The Power Grid as a Complex Network: a Survey

Giuliano Andrea Pagani and Marco Aiello

Distributed Systems Group
Johann Bernoulli Institute for Mathematics and Computer Science
University of Groningen
Groningen, The Netherlands

email: {g.a.pagani,m.aiello}@rug.nl
http://www.cs.rug.nl/ds/

March 22, 2012

Abstract

The statistical tools of Complex Network Analysis are of great use to understand salient properties of complex systems, may these be natural or pertaining human engineered infrastructures. One of these that is receiving growing attention for its societal relevance is that of electricity distribution. In this paper, we present a survey of the most important scientific studies investigating the properties of several Power Grids infrastructures using Complex Network Analysis techniques and methodologies. We categorize and explore the most relevant literature works considering general topological properties, differences between the various graph-related indicators and reliability aspects.
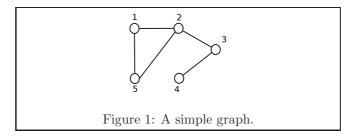
1 Introduction

Complex Network Analysis (CNA) is a relatively young field of research. The first systematic studies appeared in the late 1990s [1, 2, 3, 4] having the goal of studying the properties of large networks that behave as complex systems. The research owes a great deal of its foundations to the seminal work on Random Graphs of Erdős and Rényi [5, 6] who studied asymptotic properties of stochastic graph processes. Complex Network Analysis has been used in many different fields of knowledge, from biology [7] to chemistry [8], from linguistics to social sciences [9], from telephone call patterns [10] to computer networks [11] and web [12, 13] to virus spreading [14, 15, 16] to logistics [17, 18, 19] and also inter-banking systems [20]. Men-made infrastructures are especially interesting to study under the Complex Network Analysis lenses, especially when they are large scale and grow in a decentralized and independent fashion, thus not the result of a global, but rather of many local autonomous designs. The Power Grid is a prominent example. But what do we mean by Power Grid in the context of the present treatment?

We focus on the electricity transmission and distribution Power Grid as it is essential for today's society as an enabling infrastructure, but also its efficiency and working has major consequences, among other things, for the environment. Blackouts seem to have a special role in reminding us of the importance of the Grid and how much we give its availability for granted. From the technological point of view, the electrical system and Power Grid involve many scientific knowledge areas that contribute to the design, operations and analysis of power systems: Physics (electromagnetism, classical mechanics), Electrical engineering (AC circuits and phasors, 3-phase networks, electrical systems control theory) and Mathematics (linear algebra, differential equations). Traditional studies tend to have a "local" view of the Grid, e.g., defining how to design a transformer and predicting its functioning. Typically, studies tend to focus on the physical and electrical properties (e.g., [21]), or the characteristics of the Power Grid as a complex dynamical system [22], or again, the control theory aspects [23]. The move from a "local" to a "global" view of the Power Grid as a complex system is possible by resorting to Complex Network Analysis and statistical graph theory.

The goal of the present treatment is to provide a survey and compare the most well-known scientific studies conducted using Complex Network Analysis techniques concerning Power Grid systems. We consider several parameters to assess the differences between the various studies and try to enucleate the most important aspects of each study. We start by introducing the methods and metrics that are evaluated in this work (Section 2); the section contains the basic definitions and simple examples in order to establish a common background. Section 3 provides the main characteristics of all the studies. The actual comparison of these using CNA metrics are reported and discussed in Section 4 which also provides a conclusion to the paper.

2 Background and Survey Methodology

Before going further analyzing the various studies in detail, some common definitions need to be stated in order to have a common ground. The essential concepts from the graph theory are also given to provide a common basis. For each graph property described, a concrete example on a small graph, as shown in Figure 1, is provided to better understand property's application.

Figure 1: A simple graph.

As described in Section 1, all the works that are examined in the present manuscript consider the Power Grid networks as graphs following the mathematical meaning of the term.

Definition 1 (Graph). *A graph G is a pair of sets $G(V,E)$ where V is the set of vertexes and E is the set edges. An edge $e_{i,j}$ is a pair of vertexes (v_i, v_j). If $(v_i, v_j) \in E$ then v_i and v_j are said to be* adjacent *or* neighboring *and are called* end-vertexes *of the edge.*

Considering the Power Grid, the sets composing the graph assume particular interest from an operational and physical point of view. The physical components of the Power Grid assume a meaning in the theoretical representation of the Power Grid as a graph according the following interpretation.

Definition 2 (Power Grid graph). *A Power Grid graph is a graph $G(V, E)$ such that each element $v_i \in V$ is either a substation, transformer, or consuming unit of a physical Power Grid. There is an edge $e_{i,j} = (v_i, v_j) \in E$ between two nodes if there is a physical cable connecting directly the elements represented by v_i and v_j.*

Therefore, following the properties of the graphs is an interesting first way of categorizing the Power Grid under analysis. In particular, distinctions can be made regarding the *order* and *size* of the graph. *Order* is the number of vertexes composing the graph, while *size* is the number of edges in the same graph. More formally:

Definition 3 (Order and size of a graph). *Given the graph G the order is given by $N = |V|$, while the size is given by $M = |E|$.*

Example:
The graph G shown in Figure 1 is characterized by the set of vertexes V:

$$V = \{1, 2, 3, 4, 5\}$$

and by the set of edges E:

$$E = \{(1, 2), (1, 5), (2, 3), (2, 5), (3, 4)\}$$

the order and size of G are $|V| = 5$ and $|E| = 5$ respectively.

An important property that characterizes a vertex of a graph is the degree, that is the number of vertexes the node is adjacent to. More formally, this quantity is defined as:

Definition 4 (Neighborhood and degree). *The set of vertexes adjacent to a vertex $v \in V$ represents the* neighborhood *of v that is denoted by Γ_v. The* degree *of v is $d(v) = |\Gamma_v|$.*

From *order* and *size* it is possible to have a global value for the connectivity of the vertexes of the graph, known as *average node degree* . That is $<k> = \frac{2M}{N}$.

Usually, it is not essential to have the specific information regarding the node degree of only a certain node or the average degree. It is more interesting understand the the overall characteristics of a graph considering its statistical measures. In particular, one measure is the node degree probability distribution. More formally,

Definition 5 (Node degree distribution). *Consider the degree k of a node in a graph as a random variable, the function*

$$N_k = \{v \in G : \ d(v) = k\}$$

is called probability node degree distribution.

The shape of the distribution is a salient characteristic of the network. For the Power Grid, the shape is typically either exponential or a Power-law. More precisely, an exponential node degree (k) distribution has a fast decay in the probability of having nodes with relative high node degree. It follows the relation:

$$P(k) = \alpha e^{\beta k}$$

where α and β are parameters of the specific network considered. While a Power-law distribution has a slower decay with higher probability of having nodes with high node degree. It is expressed by the relation:

$$P(k) = \alpha k^{-\gamma}$$

where α and γ are parameters of the specific network considered. We remark that the graphs considered in the Power Grid domain are usually large, although finite, in terms of order and size thus providing limited and finite probability distributions.

The node degree distribution gives some information about the static situation of the network, but it does not give any information about the paths that can be followed in the graph to move from one node to another. To investigate properties dealing with paths between nodes, that for a Power Grid graph are important to assess which nodes have to sustain the highest flow of energy, some further concepts are essential.

The concepts of *path* and *path length* are crucial to understand the way two vertexes are connected.

Definition 6 (Path and path length). *A* path *of G is a subgraph P of the form:*

$$V(P) = \{x_0, x_1, \ldots, x_l\}, \qquad E(P) = \{(x_0, x_1), (x_1, x_2), \ldots, (x_{l-1}, x_l)\}.$$

such that $V(P) \subseteq V$ and $E(P) \subseteq E$. The vertexes x_0 and x_l are end-vertexes *of P and $l = |E(P)|$ is the* length *of P. A graph is* connected *if for any two distinct vertexes $v_i, v_j \in V$ there is a finite path from v_i to v_j.*

Definition 7 (Distance). *Given a graph G and vertexes v_i and v_j, their* distance *$d(v_i, v_j)$ is the minimal length of any $v_i - v_j$ path in the graph. If there is no $v_i - v_j$ path then it is conventionally set to $d(v_i, v_j) = \infty$.*

Definition 8 (Shortest path). *Given a graph G and vertexes v_i and v_j the* shortest path *is the the path corresponding to the minimum of to the set $\{|P_1|, |P_2|, \ldots, |P_k|\}$ containing the lengths of all paths for which v_i and v_j are the end-vertexes.*

Example:

Example of paths between vertex 1 and vertex 4 are:

$$P_{1,4} = \{1 - 2 - 3 - 4\}$$

whose length $l_P = 3$ but also:

$$P'_{1,4} = \{1 - 5 - 2 - 3 - 4\}$$

is a valid path whose length $l_{P'} = 4$, therefore the shortest path between vertex 1 and vertex 4 is $P_{1,4}$.

The distance between vertex 1 and vertex 4 is:

$$d(1,4) = 3$$

while distance between vertex 1 and vertex 5 is:

$$d(1,5) = 1$$

To describe the importance of a node with respect to minimal paths in the graph, the concept of betweenness helps. Betweenness (sometimes also referred as *load*) for a given vertex is the number of shortest paths between any other nodes that traverse it. More formally,

Definition 9 (Betweenness). *The betweenness $b(v)$ of vertex $v \in V$ is*

$$b(v) = \sum_{v \neq s,t} \sigma_{st}(v)$$

where $\sigma_{st}(v)$ is 1 if the shortest path between vertex s and vertex t goes through vertex v, 0 otherwise.

Example:

Vertex 2 is involved in the following shortest paths:

$$P_{1,3}, P_{1,4}, P_{3,1}, P_{3,5}, P_{4,1}, P_{4,5}, P_{5,3}, P_{5,4}$$

therefore betweenness of vertex 2 is:

$$b(2) = 8$$

In the Power Grid domain however it is not always true that most central node are the most important nodes for the Grid since they have to sustain the most number of shortest paths. It is not a priori true that the nodes that sustain the greatest electricity flow between two nodes in a real Power Grid are the one with highest betweenness since electricity flow could not follow the topological shortest path. Although it is important to know what is the betweenness of the most important nodes, to shift from a local to a global graph measure it is also useful to have a high level picture of the state of betweenness for the whole graph. A statistic measure is then used:

Definition 10 (Betweenness distribution). *Consider the betweenness value l of a node in a graph as a random variable, the function*
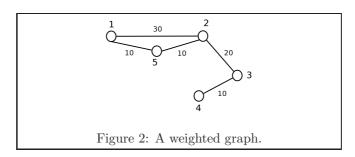
$$L_k = \{v \in G : b(v) = l\}$$

is called betweenness probability distribution *of the graph G.*

Another distinction that is available between graphs is their characterization as unweighted (as considered in Definition 1) or weighted. The edges are not always the same, in fact, different importance can be associated to them considering their role in the network or physical properties connecting vertexes. From a formal point of view:

Definition 11 (Weighted graph). *A weighted graph is a pair $G(V, E)$ where V is the set of vertexes and E is the set of edges. An edge $e_{i,j,w} = (v_i, v_j, w)$ is a triple where $v_i, v_j \in V$ and $w \in \mathbb{R}$. w is called* weight *of the edge.*

> **Example:**
> Figure 2 represents a weighted graph: each edge is characterized by a certain weight. The weight associated to the edge might be related to physical properties of the links (e.g., resistance of the cables in a Power Grid network) connecting the objects represented as vertexes in the graph.

Figure 2: A weighted graph.

Specific properties of a graph are best understood by resorting to the matricial counterparts, in particular, to the Adjacency matrix and Laplacian matrix graph representations.

