] approached the modeling of blackouts of different sizes through the use of graphs with heavy-tailed sinks that represented the power demand from users and created a model of Germany’s power grid that linked scale-free blackout sizes with scale-free city sizes. On the other hand, Carreras et al. [7] created a dynamical numerical model for blackouts in power transmission systems with minimal implementation and assumption of uniformity of the components throughout the system, with an idea to further expand it through the introduction of non-homogeneous components. European power grids are studied in [8], and modeled as the weighted graphs, with a focus on their topology and scalability. Self-organized criticality is studied by Zhao et al. [9] to ex-
plain the dynamic behavior and investigate the China power grid blackouts, targeting the small disturbances in the network that may trigger the chain reactions. Finally, Pagani and Aiello [10] reviewed the current state of the power grid studies performed by the statistical tools of complex network analysis in different states and regions over the world. The frequently studied grid infrastructure with these tools, at the time of review, was located in the United States, most of the European Union countries, China, and India.

As a case study, the US Power Grid network [11] is considered, where elements and their connections play different roles in the network's functioning. Due to the lack of information about the exact roles of nodes within the considered power grid, we have artificially selected the set of delegates which moves up to higher hierarchical levels. Namely, the goal is to reconstruct the latent hierarchical structure under selection criteria that distinguish node's participation in the formation of nontrivial structure. As the criteria for delegate nodes selection the so called topological dimension [12] is chosen and it quantifies the node's influence on their local higher-order neighborhoods. Higher-order structural properties of the US Power Grid network have already been considered in [13], where the calculated quantities displayed the statistical invariance. On the other hand, regarding the power grid vulnerability [14], research so far is characterized by an interest in a lack of robustness and resilience of a system, and was focused on considerations of random, as well as selective failures, through the percolation theory rooted in numerical simulation of different scenarios.

Results suggest that the underlying network has an internal structure of the most significant delegates, see 3, rather with higher-order aggregation than at the level of the initially obtained simplicial complex from a mathematical graph. A node that naturally originates on a lower hierarchical level can be found mixed within the same simplex with nodes that belong on a higher hierarchical level, and vice versa, presenting the potential problem, hence indicating roads towards future research. Generally, the outcomes of presented research contribute to the energy informatics research, as well as to building smart grids [15]. The proposed approach offers a broad framework for applications, not necessarily restricted to power grid networks, but also such as for example, the thermal storage systems [16].

2 Phenomenological background and its modeling

The underlying mathematical graph of the US Power Grid network is formed by nodes with different roles in the network's functioning. These various physical objects can be distributed on different levels in the hierarchical organization of a power grid network (see Fig. 1). Due to the authors' lack of knowledge on specific roles of nodes in the considered US Power Grid network, and with the aim to reconstruct the inherent hierarchical organization, the topological
criteria are applied in order to distinguish and select the influential vertices, termed delegates. Using selected delegate vertices in a clique complex, after a few iterations, different hierarchical levels are extracted. Hence, when disruption occurs, the scope of malfunction depends on the position of the level and the structure that is formed there.

Figure 1: Depiction of a segment of a layer-architecture of an arbitrary power grid network. Dashed circles display the same layer. Every type of node is found at the layer to which it naturally belongs or it may cause disruption in network stability and sustainability, so the proposed layer-architecture may be of great significance.
The notions of basic, and derived, quantities that will be introduced within this section are originating from the mathematical field of algebraic topology, more specifically, the Q-analysis \[17\], \[18\]. Let \( V = \{v_1, v_2, \ldots, v_k\} \) and \( S = \{s_1, s_2, \ldots, s_l\} \) be two sets assembled by \( k \) and \( l \) elements, respectively, and suppose that exists a binary relation \( \lambda \) that by some rule associate elements of sets \( V \) and \( S \). In other words, an element \( s_i \) from the set \( S \) is \( \lambda \)-related to some subset \( v_{j_0}, v_{j_1}, \ldots, v_{j_q} \) from the set \( V \), and vice versa, an element \( v_j \) is \( \lambda^{-1} \)-related to some subset \( s_{i_0}, s_{i_1}, \ldots, s_{i_p} \) from the set \( S \), where \( \lambda^{-1} \) is the inverse of relation \( \lambda \). Consequently, the set \( S \) and relation \( \lambda \) define a subset \( C \) of the power set \( P(V) \) of \( V \), that is \( C \subseteq P(V) \), and each element \( \{v_{j_0}, v_{j_1}, \ldots, v_{j_q}\} \in C \), \( (q \leq l) \), can be assigned to an element \( s_i \in S \), labeled like \( s_i \lambda v_{j_0}, s_i \lambda v_{j_1}, \ldots, s_i \lambda v_{j_q} \). The elements of set \( C \) will be called simlices labeled by the relation-associated elements of set \( S \) like \( \sigma(s_i) \), whereas the elements from set \( V \) will be called vertices. This notation emphasizes the distinction between elements from set \( S \) and set of elements \( C \) to which they are \( \lambda \)-related. Further, elements \( \sigma(s_i) \) are called \( q \)-dimensional simlices or just \( q \)-simlices, labeling them like \( \sigma_q(s_i) = \langle v_{j_0}, v_{j_1}, \ldots, v_{j_q}; \lambda \rangle \). The dimension of \( q \)-dimensional simplex is equal to the number of associated vertices minus 1. The \( r \)-dimensional face of a \( q \)-simplex is its subset of \( r + 1 \) elements, where \( r \leq q \), and it is also a simplex. As a result, sets \( S \), \( V \) and relation \( \lambda \) build a mathematical object called the simplicial complex \( K_S(V, \lambda) \) and it represents a collection of all simlices, together with all their faces, hence building the structure by attaching simlices via their mutual faces. Dimension of a simplicial complex is determined by the maximal dimension of simlices in the simplicial complex. In geometrical representation, simlices are displayed as \( q \)-dimensional polyhedra, e.g. 0-simplex is represented as a point; 1-simplex as an edge; 2-simplex as a triangle; 3-simplex as a tetrahedron etc.

Mutual \( q \)-face between two simlices \( \sigma_r(i) \) and \( \sigma_p(j) \), having dimensions \( r \) and \( p \), respectively, is a subset of their common \( q + 1 \) vertices, where \( q \leq r, p \), and they also share \((q−1)\)-, \((q−2)\)-, ..., 0-face also. The latter property assigned to these two simlices is called \( q \)-nearness. The degree of connectivity defined by the \( q \)-connectedness is the chain of simlices between any pair of simlices in the chain of simlices such that any adjacent pair of simlices share a mutual face, is determined by the lowest \( q \)-nearness, which is the smallest dimension of shared face. The \( q \)-connectedness property between simlices in the simplicial complex \( K \) generates the relation that satisfies mathematical equivalence \( \phi_q \) in a sense that it displays reflexivity, symmetry, and transitivity properties. Therefore, the subcomplex \( K_q \) of simplicial complex \( K \) is a set of all simlices with dimension greater or equal to \( q \) and the relation \( \phi_q \) partitions \( K_q \) into equivalence classes of simlices based on the \( q \)-connectedness. The number of \( q \)-connected components, or \( q \)-connectivity classes, represents the entries of \( Q \)-vector \[17\], or the first structure vector, and it is given as a vector starting from the largest \( q \) dimension in descending order.

3 Methods
of entries:
\[
Q = \{Q_{\text{dim}(K)} \ Q_{\text{dim}(K)-1} \ldots \ Q_1 \ Q_0\}. \tag{1}
\]
The number of simplices that have a dimension greater or equal to \( q \) represents the entries of the second structure vector \( N \), and it is written in the same descending order as the \( Q \)-vector starting from the largest dimension:
\[
N = \{N_{\text{dim}(K)} \ N_{\text{dim}(K)-1} \ldots \ N_1 \ N_0\}. \tag{2}
\]
The third structure vector describes the degree of connectedness at all dimensions of simplicial complex and it is defined like:
\[
\hat{Q}_q = 1 - \frac{Q_q}{N_q}. \tag{3}
\]
Structure vectors describe the global properties of a simplicial complex revealing its geometrical and combinatorial aspects, as well as behavior. On the other hand, in order to characterize immediate environment and, inner hidden structure of a single simplex, other quantities can be also derived within the mathematical framework of \( Q \)-analysis. One of these quantities is the so called node’s \( Q \)-vector \([12]\), \( Q^i \), that describes the environment of vertex \( i \) in a simplicial complex and it is defined as:
\[
Q^i = \{Q_{\text{dim}(K)^i} \ Q_{\text{dim}(K)-1}^i \ldots \ Q_1^i \ Q_0^i\}, \tag{4}
\]
where entries \( Q_q^i \) enumerate the number of \( q \)-simplices that contain the node \( i \). Topological dimension \( \text{dim}Q^i = \sum_q Q_q^i \) of vertex \( i \) is derived in order to quantify the total number of simplices in which the vertex \( i \) participates. Term "dimension" originates from conjugate complex \( K^{-1} \) \([17] \ [18]\), in which vertices and simplices have swapped roles - vertices are simplices and vice versa, and in conjugate complex the dimension of simplex associated to node \( i \) from underlying a simplicial complex is equal to the topological dimension of node \( i \), \( \text{dim}Q^i = \sigma(v_i) \).