Definition 12 (Adjacency matrix). *The adjacency matrix $A = A(G) = (a_{i,j})$ of a graph G of order N is the $N \times N$ matrix given by*

$$a_{ij} = \begin{cases} 1 & if\ (v_i, v_j) \in E, \\ 0 & otherwise. \end{cases}$$

Definition 13 (Laplacian matrix). *Let $D = (D_{ij})$ be the $N \times N$ diagonal matrix with $D_{ii} = d(v_i)$ the degree of v_i in G and A the adjacency matrix of G.*
The matrix $L = D - A$ is the Laplacian matrix of graph G.

> **Example:**
> The Adjacency matrix for graph G is:
>
> $$A(G) = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{pmatrix}$$
>
> The Laplacian matrix for graph G is:
>
> $$L(G) = \begin{pmatrix} 2 & -1 & 0 & 0 & -1 \\ -1 & 3 & -1 & 0 & -1 \\ 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ -1 & -1 & 0 & 0 & 2 \end{pmatrix}$$

Another interesting property that is investigated for networks and graphs is the so called small-world property. Although a complete coverage of the small-world problem is beyond the scope of the present work, we recall the basic definitions and refer to [1, 24]. We begin with the clustering coefficient.

Definition 14 (Clustering coefficient). *The* clustering coefficient *of a node v is*

$$\gamma_v = \frac{|E(\Gamma_v)|}{\binom{k_v}{2}}$$

where $|E(\Gamma_v)|$ is the number of edges in the neighborhood of v and $\binom{k_v}{2}$ is the total number of possible edges in Γ_v. The clustering coefficient *of graph G is γ, the average value of the clustering coefficient of all nodes of the graph.*

Definition 15 (Characteristic path length). *Let $v_i \in V$ be a vertex in graph G, the* characteristic path length *for G, L_{CP} is the median of d_{v_i} where:*

$$d_{v_i} = \frac{1}{(N-1)} \sum_{i \neq j} d(v_i, v_j)$$

is the mean of the distances connecting v_i to any other vertex v_j in G whose order is N.

Example:
The clustering coefficient for vertex 1 is:

$$\gamma_1 = 1$$

while for the entire graph it is the average of $\{\gamma_1 = 1, \gamma_2 = \frac{1}{3}, \gamma_3 = 0, \gamma_4 = 0, \gamma_5 = 1\}$ that is:

$$\gamma_G = 0.467$$

The characteristic path length is the median of

$$\{d_{v_1} = \frac{7}{4}, d_{v_2} = \frac{5}{4}, d_{v_3} = \frac{6}{4}, d_{v_4} = \frac{9}{4}, d_{v_5} = \frac{7}{4}\}$$

that is:

$$L_{CP} = \frac{7}{4} = 1.75$$

Definition 16 (Random Graph). *A graph $G(V, E)$ of a given size is a* random graph *if it is the result of a random process where the edges between any two nodes have been chosen independently with probability $p \in [0, 1]$ among all possible $\binom{N}{2}$ edges.*

We remark that there exist several definitions of random graphs as provided by Erdős and Rényi [6], Bollobas [25] or Durrett [26]. These mostly lead anyhow to the same asymptotic properties. This holds also for the clustering coefficient e.g. [27] which differs from the definition we proposed (in line with Watts and Strogatz in [1] and Watts in [24]).

Small-world networks (SW), proposed by Watts and Strogatz in [1], own two important properties at the same time: the characteristic path length is close in value to the one of a random graph (RG) ($CPL_{SW} \approx CPL_{RG}$) and they have a much higher clustering coefficient ($CC_{SW} \gg CC_{RG}$). Small-worlds are a better model than random graphs for social networks and other phenomena [9, 28, 29, 30] and thus a model to keep in mind for the Power Grid, too.

Another investigation usually performed when analyzing Power Grid and that is almost always the motivation that drives Complex Network Analysis studies related to electrical infrastructures is the investigation of reliability. Usually, the investigation involves evaluating the disruption behavior of the graph when its nodes or edges are removed. There are basically two ways to perform this analysis: choosing the nodes to be removed randomly or selecting the nodes following a certain property or metric significant for the network. Commonly the metric used to remove nodes follows the highest degree or highest betweenness to simulate *targeted attacks* that focus on specific nodes with certain properties or importance for the network.

Other terms to compare the various Power Grid studies involve more general characteristics of the network under analysis. In particular, the geographical location of the analyzed Grid is responsible for topological properties due to the different morphological characteristics of different countries. Another relevant aspect deals
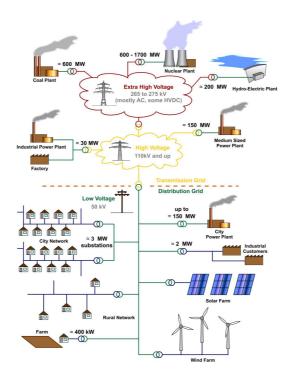
Figure 3: Organization of the Power Grid layers.

with the layer of the Power Grid under investigation since differences can emerge from a topological perspective investigating the different ends in which the Grid is usually partitioned: High, Medium and Low Voltage. An example on how the Power Grid is organized is shown in Figure 3. It is also important to have information if the type of Power Grid graph under analysis comes from a real network infrastructure or it is a synthetic sample extracted from blueprint models for the Power Grid such as the Bus models of IEEE.

The motivations to include the works in this survey are based on the quality of the research performed, the rigor in the application of Complex Network Analysis methodologies and the geography of the Power Grid analyzed in order to cover a broad spectrum of the infrastructure realized in the different countries and identify possible differences.

3   The Power Grid as a Complex Network

Complex network analysis studies are becoming more and more popular given the amount of natural and human complex systems. The Power Grid is clearly amenable to such studies and a number of these have been performed on the High Voltage Grid. Here we describe the most important aspects of each work under investigation. In particular, the works that are considered in this review are: [31, 32, 33, 34, 35, 36, 37, 38, 39, 24, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62]. These have been chosen based on the following factors: they are specifically about the Power Grid, they cover US, European, Chinese Grids or synthetic topologies of electrical engineering literature, they have samples of different sizes and, most importantly, these are the best-known and most representative works on the topic of CNA and Power Grid.

3.1   Basic Power Grid characteristics

The aspects considered in this first basic assessment of the studies take into account general and non-technical aspects so to give a global idea of the Grid considered, see Table 1. Several aspects of comparison are considered: the number of nodes and lines composing the Grid (second and third column); the type of sample considered either a real Grid or synthetic samples, for instance, coming from IEEE literature such as IEEE Bus systems (fourth column); the type of Grid analyzed (fifth column) in belonging either to the transmission part (High Voltage) or to the distribution part (Medium and Low Voltage); another essential information deals with the geography of the Grid (last column).

Albert *et al.* [31] study the reliability aspects of the United States Power Grid. They build a graph based on the information of the POWERmap system (developed by Platts company) consisting of 14099 nodes representing power plants and substations and 19657 edges representing High Voltage lines (115-765 kV). This same Power

---

[1]The values for nodes and lines in this table refer only to a snapshot of Shanghai Power Grid

Work	Number of Nodes	Number of Lines	Sample Type	Network Type	Geography
[31]	~14000	~19600	Real	HV	North America
[32]	~300	~500	Real	HV	Italy
[33]	~314000	N.A.	Real	HV	North America
[34]	~4800	~5500	Real	HV	Scandinavia
[35]	~2700	~3300	Real	HV	Europe
[36]	~3000	~3800	Real	HV	Europe
[37]	~3000	~3800	Real	HV	Europe
[38]	~370	~570	Real	HV	Italy, France and Spain
[39]	~370	~570	Real	HV	Italy, France and Spain
[24]	~4900	~6600	Real	HV	Western US
[63]	~8500	~13900	Synthetic and real	HV	Western US and New York State Area
[40]	~4850	~5300	Real	MV/LV	Netherlands
[41][1]	~210	~320	Synthetic and real	HV	China
[42]	N.A.	N.A.	Real	HV	Europe
[43]	300	411	Synthetic	HV	
[44]	~6400	~8700	Synthetic and real	HV	North America, Scandinavia and Korea
[45]	300	411	Synthetic	HV	
[46]	~8500	~13900	Synthetic and real	HV	Western US and New York State Area
[47]	~30	~13900	Synthetic and real	HV	Western US and New York State Area
[48]	~900	~1150	Real	HV	China
[49]	~3200	~7000	Synthetic and real	HV	New York State Area
[50]	~4900	~6600	Real	HV	Western US
[51]	~1700	~1800	Real	HV	China
[52]	~39	~46	Syntethic	HV	
[53]	~39	~46	Syntethic	HV	
[54]	~2500	~2900	Real	HV	China
[55]	~15400	~18400	Real	HV	North America and China
[56]	~550	~800	Synthetic	HV	
[57]	~14000	~19600	Real	HV	North America
[58]	~90	~120	Synthetic	HV	
[59]	~550	~700	Synthetic and real	HV	Italy
[60]	~29500	~50000	Synthetic and real	HV	North America
[61]	~400	~700	Synthetic		
[62]	~900	~1300	Synthetic and real	HV	South-East US
[64]	~60	~110	Real	HV	India

Table 1: Comparison between studies using CNA for the Power Grid.

Grid sample is used also in the study by Kinney *et al.* [57] where the reliability of the Grid is analyzed with a different set of metric compared to those used by Albert *et al.*.

Crucitti *et al.* [32] analyze the Italian High Voltage Power Grid from a topological perspective. They build a model based on data from GRTN (the Italian Grid manager at that time) consisting of 341 substations (nodes) and 517 transmission lines (edges) belonging to the High Voltage segment (220-380 kV).

Chassin *et al.* [33] analyze the North American Power Grid. They treat the North American Grid as composed by two different networks, i.e., the Eastern Grid and the Western Grid, that are analyzed separately. This assumption is justified by the authors by the presence of a small linking between the two Grids (this is realized on purpose to avoid the spreading of blackouts across the entire country) realized in direct current technology. In addition, the data they use contain much more nodes and links (almost 236000 nodes for the Eastern and more than 78000 for the Western Grid, and it is based on the estimations of typical load and maximal capacity of distribution electrical feeders); data come form Western Electricity Coordinating Council (WECC) and North American Electric Reliability Council (NERC), respectively.

Holmgren [34] analyzes the Nordic Power Grid involving the High Voltage Grids of Sweden, Finland, Norway, and the main part of Denmark which give an overall graph composed by almost 4800 nodes and more than 5500 edges. The author compares the Nordic network with the Western U.S. Power Grid.

Casals *et al.* [35] analyze the whole European Power Grid and try to extract non-topological reliability measures investigating the topological properties of the network. The Power Grid analyzed is the High Voltage end composed of almost 2800 nodes that span across all European continent. Casals *et al.* [36] consider the High Voltage Grids of many European countries analyzing them together and as separate entities having an overall sample of more than 3000 nodes and around 4300 edges. Solé *et al.* [37] go further in exploring the same Power Grid data analyzed in [36], in particular, they focus on analyzing the targeted attacks to European Power Grids. Rosas-Casals [42] once again considers the European High Voltage Power Grid to understand the distribution of failures that have characterized it from 2002 to 2008.

Crucitti *et al.* [38] analyze the High Voltage Power Grid of Italy (127 substations and 171 lines belonging to 380kV network), France (146 substations and 223 lines belonging to 400kV network) and Spain (98 substations and 175 lines belonging to 400kV network). This same sample is analyzed by Rosato *et al.* [39] to investigate the main topological properties of these Grids.

Watts [24] dedicates a subsection to explore the properties of the Western States Power Grid of the U.S. This Grid is treated as an undirected unweighted graph in which all the nodes of the network no matter their task (e.g., generators, transformers, substations) are equally considered as nodes. The same assumption is done for the edges: the transmission lines are considered equal even if the voltages they involve can be extremely different (the Grid considered belongs only to High Voltage segment with lines varying from 345 to 1500 kV). The overall graph is quite large (4941 nodes). This same network is the base for the study by Wang *et al.* [50] which focuses on the vulnerability of the Grid.

Wang *et al.* [63, 46] investigate both on real Power Grid samples for about 8000 nodes (the networks analyzed are the American NYISO and WSCC) and 4 synthetic reference models belonging to the IEEE literature that account for about 500 nodes. In their other work, [49], Wang *et al.* consider once again the NYSO-2935 Bus and the synthetic IEEE model: IEEE 300-Bus system.

The work of Bompard *et al.* [43] refers to the IEEE Bus blueprint infrastructure which applies an enhanced betweenness analysis to the IEEE 300-Bus. Similarly the work of Dwivedi *et al.* [52, 53] which study respectively the IEEE 39-Bus and the IEEE 118-Bus. Also the study of Pahwa *et al.* [56] refers to the synthetic Buses from IEEE (IEEE 300-Bus and IEEE 118-Bus) and the WSCC 179-Bus equivalent. Arianos *et al.* [58] study smaller IEEE networks: the IEEE 30-Bus and the IEEE 57-Bus.

In our work [40], we study the Medium Voltage and Low Voltage end of the Power Grid with special focus on the northern Netherlands situation. The overall sample is almost 700 nodes for the Low Voltage part and 4200 for the Medium Voltage one.

Mei *et al.* in their book [41] perform various simulation related to Power Grid vulnerability and blackout conditions considering the synthetic models of IEEE literature (e.g., IEEE 14-Bus, IEEE 30-Bus, IEEE 39-Bus and IEEE 118-Bus) and also real samples of Chinese Power Grid. In particular they analyze the 500kV Northeast Power Grid of China and a snapshot of Shanghai High Voltage Power Grid which consist of more than 200 nodes and 300 lines.

Kim *et al.* [44] analyze basic metrics of several Grids both the synthetic ones belonging to IEEE literature such as 14-Bus and 118-Bus (respectively 14 nodes and 20 edges and 118 nodes and 179 edges) and real High Voltage networks. The real samples are the American Mid-Continental Area Power Pool (MAPP) of 230kV and higher composed of 575 nodes and 754 edges, the Electric Reliability Council of Texas (ERCOT) of 345kV composed of 148 nodes and 209 edges and the Western Systems Coordinating Council (WSCC) from 100kV and above which has about 4600 nodes and more than 6200 edges. In addition, they analyze the Nordel network in Scandinavia (100kV and above) characterized by 410 nodes and 564 edges and the Korea Electric Power Corporation (KEPCO) (66kV and above) made of 553 nodes and 783 edges.

Hines *et al.* [45] uses the High Voltage synthetic model IEEE 300-Bus system as the reference network on

which compute centrality metrics exploiting the impedence representation of lines connecting Buses (nodes).

Han *et al.* [48] investigate the topological characteristics of two Power Grids sections belonging to the Chinese High Voltage network: the Anhui Power Grid (84 nodes and 112 edges) and the bigger East China Power Grid (769 nodes and 1029 edges). Also focused on the Chinese High Voltage Grid is the work of Ding *et al.* [51] analyze the topological properties of two samples of the Chinese High Voltage Power Grid. In particular, the Sichuan-Chongqing Grid (724 nodes and 771 edges) and the Guangdong Grid (956 nodes and 1049 edges). Guohua *et al.* [54] analyze another sample of the Chinese Power Grid, in particular the North China Power Grid which is composed by 2556 nodes and 2892 edges.

In one of the first comparison of power Grids [55], Sun compares the Western American Grid ($\approx$5000 nodes and $\approx$6600 edges), the North China Power Grid ($\approx$8000 nodes and $\approx$9000 edges) and the Center China Power Grid ($\approx$2400 nodes and $\approx$2800 edges). The model used considers the transmission lines (above 110kV for the Chinese samples and above 115kV for the American one) as unweighted and undirected edges, while generators, substations and transformers are considered all equal nodes of the graph.

The Italian transmission is once again the example of network used to test Power Grid vulnerabilities [59], comprising 521 nodes and 679 transmission lines. In addition in the same work a synthetic Bus with 34 nodes is tested as well.

The work of Hines *et al.* [60] investigates the vulnerability aspects both of synthetic samples (IEEE 300-Bus) and 40 control areas (out of the 136 areas) of the North American Eastern Interconnect Grid which represent about 29000 nodes and 50000 transmission lines of the Grid.

The work of Pepyne [61] focuses on evaluating the cascading effects of synthetic models coming from the IEEE literature (e.g., IEEE 57-Bus and IEEE 118-Bus) and on a sample satisfying the small-world model [1, 24] of 200 nodes and 400 edges.

Brummitt *et al.* [62] investigate the cascading effects in interdependent networks (i.e., networks which are loosely coupled) and the effects of adding connectivity (e.g., power lines) to mitigate or aggravate cascade effects. For this purpose two loosely connected samples of the South-East region of the United States are used as samples, for a total of about 900 nodes and 1300 edges.

A new geography is explored by Chaitanya *et al.* [64] who investigate the properties of an High Voltage Grid of the eastern region of India. The network is built on 63 nodes and 113 transmission lines.

Notice that the numbers in the second and third column are not the exact numbers, but they are an approximation to give the idea of the importance of the sample.

Summary:
In summary, the data are in the most cases extracted from real samples, that is, they represent real electric infrastructures deployed other works in addition to real Power Grids while fewer studies consider synthetic models as shown in Figure 4. Most of these synthetic approaches consider samples coming from IEEE blueprints such as IEEE Bus systems (a representation of the various IEEE Bus models used in the surveyed articles is shown in Figure 5), while very few concentrate only on other synthetic samples (e.g., non-IEEE models, small-world models, random graphs); the number of synthetic model used is shown in Figure 6. Almost all samples belong to the High Voltage end of the Power Grid that are the lines used for long range transmission to which big power plants are attached too; the only exception is our study [40] that is focused on the distribution part of the Grid (i.e., Medium and Low Voltage network). From a geographical perspective the samples are mainly localized in the United States or in Europe with some studies that consider Chinese High Voltage samples; a map of the countries whose Grids are analyzed is represented in Figure 7 and the number Grid analyzed for a given country is shown in Figure 8. Another main commonality is to treat the Grid as an undirected graph where each substation or transformer represents a node and each line transporting electricity is an edge.

3.2   Statistical global graph properties

The main characteristics from a graph and Complex Network Analysis perspective of the Grids under analysis are summarized in Table 2. Several aspects of comparison are considered: the *order* (N) and *size* (M) of the graph (second and third column) corresponds to the number of nodes (*order*) and number of lines (*size*) actually in the Power Grids. The average degree, computed as $<k> = \frac{2M}{N}$, gives a general idea of how many vertexes
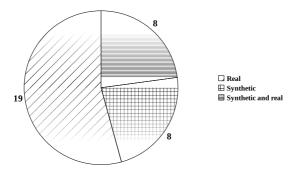
Figure 4: Number of studies that consider real Power Grid samples or synthetic models.
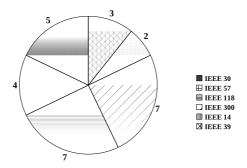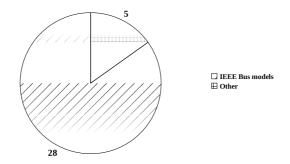

Figure 5: IEEE literature bus model analyzed.

Figure 6: Number of models used coming from the IEEE literature compared to other models.

is an average vertex connected to (fourth column). Fifth, sixth and seventh column give information about the type of statistical analysis performed on the graph, in particular, the assessment of node degree distribution and betweenness distribution together with an evaluation of the path length are considered. Another term of comparison deals with the type of graph analyzed taking into account weights or simply use the unweighted definition of graph. Last two columns of the table consider the type of aim of the graph analysis either an investigation of the disruption behavior of the graph or the evaluation of the small-world properties.

Many studies [40, 63, 24, 37, 42] remark the limited value of the average node degree, generally between 2 and 3, for the Power Grid especially if compared to other types of Complex Networks (e.g., the Web, social networks). This is due to the physical, geographical and economical constraints that are associated to the substations and power cables.

Half of the studies focuses on the investigation of the node degree distribution statistics since this information is a key to find what kind of theoretical probability model is beyond the sample, allowing to establish which kind of network the sample can be associated with. The other studies that do not take into account node degree distribution statistics focus on other very specific aspects of the Complex Network Analysis. Betweenness distribution statistics are less common, in fact, only few studies dig into this property, although it provides essential information related to the load sustained by the nodes of the network.

Almost half the studies take into account the path length to investigate the effort that it takes to move from one node to any other one. The study of the path length is usually not performed per se, but it is essential to then

---

[2]The values for nodes and lines in this table refer only to a snapshot of Shanghai Power Grid.

Work	Sample Order	Sample Size	Average degree	Node Degree Distribution Statistics	Betweenness Distribution Statistics	Path Length Analysis	Weighted/Unweighted Analysis	Resilience Analysis	Small-world Investigation
[31]	~14000	~19600	~2.80	✓	✓		Unweighted	✓	
[32]	~300	~500	~3.33	✓	✓	Indirectly through efficiency metric	Weighted not based on physical properties	✓	
[33]	~314000	N.A.	N.A.	✓			Unweighted	✓	
[34]	~4800	~5500	~2.29	✓		✓	Unweighted	✓	✓
[35]	~2700	~3300	~2.44	✓			Unweighted	✓	
[36]	~3000	~3800	~2.53	✓		✓	Unweighted	✓	✓
[37]	~3000	~3800	~2.53	✓			Unweighted	✓	
[38]	~370	~570	~3.08			Indirectly through efficiency metric	Unweighted	✓	
[39]	~370	~570	~3.08	✓		✓	Unweighted	✓	
[24]	~4900	~6600	~2.69			✓	Unweighted		✓
[63]	~8500	~13900	~3.27	✓		✓	Unweighted and impedance analysis		✓
[40]	~4850	~5300	~2.18	✓	✓	✓	Both	✓	✓
[41]2	~210	~320	~3.05			✓	Both	✓	✓
[42]	N.A.	N.A.						✓	
[43]	300	411	2.74				Both	✓	
[44]	~6400	~8700	2.72			✓	Unweighted	✓	✓
[45]	300	411	2.74	✓(chart only)	✓		Both	✓	
[46]	~8500	~13900	~3.27	✓			Unweighted	✓	
[48]	~900	~1150	~2.55	✓	✓	✓	Weighted	✓	✓
[49]	~3200	~7000	~4.375	✓(chart only)	✓		Weighted	✓	
[50]	~4900	~6600	~2.69				Unweighted	✓	
[51]	~1700	~1800	~2.12			✓	Both		✓
[52]	~39	~46	~2.36				Weighted	✓	
[53]	~150	~46	~2.36				Weighted	✓	
[54]	~2556	~2892	~2.26				Weighted	✓	
[55]	~15400	~18368	~2.39	✓(for one sample only)		✓	Unweighted	✓	✓
[56]	~550	~800	~2.91			✓	Unweighted	✓	
[57]	~14000	~19600	~2.80				Weighted not based on on physical properties	✓	
[58]	~90	~120	~2.67				Weighted	✓	
[59]	~550	~700	~2.55				Weighted	✓	
[60]	~29500	~50000	~3.39			✓	Weighted	✓	
[61]	~400	~700	~3.5			✓	Weighted	✓	✓
[62]	~900	~1300	~2.89				Unweighted	✓	
[64]	~60	~110	~3.67				Unweighted		

Table 2: Comparison of the main characteristics of the graphs related to Power Grids.
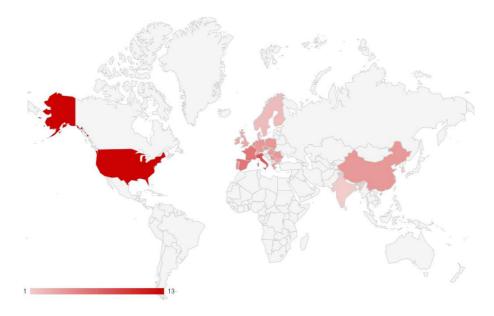
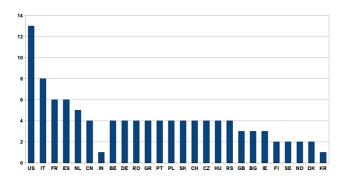Figure 7: Map of the Power Grid infrastructure studied using CNA approach.