Alongside the above, in this work two important properties of a simplicial complex are known as eccentricity and vertex significance of simplex \([19]\) are calculated. Let’s define \( \hat{q} \) as a dimension of a \( q \)-simplex \( \sigma(s_i) \) and \( \check{q} \) as a dimension at which it is joint to connectivity class for the first time, i.e. the dimension of a minimal face that attaches simplex \( \sigma(s_i) \) to other simplices. Now, eccentricity can be defined like:
\[
ecc(\sigma(i)) = \frac{\hat{q} - \check{q}}{\check{q} + 1}, \tag{5}
\]
which takes values between 0 and 1 and quantifies the degree of integration of a simplex in a complex. High values of eccentricity suggest low integration of simplex in local environment,
whereas low values imply high integration of simplex in the simplicial complex. On the other hand, the vertex significance quantifies the significance of simplices with the respect to vertices that build them. Therefore, let weight $\theta_i$ be the number of simplices that contain the vertex $i$, and $\Delta(i) = \sum_{j \in i} \theta_j$ represents the sum of $\theta_j$ for simplex $\sigma(i)$, where $j$ goes through all vertices in $\sigma(i)$. Vertex significance of simplex is defined as:

$$\nu_s(\sigma(i)) = \frac{\Delta(i)}{\max(\Delta)},$$

where $\max(\Delta)$ represents the maximal value from $\Delta$, and it represents the normalization constant so that vertex significance of simplex has maximal value 1, i.e. most significant simplex has $\nu_s = 1$.

In this paper, the inquiry of the considered phenomena is divided in two parts, one considering graph theory measures and the second part represent algebraic topology analysis of so called clique complex [20] constructed from the underlying mathematical graph $G$ (object which consider only pairwise interaction between vertices). In graph theory, a mathematical graph is presented via the adjacency matrix $A(G)$ which is a square matrix ($n \times n$) with rows and columns presenting vertices and its elements are commonly $A_{ij} = 1$ if vertices $i$ and $j$ are connected by an edge, otherwise $A_{ij} = 0$ (if an edge exists, element of adjacency matrix can also be the weight of an edge). Measures from the graph theory that are used are: $n$ - number of vertices in mathematical graph; $e$ - number of edges in mathematical graph; $\langle k \rangle$ - average degree of nodes (vertex's degree - total number of edges connected to a node); $cc$ - average clustering coefficient (the clustering coefficient is calculated as the ratio between the number of triangles around a vertex and the maximum number of triangles that could possibly be formed around that same vertex); $d$ - diameter (maximal distance between any pairs of vertices); $\langle p \rangle$ - average path length (vertex’s path length - average distance from a vertex to all other vertices); $modul$ - modularity (represents a measure of the structure of mathematical graph - determines the strength of division of a graph into communities or subgraphs).

Adjacency matrix is defined by binary relation between vertices of a mathematical graph. On the other hand, the incidence matrix $\Lambda(K(G))$ which defines the clique complex is the $n \times m$ matrix, where $m$ is the number of cliques in the underlying mathematical graph. In the clique complex $K(G)$, vertices are the same as vertices of the mathematical graph $G$, and simplices are identified as cliques (maximal fully connected subgraphs) which also represents the relation that form clique complex from the binary related mathematical graph. Bron-Kerbosch algorithm [21] is used for the detection of maximal cliques in the underlying mathematical graph.

For the purpose of the presented research, new criteria for layer-reconstruction of a simpli-
cial complex will be introduced. A delegate of a simplex is a vertex with the highest value of the node’s topological dimension. After determining all delegates within an underlying simplicial complex, a higher layer of a simplicial complex is constructed by connecting those delegates whose underlying simplices share a mutual face. This criterion can be applied multiple times in order to determine the inherent layers of vertices and simplices in an underlying complex.

4 Results

The network considered in this paper is the power grid of the Western States of the USA initially studied in [11], whereas the dataset is openly available in the KONECT project. Nodes are considered to be either as a generator, a substation or a transformer units, while links represent power supply lines. The network is formed by \( n = 4941 \) nodes and \( e = 6594 \) edges. As previously mentioned, different kind of elements in power grid network naturally operate at different hierarchical layers, hence it is preferable to identify them. In order to do that, the first step is to construct a clique complex from the underlying mathematical graph of the power grid network. Since this network has 3 different types of nodes, 3 iterations of algorithm are applied to extract delegates (see Fig. 2). Table 1 presents graph theory measures mentioned in the previous section for the original, as well as mathematical graphs of delegates after 3 iterations when the criteria for layer-reconstruction is applied.