Figure 8: Number of times Power Grid of a given country are investigated.

proceed in the investigation of the small-world property of the network. Almost all the studies that investigate path properties then go further and evaluate the small-world characteristics as well.

Half of the studies performs analysis considering the unweighted definition of graph, while the other half takes into account weights for investigating the graph. Crucitti *et al.* [32] use a weighted graph that anyway has no relationship with the physical properties of the considered Power Grid. The weight used is related to the betweenness managed by a node and it is partitioned between the edges it is connected to. Wang *et al.* [63] consider an impedance analysis therefore dealing with the physical properties of the lines, but the Power Grid graph are then not considered with the weighted definition, but only a probability distribution of the impedance is computed. Again Wang *et al.* [49] use a weighted representation of the Grid by using an admittance matrix which is based on the Laplacian matrix representation of the Power Grid enriched with the information about the admittance of the lines connecting the nodes. In [40], we perform a weighted analysis considering the resistance of cables as weights for the edges in the graph and we compute the same set of statistics for the unweighted and weighted definition of the graph. One reason for the lack of weighted Complex Network Analysis analysis on real samples of the Power Grid is probably due to the difficulty, first of having Grid data, and second of having the detailed information of the cables involved. The work of Mei *et al.* [41] is particularly rich and complete in the type and characteristics of the graph that are considered. Several models are considered based on the specific aspect of the Power Grid that are analyzed in the various sections of the book. For instance the evaluation of the small-world properties of the Power Grid and the power flow characterizing the Grid the graph is built considering the admittance values that characterize the lines. In another chapter which analyzes the vulnerability when considering the static behavior of the Power Grid the authors build a a directed graph with weights on the edges that represent the length of the lines. Again in another chapter describing the growth and evolution models of the Power Grid the graphs generated are considered undirected and unweighted. The weights considered in [43] represent the power flows that move through the lines. Weights are also considered in [45] where an impedance matrix representing the electrical properties of the physical network is used to assign weights to edges in the IEEE Bus structure used. In [48] weights are represented by the reactance of the transmission lines which are assigned as the values of the adjacency matrix representing the transmission lines. Another study that uses

13

weights assigned to the edges representing the impedence of the transmission line is [51] where two samples of the High Voltage Chinese Grid are analyzed. Other examples that take into account weights are [52, 53] where the the graph representing the network is considered directed, since the Power Grid is assumed in a steady state, and the weights are represented by the admittance matrix characterizing the lines. Guohua *et al.* [54] analyze the North China Power Grid and the graph model that is used considers a weighted and directed graph, associating weights both on nodes (amount of input power in the node) and edges (power transmitted on the transmission line). The work of Sun [55] is mainly a comparison of the first results obtained considering Complex Network Analysis techniques applied to the Power Grid field for the American Power Grid and two samples of the Chinese Grid (North and Center China Power Grid). The analysis performed by Kinney *et al.* consider once again the U.S. Power Grid used in [31], but this time the graph is considered weighted both for edges and for nodes. In particular a value of efficiency in the range of [0,1] is assigned to the edges representing their ability to transmit power; for nodes a value of capacity is provided which is proportional to the betweenness of the same node. Arianos *et al.* [58] evaluate the vulnerability of two samples coming from the IEEE Bus literature (IEEE 30-Bus and IEEE 57-Bus). In their analysis the authors use impedences values and power distribution factor associated to the transmission lines to have a model closer in his behavior to the real electric flow. The model of the electrical flows considered is a direct current model which avoids much of the complexities behind the real energy flow exchanges, but that in a Complex Network Analysis study is sufficient to have a good idea and approximation of the power flows in the network. This same approach of characterizing transmission lines with their physical parameters representing impedances and power distribution factor is used by Bompard *et al.* [59] where the vulnerability results obtained with just topological measures are compared with a model that takes into account enhanced physical aspects as well is applied both to a synthetic network (a Bus with 34 nodes) and to the Italian High Voltage Grid. Hines *et al.* [60] consider for their reliability evaluation the IEEE 300-Bus and a considerable sample of the North American Eastern Interconnect (almost 30000 nodes and 50000 edges). Their aim is to evaluate the adequacy of pure topological measures compared to measures that take into account also the electrical and physical properties of the Grid. Therefore, the authors consider a simplification of the Grid that exploits the direct current model for calculating the power flow in the network. To build this model of power flow physical/electrical properties are added to the pure topological description of the Grid. In particular, reactance values are assigned to the links and at each node a voltage phase angle is attributed so that the power flow flowing through a transmission line can be computed knowing the amount of power flowing through the neighboring nodes. The study of Pepyne [61] focuses on the investigation of cascading effects in IEEE model buses (i.e., IEEE 57-Bus and IEEE 118-Bus) and on a small-world based network (200 nodes and 400 edges) using a DC power flow model to represent power flows to assess the lines affected in a cascading situation. Brummitt *et al.* [62] consider a small section of the South-East US Grid in order to evaluate the effects of enhanced connectivity of networks that are loosely connected. The two samples are about 400 and 500 nodes and they are connected with each other by just eight edges. The Indian Power Grid is explored in [64] where the very basic properties (e.g., average node degree, clustering coefficient) of the directed graph used to represent the eastern region Indian High Voltage Power Grid are investigated.

Summary:
From Table 2 a difference appears: the studies closer to a topological characterization uses unweighted representation of the edges of the Grid and consider always the node degree distribution in the analysis, since it is an important element to define the type of network under study (e.g., scale-free network). On the other hand, the studies that apply a weighted representation of the graph do not consider the node degree distribution statistics neither considering the unweighted definition of node degree nor using a definition that takes into account weights as proposed in [65]. This last aspect might worth to be considered since the degree distribution properties of the network might change the picture of the node degree distribution in comparison with unweighted studies.
Centrality measures are not often used, exceptions are [41, 49, 43, 52] and the few other that compute between distribution statistics, to identify the statistical distribution of critical nodes. More attention to centrality measures especially using weighted representation of the Power Grid graphs or models that provide the capacity or energy flows through the Grid might be beneficial in understanding the most critical nodes or lines in the power system. Another recurring theme in the Complex Network Analysis involving the Power Grid is the reliability analysis, and actually it is the main motivation that drives these kind of studies. In fact many works were performed after major blackout occurred, such as the North American black-out of 2003[2] or the Italian one of 2003[3] (e.g., [66, 67, 32, 33]) or anyway mention blackouts as the main motivation for the work. The fragility and resilience properties of the Power Grid has been the major reason of concern that has determined the focus of such Complex Network Analysis studies on the High Voltage network. In fact almost all studies consider the behavior of the Grid to various attacks to its nodes or edges.

3.3   The small-world property

The small-world property in network has received lots of attention starting with sociological studies [68, 9], but more recently with application of this concept and model to many more classes of networks [1, 24, 2, 69]. Among the studies analyzed small-world property investigation is performed by ten out of the thirty two. The various studies look for the satisfaction of the small-world property described by Watts [24] then together with Strogatz [1].

Holmgren [34] performs a comparison of the Nordic network with a random graph with the same number of nodes and edges. The results show an average path length for the sample double compared to the random graph, but a clustering coefficient almost one order of magnitude bigger than the random graph one, so the author concludes the Nordic Grid belongs to the small-world class network.

Casals *et al.* [36] perform a comparison between each sample of the European Grid and random graphs. The results show that the majority of samples satisfy the small-world conditions.

Watts [24] dedicates his book to illustrating the small-world phenomenon and a section is dedicated to the analysis of the Western States Power Grid of the American network. He notes that sparseness of the graph corresponding to the Grid violates one of the assumption usually necessary to have a small-world. Compared to other types of networks (e.g., social network of actors' interactions in movies), the clustering coefficient is quite small and the characteristic path length quite big, however, despite these characteristics, he states that the small-world property holds for the Western United States Power Grid. A statistic that emerges is the high fraction of edges that are also *shortcuts* (i.e., if the edge is removed the shortest path between the same edge's ends is increased more than 2) for the graph which is around 80%. It might be due to the sparseness of the graph which is implied by physical and economical limitations of adding lines to substations. Watts also notices that also the way the Grid has developed supports the small-world concept: many independent and disconnected Grids have been connected together with the aim of sharing and exchanging power excess between remote locations, enhancing reliability and efficiency. Watts also notices that the model underlying the Western States Power Grid is closer and better explained quantitatively from a *relational model* than a *dimensional model*. The former model consider the creation of edges as a function of the preexisting edges in the graph, as if the previous relationships between nodes were to a certain extent kept. The latter, on the other hand, considers the creation of edges as a function of the particular spatial location of the vertex and the physical distance to another vertex. The better explication by a relational model is quite surprising and counter-intuitive, but can be justified by the inability

---

[3]`http://news.bbc.co.uk/2/hi/americas/3152451.stm`
[4]`http://news.bbc.co.uk/2/hi/3146136.stm`

of the dimensional model to admit occasional global edges spanning across nodes otherwise very far apart. This last property is the key to keep the small-world characterization.

Wang *et al.* [63] also investigate the small-world properties for their samples and they state that the model proposed by Watts and Strogatz [1], is only able to capture some features of the Power Grid, since the Power Grid is sparsely connected compared to small-world networks. They also notice that the basic condition required by Watts and Strogatz's model is not satisfied by the Power Grids under test.

In our study [40], we perform an investigation about the small-world properties of the Medium and Low Voltage network comparing the sample topologies with random graphs with same order and size. The results show that this end of the network seems even less close to small-world properties than the High Voltage due to a general very small clustering coefficient.

In [41] the small-world concept is an essential aspect across the whole book. Actually the authors take as an assumption the fact that the High Voltage Power Grid is a small-world network. This is justified by the authors considering the type of evolution process that the Grid is subject to in its development. In fact, it consists in a higher connectivity in a "local-world" where nodes (substations and power plants) are physically closer, therefore forming tighter connections locally, and only few long distance links that help in reducing the average path length of the network. In the small-world model also load and physical parameters are applied to make the model closer to the real Power Grid system. This model is analyzed to investigate the vulnerability of the network and its disruption behavior.