Table 1: Common network measures for the original power grid networks and delegate networks after 1, 2 and 3 iterations of layer-reconstruction

<table>
<thead>
<tr>
<th></th>
<th>Original</th>
<th>D1</th>
<th>D2</th>
<th>D3</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n )</td>
<td>4941</td>
<td>2373</td>
<td>926</td>
<td>298</td>
</tr>
<tr>
<td>( e )</td>
<td>6594</td>
<td>7710</td>
<td>6462</td>
<td>5271</td>
</tr>
<tr>
<td>( \langle k \rangle )</td>
<td>2.67</td>
<td>6.50</td>
<td>13.96</td>
<td>35.38</td>
</tr>
<tr>
<td>( cc )</td>
<td>0.11</td>
<td>0.52</td>
<td>0.63</td>
<td>0.73</td>
</tr>
<tr>
<td>( d )</td>
<td>46</td>
<td>25</td>
<td>13</td>
<td>6</td>
</tr>
<tr>
<td>( \langle p \rangle )</td>
<td>18.99</td>
<td>10.05</td>
<td>5.40</td>
<td>2.73</td>
</tr>
<tr>
<td>modul</td>
<td>0.93</td>
<td>0.9</td>
<td>0.82</td>
<td>0.59</td>
</tr>
</tbody>
</table>

These results suggest that the size of network is shrinking through keeping only nodes that should exist at higher layers. From the presented results, the number of links stays more or less the same, which means delegate nodes that are connected at higher layers are connected via other nodes from lower layers and they represent some kind of binders in the underlying network. In this case, generators are some kind of network binders, but there is no strongly distinguishable community structure among them, as the modularity for network \( D3 \) suggests.

\footnote{1\text{http://konect.cc/networks/opsahl-powergrid/}}
On the other hand, in the underlying original network, smaller communities are formed around generators and those clusters have denser inter-connection compared to the connection between clusters. From the Table 1, it can be inferred that the values of clustering coefficient of these networks indicate strong pseudo-connection between nodes that belong to higher layers, that is generators and substations in $D_2$ network, and among generators in $D_3$ network even though they are not directly connected. The presence of those pseudo-connections between nodes that naturally exist on the same layer preserves the stability and sustainability of a power grid network and keeps it from creating unintentional islanding and severe disruptions in the infrastructure, hence having a role of some kind of bypass. When the random failure of a fraction of nodes occurs, it is likely that still active delegate nodes will take the role of connections that will keep the significant portion of the network in function.

Regarding the results obtained for structure vectors (see Fig. 3), it can be concluded that inner-connections and degree of connectedness increases as simplicial complexes progress toward higher layers of the power grid network. First two simplicial complexes of the underlying network of delegates display small changes compared to the original power grid network. On
the other hand, the structure of simplicial complex obtained in the last iteration reveals hidden properties that simplicial complexes in previous iterations do not possess, and it is another confirmation of the existence of previously mentioned pseudo-connections of generators. Therefore, the network of delegates which may have a role of the network of generators presents a backbone of the power grid network. In this sense, the intertwine pseudo-connections, which form bigger simplices, suggest high stability regarding disruptions and islanding in the infrastructure. If there was a poor connection of the base of one power grid network, there would be a high risk of creation of unintentional islanding and blackout of parts of a power grid.

From the Fig. 4, the simplicial complex of \( D3 \) network has less average eccentricity, and higher average topological dimension and vertex significance than networks at other 3 layers, as expected, and it is complementary to previous results. If carefully observed, topological
dimension and vertex significance of simplices in simplicial complexes $D_2$ and $D_3$ have rapid decline intersecting $D_1$ around $500^{th}$ and $200^{th}$ ranked node respectively. This can suggest that some of the nodes and simplices constructed by these nodes, do not naturally belong to that layer, therefore they should be restrained in lower layers. Tails of above rankings demand further inspection because if node, that naturally lives on lower layers, is progressing to the higher layers as delegate, it can be the one that has potential of causing damage to the whole cluster to which it belongs, and therefore the power grid network's integrity and stability.

5 Conclusion and discussion

Results reveal latent embedment of the most significant delegates with higher-order aggregation than at the level of the initially obtained simplicial complex from an original network. Thus, they represent the backbone of a power grid network, and higher-order clustering of the same delegates via pseudo-links refers to the high stability of the network.

Topological analysis of local architecture should classify structural layers where nodes naturally belong. The study of the tails of rankings and local properties of simplicial complex of delegates in a power grid network can reveal whether there are potential issues that can affect network stability when nodes are located at the inaccurate layer. In other words, if a node that naturally originates in a lower hierarchical layer is found mixed within the same simplex of delegates with nodes that belong on a higher layer, and vice versa, it would present a potential issue in power grid stability and can cause cascade disruptions ensuing blackouts and islanding. Therefore, this model and algorithm have great potential for testing power grid stability whether it is a network in construction or an original network, and can suggest nodes reconnecting in order to achieve a network with a long term stability and sustainability.

Acknowledgements

The research was funded by the Ministry of Education, Science and Technological Development of the Republic of Serbia.

References


Submitted: 16.06.2022.
Revised: 29.07.2022.
Accepted: 11.08.2022.