Kim *et al.* [44] investigate the small-world property of several High Voltage samples in particular comparing the metrics of characteristic path length and clustering coefficient to the theoretical ones for random graphs. Many of the samples (MAPP, Nordel, KEPCO and WSCC networks) satisfy the small-world conditions. The authors also define a so called "small world-ness index" as $SW_I = \frac{\gamma/\gamma_{rnd}}{L_{CP}/L_{CP_{rnd}}}$; the more the index gets high the more the network gets close to a small-world. The comparison between effects of line removal shows that the topological measures of characteristic path length increases substantially (it almost doubles) between the situation that do bring to outages and the ones that do not. Therefore the claim of the authors is that static topological analysis can provide an insight about the faults on lines that might lead to outages problems.

Han *et al.* [48] perform a small-world analysis of two samples of the Chinese network. The small-world analysis applied to the Anhui Grid does not satisfy the clustering coefficient condition which is slightly higher than the one of a random graph with same *order* and *size*. On the other hand the other sample analyzed representing East China Grid completely satisfies the small-world conditions.

Ding *et al.* [51] in addition to compute the usual metrics to evaluate the satisfaction of the small-world property for the two Chinese Grid under investigation, concentrate their efforts on modifying and adapting the small-world model to consider a weighted graph. They adapt the clustering coefficient by weighting the traditional value of clustering coefficient by the ratio between the average impedance of the line in the neighborhood of the considered node and the average impedance of all the lines in the network. Instead of considering the Characteristic Path Length the authors use a similar quantity that can be applied also to a weighted condition which is the characteristic node number (the median of the mean of the number of node traversed for all the shortest paths connecting each vertex to all the others). The authors claim that the improved model better captures the real characteristics of the Grid and it shows more than the unweighted model the compliance to the small-world model in both the two samples analyzed.

Sun [55] evaluates the small-world properties in American, North and Center China Power Grid and the results show that all these samples can be considered examples of small-world networks. The conditions of small-world are satisfied for the three samples considered, but the Center China Power Grid actually is a borderline condition especially for the clustering coefficient which is just 4 times higher than a random graph with same *order* and *size*. The author stresses that although the Power Grid are systems that are developed in a careful manner and each step of development is planned in detail, still some characteristics of randomness remain which might be due to the many practical factors that involve the Power Grid plan. The author also remarks a known property of small-world networks that is the ability in easily spread information in the same network. This aspect in the Power Grid to a certain extent is not always beneficial since this also eases the spread of cascading failures that lead to Power Grid blackouts.

Pepyne [61] uses a small-world network model to evaluate the cascading effects of networks with different values of the rewiring probability parameter. The author also investigates the small-world properties of the IEEE 57 and 118-Bus by verifying the conditions for these type of network defined by Watts and Strogatz [1, 24]; the two IEEE samples satisfy such conditions.

Summary:
In general, the various studies tend not to have a common answer for the general question regarding the membership of Power Grid networks to the small-world group. It is indeed very specific to the samples analyzed and no conclusion can be drawn, this seems especially true for the High Voltage Grid, while the Medium and Low Voltage networks seem far from being a small-world network [40].

3.4 Node degree distribution

The degree of a node is a property to understand how many other nodes it is connected to. However, this information is not particularly important for big graphs since keeping track of each node degree may not be manageable, instead it is better to have a general idea of the statistics of the node degree. In particular, its probability distribution gives us some insights of the general properties of the networks such as the likely or unlikely presence of nodes with very high degree (sometimes also referred as hubs). Table 3 shows the main information about the degree distribution in the works that perform this study. The second column gives a general idea of the type of cumulative node degree distribution that is investigated in the articles under review. What is interesting is to fit the distribution to a class of curves. This is shown in the third column.

Work	Cumulative Node Degree Distribution Probability Type	Fitted Distribution
[31]	Exponential	$y(x) \sim e^{-0.5x}$
[32]	Exponential	$y(x) = 2.5e^{-0.55x}$
[33]	Power-law	$y_1(x) = 0.84x^{-3.04}$ $y_2(x) = 0.85x^{-3.09}$
[35]	Exponential	$y_1(x) \sim e^{-0.81x}$ $y_2(x) \sim e^{-0.54x}$
[36]	Exponential	$y(x) \sim e^{-0.56x}$
[37]	Exponential	$y(x) \sim e^{-0.61x}$
[39]	Exponential or sum of exponential terms	$y_1(x) = e^{-0.18x^2}$ $y_2(x) = e^{-0.21x^2} + 0.18e^{-0.25(x-4)^2}$ $y_3(x) = 0.96e^{-0.17x^2} + 0.25e^{-0.19(x-3.9)^2}$
[63, 46]	Sum of truncated geometricaland irregular discrete terms	$y_1(x) \sim f_1(x)$ $y_2(x) \sim f_2(x)$
[40]	Power-law (unweighted) and sum of exponential terms (weighted)	$y_1(x) \sim x^{-1.49}$ $y_2(x) \sim 0.15e^{-21.47x} + 0.84e^{-0.49x}$
[48]	Exponential	$y_1(x) \sim e^{-0.65x}$ $y_2(x) \sim e^{-0.58x}$
[55]	Exponential	$y(x) \sim e^{-0.5x}$

Table 3: Comparison of the node degree cumulative distribution probability functions.

As seen in the table, the results do not completely agree on the type of the distribution followed by the Power Grid networks, but generally they are close to an exponential decay. Figures 9 and 10 represent the fitted node degree cumulative distribution reported in the third column of Table 3. For [63, 46] presented in the table the functions $f_1(x)$ and $f_2(x)$ are not reported in the table for size reason, but in footnote.[1] The plots in Figures 9 and 10 give a general idea of the shape of the distribution. The charts have to be interpreted in a qualitative way since the details concerning the coefficients are not always available in the reviewed studies. In addition, for studies concerning multiple samples (i.e., [35, 36, 40]) averages between all samples, or particular significant samples have been chosen among the many available.

The investigation in [34] and the charts shown report a node degree distribution for the Western U.S. and for the Nordic Grid that both seem to follow an exponential distribution.

The results found in [45] show that the properties of the network (the IEEE 300-Bus blueprint with 300 nodes and 411 edges) change once the weights, in the specific case the impedance characterizing the lines, are considered in the network instead of the traditional topological ones. With impedance as weights on the edges the shortest paths that are computed are actually the ones that the current flow follows according Kirchoff's

---

[1] $f_1(x) = \sum_{x_i < x} 0.2269(0.7731)^{x_i} * \{0.4875\delta(1), 0.2700\delta(2), 0.2425\delta(3)\}$ $x_i = 1, 2, \ldots 34$
$f_2(x) = \sum_{x_i < x} 0.4085(0.5916)^{x_i} * \{0.3545\delta(1), 0.4499\delta(2), 0.1956\delta(3)\}$ $x_i = 1, 2, \ldots 16$
The $*$ symbol is here to be considered as the convolution operator and the δ is the Dirac delta function.

electrical laws. In fact, if the graph is rebuilt considering only the 411 node-to-node connections (which are the 411 paths with minimal cost in the original graph representation) with smallest impedance (defined as electrical connectivity) the resulting graph is dominated by few hubs, thus suggesting a scale-free topology when electrical connectivity is considered.

The charts shown in [49] that analyze the node degree centrality distribution give an idea about the type of relationship that the node have resembling a power-law. A similar tendency is shown in the chart representing the node degree centrality distribution taking into account the electrical parameters of the network (impedence of transmission lines).

Summary:

In general the various studies focusing on the High Voltage Grids agree on a statistical distribution for node degree that follows an exponential (or exponential based) distribution with characteristic parameters of the exponential curve that depend on the specific Grid analyzed. While High Voltage Grid have been quite extensively analyzed the Medium and Low Voltage Grids have not found much attention so far and a deeper and wider investigation needs to be performed in different geographies since the only study (i.e., [40]) is representative of the Northern part of the Netherlands. In addition, the distribution Grid will be the section of the Power Grid mostly impacted by Smart Grid technology [70].
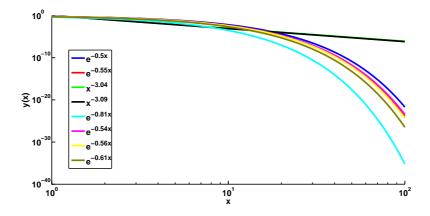
Figure 9: Log-log plot of fitted node degree cumulative probability distribution corresponding to the first six rows of Table 3.
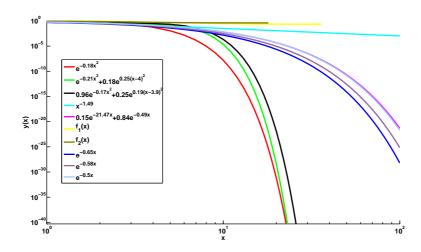

Figure 10: Log-log plot of fitted node degree cumulative probability distribution corresponding to the last five rows of Table 3.

3.5 Betweenness distribution

Betweenness is an important measure to assess how a node is central in a network. This metric in fact computes how many shortest paths traverse a node, therefore giving an information of the importance of the node in the path management. The main characteristics of the betweenness study are summarized in Table 4 where the second column shows the type of followed distribution, while the analytical function is represented in the third column. Unfortunately, this metric is computed by only five studies ([31, 32, 40, 48, 49]).

Work	Cumulative Betweenness Distribution Probability Type	Fitted Distribution
[31]	Power-law	$y(x) \sim (2500 + x)^{-0.7}$
[32]	Power-law	$y(x) \sim 10000(785 + x)^{-1.44}$
[40]	Power-law and exponential	$y_1(x) \sim x^{-1.18}$ $y_2(x) \sim 0.68e^{-6.8 \cdot 10^{-4}x}$
[48]	Power-law	$y(x) \sim x_1^{-1.71}$ $y(x) \sim x_2^{-1.48}$

Table 4: Comparison of the betweenness cumulative distribution probability functions.

Although the studies that perform this type of analysis are only few, one can see that there is a tendency for the High Voltage network to have a betweenness distribution close to a Power-law. For the Medium and Low Voltage the situation is less clear: some samples analyzed in [40] follow an exponential decay, especially the Low Voltage ones, while other, usually the bigger belonging to the Medium Voltage, follow a Power-law. In Figure 11 the plot of the distributions is represented to show the difference between the trend of the Power-law and exponential decay: after a certain point the exponential distribution has a faster decay.

The work of Hines *et al.* [45] reinforces the idea that betweenness has a Power-law trend in its probability distribution. In particular, the author analyzes betweenness in a condition where weights representing lines impedance are considered. The results show, even without providing an analytical representation, a tendency for this metric of being dominated by few nodes that support much of the weighted paths, while the majority of the nodes is only slightly involved. This suggests a Power-law distribution.

Wang *et al.* [49] investigate several centrality measures (i.e., degree centrality, eigenvector centrality, closeness, betweenness both on vertexes and edges). The normalized vertex betweenness whose chart is given in [49] suggests the presence of a Power-law although the analytical description of the best fitting curve is not given. The authors interestingly note that there is a very good accordance in this metric between the values computed with the electrical parameters (weighted analysis based on admittance matrix) and purely topological one; there is in fact an 80% overlap in the first 10 nodes most with highest betweenness.
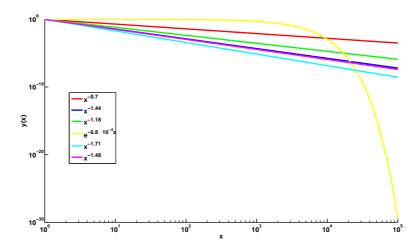

Figure 11: Log-log plot of fitted betweenness cumulative probability distribution corresponding to Table 4.

19

Summary:
Power-law seems the dominating rule for betweenness probability distribution even if few studies consider this statistical property of graphs. To draw a general definitive conclusion regarding this property for the Power Grid more studies are required. Another aspect to be considered which has not received much attention so far is the study of betweenness statistics in weighted Grid models or when power flows are considered instead of the pure topological analysis.

3.6 Resilience analysis

The characterization of resilience is the main motivation for the studies involving Complex Network Analysis and Power Grid. In fact, the behavior in terms of connectivity of the network when nodes or edges are removed is the primary question in many works considering failures that happen in a random fashion or following an attack strategy. Table 5 describes the different types of resilience analysis that are performed by the various authors. In particular, the second column contains the metric that is used to assess the reliability of the network. The focus of the attack either it is related to nodes or edges is considered in third and fourth column, while the fifth column remarks the studies that in addition to the resilience analysis also propose a mitigation strategy for improving Grid reliability.

Work	Resilience Analysis Type	Node attack	Edge attack	Grid improvement
[31]	Connectivity loss	✓		
[32]	Efficiency	✓		
[33]	Loss of load probability	✓	✓	
[34]	Influence on largest component size and path length	✓		✓
[35]	Robustness through mean degree, motifs and patch size analysis			
[36]	Influence on largest component size	✓		
[37]	Influence on largest component size e comparison with theoretical results	✓		
[38]	Damages and improvements		✓	✓
[39]	Nodes disconnection and improvements		✓	✓
[40]	Influence on largest component size	✓	✓	
[41]	Several criticality analysis and black-out models	✓	✓	
[42]	Reliability and disturbances			
[43]	Unserved energy/load	✓	✓	
[44]	Critical Path Length and clustering coeff.		✓	
[45]	Sensitivity			
[46]	Influence on largest spanning cluster size	✓		
[48]	Loss of load and failure endurance	✓		
[50]	Avalanche size	✓		
[52]	Flow availability		✓	
[53]	Efficiency		✓	
[54]	Largest power supply region	✓	✓	
[56]	Influence in network connectivity and power degradation		✓	
[57]	Efficiency	✓		
[58]	Efficiency, net-ability, overload		✓	
[59]	Efficiency, net-ability, overload		✓	
[60]	Path length, connectivity loss,	✓		
[61]	Line overload, cascade effects, network disruption		✓	✓
[62]	Overload, cascade effects, blackout size	✓		✓

Table 5: Comparison of the resilience analysis and improvement studies.

Node based attack analysis

Albert *et al.* [31] show the cascading effect of the whole American Power Grid when removing a certain fraction of nodes. In particular, the authors define the concept of *connectivity loss* which expresses the magnitude of the substations that cannot receive power from any generator due to failures in the network and thus inhibiting

the end users as well to receive any power. The connectivity loss assumes different levels of severity based on the type of nodes, i.e., transmission substations, that are removed. The random removal of these nodes has a limited impact on the connectivity loss, that increases almost linearly with the removal; the situation is completely different if the removal targets the nodes with highest betweenness especially when the betweenness is re-calculated after each new removal. In this situation, the connectivity loss increases harshly after a certain number of substations are compromised and just removing 2% of the nodes brings to a connectivity loss of about 60%; the behavior is the typical non-linear one that characterizes threshold effects.

Crucitti *et al.* [32] propose a model that considers *efficiency*, a measure inversely proportional to the shortest path. In particular for each node the authors define weights both for nodes (i.e., the maximum capacity a substation can handle, taking also into account a certain level of permitted tolerance or overcapacity) and for edges (i.e., the ability of delivering power for a certain transmission line). To study cascading effects after the removal of one node, the edge weights are re-calculated to investigate what effects on the system are triggered, and what is the new ability of edges in supporting paths. They show the cascading effects and the problems the network suffers in term of *efficiency* in different situations: random node removal and highest load-based removal. The results show a dissimilar behavior in the two situations especially when the tolerance parameter is low (i.e., the substations are considered to operate with small possibility to sustain more than the nominal capacity). Also in this case the worst results are experienced when the nodes with highest load are removed one after each other. Perhaps surprisingly, there is a non-perfect correlation between node degree and betweenness: it is not implied that the nodes with highest degree have always a high load.

Casals *et al.* [36] investigate the behavior of the European Grid under failures or attacks. The analysis shows a typical pattern already known for many other type of networks: the giant component (i.e., the biggest connected component) of the network vanishes after a threshold of nodes is removed and a phase transition occurs. The theoretical threshold that is computed for exponential node degree-based graphs well suits the empirical details that are extracted from the samples. Under random failures the decrease in the *order* of the giant component is continuous until the threshold is reached; under targeted attacks (e.g., removal of nodes with highest degree) the samples show a network disruption that appears when a smaller fraction of nodes is removed. A remarkable result the authors find is the relation between the *order* of the network and the *order* of the giant component: an increase in the number of nodes of the network makes it more prone to failures, but at the same time the way the networks have evolved tend to reduce their fragility.

Sole *et al.* [37] investigate the consequences of intentional attacks European Power Grid might face. The targeted attack problem is translated into an equivalent problem of random failures such that it can be studied with percolation theory, thus identifying a threshold of nodes to be removed to breakdown the giant component of the network. The fraction of nodes to be removed in targeted attacks towards highly connected vertexes is, as intuition might suggest, much smaller than in random attacks. The results show that there is generally an acceptable matching between the theoretical parameters found and the value extracted from the samples, nevertheless some deviations are present especially for those networks that are more robust. A remarkable aspect is that there is a correlation between the critical fraction removal and the most important reliability indexes used by power engineers to measure the Power Grid performances (e.g., energy not supplied, loss of power, interruption time). In fact, the two groups of European Grids that are considered (based on the similarity of the calculated and theoretical value of the threshold) although managing almost the same amount of power and energy, show very different results in terms of failures. The networks with threshold that deviates positively from the theoretical values are much robust and experience small reliability issues, thus reinforcing the correlation between non-topological reliability indicators and the topological ones.

Wang *et al.* extend forward in the analysis done in [63] and in [46] by studying the vulnerability properties of four synthetic IEEE Bus models and two real networks (NYISO and WSCC). They evaluate the robustness of the samples in two distinct situations: random node removal and targeted attack towards node with highest node degree. They evaluate the presence of a threshold value in the number of nodes removed before the network spanning cluster (i.e., a notion equivalent to the largest connected component of a graph) is fragmented. The authors reach the non-surprising conclusion that networks are more vulnerable if the nodes with the highest degree are attacked first compared to a random attack. The main finding is the definition of the equation that computes the critical fraction of nodes that destroy the spanning cluster in targeted attack conditions. In fact, they find that this value is influenced by the node degree distribution of the network. As a better fitting to the theoretical value of the critical fraction (threshold) the authors propose to consider the node degree distribution is modeled with a truncated geometric distribution plus an irregular discrete distribution [63], rather than an exponential one as computed for the critical fraction in [37, 36].

Han *et al.* [48] consider two elements to characterize problems in the Power Grid: load operating limit and the maximal load that a node is able to support. The reactance value of lines is adjusted accordingly when the operating limit is overtaken and the line is considered discarded if the load of the node to which it is connected is higher than the maximum supported value. The mechanism for evaluating the spread of failures in the network is by rerouting the paths when the nodes (and the edges attached to them) are not available. To show the impact of failures the authors use a loss of load metric which is given by the ratio between the load of nodes which are

no more in operation (since they have overtaken their limits) and the initial load supported by all the nodes. This loss of load metric is applied to evaluate the effects of two sorts of failure modes: nodes with high degree and high load are removed first and the effects are evaluated; nodes are divided in groups and average load for each group is computed, then the first 20 groups of nodes in terms of their load are removed, a second mode of evaluation is by random group removal and after each removal an increase in capacity factor of lines by 0.0125 is performed. These removal methods are applied to both the Anhui and East China Power Grids and they show for both networks that the spread of failures when attacking high degree nodes spreads quite slowly; on the other hand, when the nodes with high load are attacked the small-world network (i.e., East China Power Grid) appears particularly vulnerable and causes an important decrease in transmission capacity. The second type of attacks (based on group of nodes) evaluates the effects on the reduction of load by a group of nodes removal compared to the effects of the removal of the first group of nodes; this is to test the endurance capability of the network. The results show that Anhui Grid (non small-world) improves its performance thanks to the increase in capacity at each round, on the other hand the improvement is extremely limited in the East China Power Grid. The authors claim that the weaknesses of the small-world network compared to the other under test is due to the very nature of the network itself because of the high heterogeneity the nodes have in their load, with few nodes (those that are attached to the "shortcut" edges) having to sustain heavy load.

Wang *et al.* [50] consider the behavior of the network and the cascading failures issues when nodes with high and low load are attacked respectively. Conversely from other works the definition the authors provide in their work for load is obtained by a product of the node degree and the sum of the node degree of the neighbors of the same node amplified or lessen by a tunable parameter that represent the initial load of a node. In addition each node is characterized by a tolerance parameter which directly influence the capacity (maximal load) that the node is able to sustain. If the limit is overtaken than the node is considered broken and the load sustained is redistributed to its neighboring nodes proportionally to their initial load. A node that is under attack has to redistribute its load to its neighbors. The metric used in to understand the problem induced by an attack are measured through an avalanche size indicator that considers how many nodes reach the failure condition after an attack is performed. The authors consider two strategy for attacks: the first strategy involves nodes with highest load that are attacked in descending order and the second strategy attacks nodes starting from those with lowest load in ascending order. The authors find that, based on the initial load that is given to nodes, it is more likely that one strategy or the other is more critical in triggering avalanche failures. In particular, when the initial load is limited the cascade effects are triggered by the nodes with the lowest one, the situation is reversed when the initial load is bigger: highest load nodes attacks generate cascade effects behaviors. Typical of this work is the presence of a threshold in the tolerance of the lines: if the tolerance is above the threshold then the network do not experience avalanche effects, but if the threshold is even slightly overtaken the cascade effect arises.

Kinney *et al.* [57] perform a comprehensive analysis of the failure behavior of the U.S. Power Grid. The metric that is assessed is the damage that is inflicted to the network computed as the loss in average efficiency (defined as the average over all paths of the harmonic composition of the efficiencies of the edges composing the path) of the network before and after the breakdown has occurred. The mechanisms considered for the simulation of the failures are two: 1) single node removal and 2) progression removal of nodes. For each of the two schemes two different methods of node selections are used: a random choice of the node or nodes to remove and a targeted selection involving the nodes that support the highest load. The results found show that in low tolerance conditions (i.e., the capacity of nodes is very close to their initial load) just the removal of a node can cause a damage equal to 25%. In general the authors find three categories in which the nodes can be fitted: nodes whose removal causes no or very little damage independently of the tolerance parameter, and these are about the 60% of the overall nodes of the U.S. Power Grid; very critical nodes whose removal causes high damages in the network even when the tolerance is high; nodes whose removal causes damages depending on the tolerance of the network, usually a threshold value exists that is responsible for a steep transition from a little damage to an high damage situation, this last type of nodes are the majority in the network. Usually, nodes that cause critical effects if removed are also characterized by topological properties. In particular, nodes causing no efficiency loss have both low betweenness and low node degree: 90% of the transmission nodes that cause almost no damage to the network have 2 as node degree and load smaller than 2000. The main finding is once again that the removal of nodes with high betweenness and high node degree are the most critical and just the removal of 0.33% of nodes (those with such characteristics) generates more than 40% of damage in the network. The work suggests that changes in topology by adding more transmission lines and substation can help in decrease the sensitivity to failures of high load transmission lines.

Hines *et al.* [60] have the goal of comparing the vulnerability analysis obtained considering Grid models based only on topological properties and Grid models that take into account also the electrical/physical properties in order to assess the goodness of topological analysis in Power Grid vulnerability problems. To perform this comparison three metrics are evaluated, two of them are typical of the Complex Network Analysis literature and the third is more related to the electrical behavior of the Grid. In particular, the first metric assessed is the Characteristic Path Length so that as the components in the Grid fail the effort in the reachability of other nodes is evaluated compared to the initial situation. The second topological parameter investigated is the connectivity

loss (metric already used in [31]) which considers the percentage of generating nodes that are unreachable due to failures in the network. The third metric takes into account the blackout size as the amount of power load that has been curtailed as a consequence of the cascading failures in the Power Grid. In order to evaluate this last metric a model considering physical/electrical parameters and the direct current power flow is proposed (see description for [60] in Section 3.2 for the details of the model). The authors consider different types of failures that the network is subject to: random failures of nodes, degree-based attack which removes nodes one after each other starting with nodes with the highest degree. Another attack is based on the maximum traffic that a node manages in terms of power that flows through it and the attack targets the node with highest traffic; the opposite situation is also considered: minimum traffic attack that removes nodes starting with those having smallest traffic. The last attack considered takes into account betweenness and nodes with highest betweenness are removed first. Considering the path length metric the attacks based on degree, max-traffic and betweenness provoke grater damage in terms of path length compared to a random removal of nodes; the min-traffic has a behavior almost the same as random failures. The same considerations can be applied for the connectivity loss metric showing that the Power Grid is more vulnerable to directed attacks than random, a behavior thus more similar to scale-free networks. Generally the same considerations apply to the blackout size metric too, but it worths to highlight that just by attacking the 10 nodes with highest max-traffic creates a blackout with average size of 72%, while a random attack on 10 nodes on average results in a blackout of just 20% while the removal of 10 nodes with min-traffic results in 5% average blackout. Although the general intuition shows that attacks that target nodes with certain properties are more damaging than random or min-flow attacks it worths to remark that for the three metrics considered (i.e., path length, connectivity loss and blackout size) there is no accordance on the type of attack that causes the most damages: in the first case it is betweenness, in the second case it is node degree and for blackout metric it is max-traffic that provokes the most damage. The authors conclude that topological measures provide some indication of the vulnerability of a network, but they remark also that if taken alone they could be misleading especially if mitigation strategies on failure reduction are based on one topological result only. The authors also stress that physics-based models are more realistic and generally more useful for infrastructure risk assessment than topological only models.

Edge based attack analysis

Kim *et al.* in [44] replicate two blackout conditions that happened in the United States in 1996 by inspecting the topological properties of these networks during the cascade events that brought to the blackout. In particular, the authors simulate on the Power Grid graph the power line trimming process events that happened and other line removal events that do not lead to outages in the network. The metrics that they inspect to assess the vulnerability are two static topological metrics: clustering coefficient and characteristic path length; they also combine them into the "small world-ness index" as described in Section 3.3.

Dwivedi *et al.* [52] consider the vulnerability that an IEEE 39-Bus is subject to when the most central lines are removed. The centrality metric developed in this work considers the flow that power lines are able to support given their physical properties (admittance values): the lines that carry more flow (appropriately normalized to make the comparison fair between the power supported by the different source and sink nodes) are the most central. The authors claim that this method of investigating the most critical lines is more precise and accurate than the purely topological one, since it takes into account the amount of power that a certain line provides. However, the authors specify that some of the lines identified as important through this weighted centrality method are the same that can be identified with a pure topological investigation, thus stressing their importance also from a topological point of view. The same authors in [53] investigate the vulnerability of IEEE 39-Bus and IEEE 118-Bus. Even in this case they consider a weighted approach for the network characterized by his reactance parameters as weights. Betweenness values for the weighted networks are computed for each line so that it is available a ranking of the most critical or vulnerable lines. For each IEEE synthetic model two strategy of network disruption are applied: random line removal and removal based on the ranking of betweenness. The results are evaluated considering the efficiency of the network computed for each line as the inverse of the the reactance. The simulations show results which are not completely unexpected and that are common for Power Grid networks: a quite good robustness towards random attacks (small reduction of network efficiency); the situation changes when the most critical lines are attacked for which there is a considerable reduction in the network efficiency (60% reduction for the 39-Bus and almost 40% for the 118-BUS) by just removing the 10 most critical edges.

Pahwa *et al.* [56] consider the effects of cascading effects in synthetic models when lines are removed. In particular in their analysis they categorize the links in two types: vulnerable and non-vulnerable depending on the damage they create in the network (more or less than 10% upon their removal). The metrics that are computed to assess the damages are the influence in network connectivity which is purely topological and the Power Degradation which is the fall in the total load supported by the network compared to the original supported load. The authors propose synthetic models of networks generated in random fashion with same *order* and *size* of the reference engineering literature (i.e., IEEE, WSCC) samples. The results show that the generated networks

achieve a new equilibrium situation after a cascading is initiated in less steps, while the literature samples stabilize at a later stage when nodes or load is removed.

Arianos *et al.* [58] analyze the vulnerability of two samples of the IEEE literature. The authors compute three set of metrics: efficiency already applied in Power Grid studies [38, 57], net-ability and overload. The net-ability takes into account the path from a generator node to a load. It is proportional to the power share of the considered path in transmitting power from generator to load and inversely proportional to the *electrical distance* (i.e., a function to weight the edges based on the path traversed, the power factor and the impedance of the involved lines) between generator and load. It turns out to be proportional to the power transmission distributor factor for the line and the impedance of the same line. An overload metric is computed by considering the sum over the whole network of the ratio between the power flow of lines and their power flow limit. The difference between the overload condition before and after the removal of a line is the assessed parameter. The test that is performed on the network consists of line removal: for each line removed separately the three metrics mentioned above are computed. The main finding of the analysis is that the metric based on efficiency do not show much difference in the reduction of efficiency of the network when different lines are removed from the network, thus it does not provide much information about the criticality of a line. On the other hand, the net-ability and the overload metric present peaks values when specific lines are removed, thus being more precise to localize potential failing lines. Using the same set of metrics just described the critical lines analysis is applied to the a small synthetic Bus (34 nodes) and to the Italian High Voltage Grid in [59]. The main idea is to show how different are the results when metrics based on pure topological parameters are applied compared to the results obtained when also the physical characteristics of lines and nodes are also taken into account. The analysis performed on the effects of a single transmission line removal both in the synthetic sample and in the Italian Power Grid show that the net-ability metric captures better the lines whose fault results in worse disrupting conditions for the network compared to the efficiency metric. The same conclusion i.e., better adequacy of the metrics that take physical quantities into account is obtained also considering the *entropic degree* compared to the traditional node degree to rate importance of nodes. *Entropic degree* is defined considering not only the number of connections that a node has, but also the weights that the lines attached to the node have considering their impedence value [59]. The conclusion is that pure topological node degree may underestimate the real importance the nodes. It is interesting to remark that the authors of [58, 59] claim that the critical lines identified by using the net-ability metric are the same that the Italian Grid operator identifies as critical using engineering and electrotechnical approach to Power Grid failure analysis.

Pepyne [61] evaluates the cascading effects on IEEE blueprint networks. The author uses a DC flow model to characterize the network and to assess the cascading effects: each line is characterized by an impedance and at each node with generating capabilities a power flow injection is randomly assigned as well as the load nodes demand. In the study the line with the highest load is then brought to the 99% of its trip point and subsequently overloaded to study the evolution of power redistribution process and eventual subsequent failures of other lines. The condition the author evaluates to consider the cascading effects are two: 1) the cascading sequence proceeds till the network breaks apart into disconnected islands; 2) the cascade is self-limiting and stops before breaking up the grid. The simulation on the IEEE 57-Bus sample shows that just after 4 lines experiencing outages, the network breaks apart. The author evaluates also the behavior of small-world networks while increasing the rewiring probability of edges from a situation close to a regular lattice (rewiring probability $p = 0.001$) and a situation closer to a random graph (rewiring probability $p = 0.5$). The author finds that with an increase in rewiring probability (i.e., which corresponds to an increase in the number of so-called "shortcut" edges) there are two effects: 1) the load of the network increases and the line congestion decreases (measured as the mean and maximum line loading); 2) the network is more fragile in cascading failures than regular topologies since it tends to break apart just after few lines are in failure state. The author also tests a solution based on the upgrade of lines that in the IEEE Bus model act as kind of "shortcut" links whose capacity is doubled. The simulations with improved lines show that the average load of lines reduces thus being able to accommodate more load on average, on the other hand since the maximum loadability of lines is higher, more load has to be redistributed in case of failures, therefore the overall effect is a slightly more fragile Grid. The author finally suggests possible improvements for the real Grids: adding new "shortcut" links and increase the capacity of the existing "shortcut" links.

Node and edge based attack analysis

Chassin *et al.* [33] define a failure propagation model that considers failure probability for nodes and failure probability for edges. For the latter, although the model is based on an undirected graph, the propagation of failures happens in a directed way, i.e., the propagation evolves when the power flows from the node and the node supports the edge. The two probabilities of failure are then combined together to determine the *Loss Of Load Probability*. The results obtained in term of Loss of Load Probability are similar to the ones obtained for other electrical systems by electrical engineering studies, supporting the validity of the model.

In [40], we perform a reliability analysis considering different node removal strategies: random, node degree-

based, betweenness-based. We evaluate the *order* of the largest connected component while nodes are removed. The results obtained show a general robustness of Medium and Low Voltage networks to random attacks while they are extremely vulnerable to targeted attacks. The study also investigates the reliability regarding attacks towards edges exploiting the Laplacian matrix spectrum method to identify the most critical edges.

The book [41] focuses on evaluating and modeling the reliability of the Power Grid considered as a system which is characterized by a self-organized criticality behavior and modeling the behavior of the Grid that drives to blackout situations. The vulnerability assessments performed are several. To evaluate the reliability of the network the authors consider the static behavior (i.e., in steady state conditions) of the Power Grid when lines are removed. The authors develop a set of indexes to evaluate 1) the impact the faults has on active power transmission 2) the influence of the fault in local balance of reactive power. The first index is called *average electric transmission distance* and considers the ratio between the weighted transmission path length (a weighted path where the weights are the product of active power transmitted and physical length of the line) and the total power received by nodes; this index is a global measure that describes the influence of network structure on transmission of active power. The second pair of indexes takes into account the local balance in reactive power in particular the first index considers the difference in the voltage experienced before and after the fault, the second considers the variation in the reactive power supplied by generators before and after the fault compared to the maximal reactive power they can provide. The global and local metrics are then combined into a single vulnerability index. Using this combined index the authors then assess the vulnerability of the transmission lines of IEEE-30 and IEEE-118 Bus systems and a snapshot of Shanghai Power Grid of 2006. They also compare the results obtained through the method proposed in traditional power electrical analysis related to lines criticality; the results show a good agreement in identifying the most vital lines for the operation of the Grids. In another section of the book an investigation of vulnerability more in line with traditional high level Complex Network evaluates the behavior of Power Grid generated following the small-world model proposed by the authors for Power Grid evolution. They consider several types of attacks on nodes (random node, static node degree, dynamic node degree which is recalculated after each removal, node betweenness) and on lines (random line, line betweenness). The metrics they evaluate are similar to the ones of other works such as efficiency in transmission calculated as the ratio between the sum of the reciprocal of the shortest path between any two nodes after the failure has occurred and before the problem has taken place. All these types of attack and the metrics computed are applied to small-world generated networks, a 300 Bus network, IEEE-30 and IEEE-118 Bus systems. The main distinction between the small-world and IEEE networks lies in a better reliability of small-world networks towards random failures compared to IEEE, on the other hand IEEE Bus systems report better scores when the attacks are targeted (i.e., betweenness of degree based). The book also models the blackout situations occurring in a power systems considering a model based on DC (direct current) power flow and AC (alternating current) from a more electrical (power flow equations and power system constraints are applied) than topological point of view. These models are respectively applied to the High Voltage Power Grid of North East China and to IEEE-30 Bus system.

Bompard *et al.* [43] considers betweenness as the metric to investigate in analyzing the vulnerability of Power Grids. Unlike the other studies [31, 32, 40, 41] which consider betweenness only from a topological point of view, the author builds a graph which has on each edge a weight that represents the value of power flow that is managed by the line; the betweenness computed with such a graph is called electrical betweenness. The metric assessed to evaluate the removal of nodes, or edges, with highest electrical betweenness is the *unserved energy*. The results obtained by removal of edges, or nodes, considering the electrical betweenness are more critical compared to the ones found in computing topological betweenness since the quantity of unserved load is higher in the first case. The author therefore claims that electrical betweenness is a more correct indicator than topological betweenness when assessing vulnerability of Power Grids.

Guohua *et al.* [54] evaluate the vulnerability of the North China Power Grid. As the metric to evaluate the effects of the disruption the authors propose the largest power supply region index which is obtained by the ratio between the maximum load that the system can handle after a cascading failure has occurred, and the initial total load of the system. To assess this metric different disruption policies are applied. Random failures happening on lines are simulated and the probability law governing the failures of the lines is a Poisson distribution. A second type of attack involves nodes and they are attacked in descending order of node degree. Third type of attack considered deals with nodes that have the highest vulnerability index: in the weighted and directed model of network the authors consider, each node has a supported input power that gives it a certain weight, but also the lines to which it is attached (that have a transmission power limit which is the line weight) are an important part to define the node vulnerability index which turns out to be the product of the node weight and the line weight. The last policy considered in the paper is a targeted attack on lines based on their descending line vulnerability index which is equal to the betweenness value for an edge when the graph is considered weighted and directed. The results on the North China Power Grid following these attacks show that the Grid supports quite well random failures of lines since no cascading effects happen. On the other hand, the selective attack on high degree nodes are more damaging increasing the spread of damages in the network. Even more critical for the Grid are the attacks based on node vulnerability index and line vulnerability index which generates serious

disruption in the network by leading to cascading effects.

Other type of robustness analysis

Casals *et al.* [35] find that each single Grid composing the European network has a characteristic parameter that is related to its specific *robustness*. The authors investigate three main elements to assess the robustness of the different Grids: mean degree distribution, motifs analysis and patch size distribution. The first indicator is related to the comparison of the average degree of each Grid sample with the average degree obtained by the best fitting Poisson distribution (Poisson distribution for node degree arises from a random graph). The assumption is that when the number of nodes is sufficiently big the exponential tail and the Poisson one are very similar; the authors suggest that, the more the average node degree deviates from the Poisson distribution value, the most fragile is the network. This might seem counter-intuitive, but the networks that deviate more from the random topology are more fragile. The second idea is the investigation of motifs (i.e., linear, stars, triangles) inside the network. The conclusion the authors draw out is that fragility increases as the elements of the Grid become more interconnected and motifs such as stars and triangles begin to appear. Patch size distribution is the third element taken into account to characterize the relation between topology and robustness of network. The results suggest that a balanced distribution i.e., having a constant frequency of patches with different areas, characterizes more robust networks. The ideas and results presented are original and interesting even if, as the authors explicitly mention, a more deep analysis and comparison with Grid's dynamical results are needed.

Rosas-Casals [42] performs a slightly different type of study to evaluate the statistics of the European High Voltage Grid. It evaluates the statistical distribution of three metrics that are common in the assessment of the performance of electrical systems adopted by electrical engineers: energy not supplied, total loss of power and restoration time. Information about the statistics of failures are obtained by the ENTSO-E (European Network of Transmission System Operators for Electricity). Based on the results of his previous study of European Grid classification as robust or fragile networks [35], the author considers the statistical distribution of the three reliability measures investigating the presence of Power-law characteristics in their distributions. The presence of Power-law distribution is plausible only for the restoration time statistics. One of the interesting findings is that robust grids accumulate less failure events than the fragile ones, but the events that strike robust networks lead usually to important consequences, on the other hand fragile networks have more frequent problems but their consequences are usually limited.

Hines *et al.* [45] consider a different metric in assessing the vulnerability of the network that is based on the physical properties (line resistance) of the network: sensitivity. Sensitivity, which is defined as a ratio between a resistance of a path from one node to any other node and the resistance of the node itself, expresses the extent to which a state change in one node propagates to another node. The results show that sensitivity decades linearly with the distance between nodes, but the more the network has a scale-free structure the more the number of nodes are involved. For the IEEE 300-Bus the results suggest a behavior in the decrease in sensitivity with distance between nodes similar to the one of a random graph with same number of nodes and edges.

The vulnerability analysis performed by Sun [55] is actually a description of results obtained by other studies (e.g., [67, 57, 71, 72]) on Complex Networks and Power Grid studies on the behavior of networks under failures of nodes or edges.

Studies with infrastructure improvement analysis

Holmgren [34] inspects the structural vulnerabilities of Power Grids considering the decrease in size of the largest component of the graph. In particular, the author compares the resilience to failures of the Nordic Grid, the Western U.S. Grid, a random graph and a Barabási-Albert scale-free network. The comparison is considered both with random node removal and with targeted removal focusing towards nodes with highest node degree and proceeding in decreasing order; the size of the largest component is the evaluated parameter. The simulation shows that removing nodes following descending node degree, especially recalculating after each removal, brings the highest disruption in the network. In particular, the Nordel network seems less robust than U.S. Grid that in turn performs close to a scale-free network for these targeted attacks. The author also shows the evolution of the average path length when nodes are randomly removed, highlighting a similar behavior of the U.S. and Nordic Grids, both dissimilar from the behavior of the Barabási-Albert and random graph models. The author then proposes a virtual scenario of a very simple Grid for which he analyzes different improvement strategies and different types of potential attacks (both natural circumstances and human-driven sabotages) involving different type of nodes and edges. The author also shows the decrease in vulnerability that different strategies of network improvement may bring.

Crucitti *et al.* in [38] aim to detect the most critical lines for Italian, French and Spanish High Voltage Grids and to propose solutions to possible vulnerabilities. They use *efficiency* as main metric to evaluate the performance of each Grid, that is, a function of the shortest path connecting two nodes. Based on this metric the vulnerability of the Grid is measured in terms of efficiency loss experienced after a damage is inflicted to the network. By computing metrics on experimental data, the authors find significant differences between the Grids.

The Italian Grid is the most vulnerable: by removing just one edge (of course the most critical) the decrease in efficiency is of 5%, while for the Spanish and the French ones the reduction is only of 3%. With this approach it is possible to go forward and identify the most critical n-edges that if removed simultaneously decrease efficiency. The worst case occurs when three lines are removed together as experienced by the Italian Grid whose efficiency decreases by 31%. In addition, the resulting Grid is then broken into two subgraphs. The authors also consider the improvements that might be brought by the addition of one new line to the Grid, the Italian Grid is the one getting the greatest benefits (8% in efficiency increase). The paper also shows the results of the probability distribution of causing a certain damage (decrease in relative efficiency) for the Grids. The differences in the shape of the distribution between French, Spanish and the Italian are significant: French and Spanish have a Gaussian shape, while the Italian damage distribution is bimodal (obtained by the sum of two Gaussian with different means), in particular, the second peak corresponds to the situations in which the removal of edges brings to the complete break of the network in two subgraphs.

Rosato *et al.* [39] use the spectral properties of the Laplacian matrix considering the second smallest eigenvalue and corresponding eigenvector to identify edges that are extremely critical. For the Italian Grid the authors find the edges, only three, that if removed break the network apart, while for French and Spanish Grids the value is seven. This result is in line with the findings described in [38]. The authors also compute the conditional probability of having a certain number of disconnected nodes from the Grid once a prefixed number of edges is removed at random. The resulting measure is a sort of robustness index that indicates the level of vulnerability of the network. The results show the most critical is the Italian Grid for which the removal of just two edges leads to the probability of disconnecting more than two nodes that is around 5.8%, while for instance the same situation for the French Grid has only a probability of 0.3%. This is due to the specific morphology of Italy that spans across a long peninsula. A solution to improve the vulnerabilities of the Italian Grid is to add an edge between two strategic nodes, then the number of edges to be removed to break apart the network in two components raises to twelve. In this work, unlike [37], the geographical structure seems to play an important role and it is a key to understand certain properties.

Brummitt *et al.* [62] analyze the cascading effect happening on a Power Grid sample of the South-East of the United States using a sand pile model to generate cascades. In synthesis sand pile model assumes that each node is characterized by a sand threshold it can support (e.g., maximum load that it can support) and at each time step a grain of sand is randomly dropped on a node; this added sand can trigger the overtaking of the threshold for a node which is considered in a failure state and its sand is redistributed among its neighbors. This redistributed sand (e.g., load) can then trigger other failures in a cascade effect. For more details on the sand pile model refer to [73, 74]. The analysis in [62] is centered on the effects that connectivity has in networks. In particular, the authors focus both on randomly generated network, with similar node degree distribution to real Power Grids and on two real Power Grid, samples. The peculiarity of the samples, which is imposed also in the synthetically generated networks, is the loose connectivity between the two Grids: the two samples of the US Grid are actually connected by only 8 edges. The authors find that adding connectivity between the two loosely connected samples is not always beneficial. In particular, an optimal connectivity exists that minimizes the chance of cascading spread. Above this threshold of connectivity, cascades, especially of large dimensions, are likely to spread across the entire connected network. The situation is even more complex when the networks loosely connected have a disparity in the load they support: for one network more external connectivity would be beneficial mitigating his internal cascades, while for the other the increased connectivity would pose more threats. This situation creates the necessity for one Grid to increase its capacity in order to diminish the possibility of inflicted cascades originating in the neighboring Grid, thus producing a situation of disequilibrium difficult to solve (e.g., each Grid might aim at achieving his local optimal situation without considering the effects on the neighbor).

Summary:
In general, the reliability is assessed by evaluating the connectivity or the ability to efficiently guarantee paths between nodes when nodes or edges in the network are removed. As a general result for failures related to nodes all the samples show a good resilience to random breakdowns. In fact, the network is always able to guarantee a certain connectivity until the number of nodes removed are the biggest part of it. On the other hand the Grids are extremely vulnerable to targeted attacks, that is failures that focus on key nodes for the entire network such as high degree nodes or nodes with high betweenness or nodes or lines that manage the highest amount of load or electricity flow. The thoroughness and space given to the vulnerability and blackout topics in [41] covering also aspects more related to the electrical domain than pure Complex Network Analysis, provide a view of the reliability of the power system that is more precise and complete. This last aspect (i.e., using also physical/electrical parameters in the analysis) is stressed also in other works [58, 59, 60, 41, 45] which provide motivations and results justifying the superiority of vulnerability analyses that have a more realistic view of the Power Grid network. Adding physical parameters to the network is beneficial to have results concerning the way networks tend to disrupt and spread failures closer to reality. However, an aspect that is missing is the cross-check to real problems experienced by the transmission operators on the lines or nodes of the network identified as most critical and the resulting cascading effects obtained by a pure topological analysis or a topological analysis enriched with physical parameters and constraints characterizing the lines and nodes.

3.7 Further studies

Some of the studies considered in this survey provide additional analyses and investigations that are not common to any other study and thus deserve special attention. In this subsection we describe the most interesting additional aspects of some works for the sake of providing a more complete and thorough picture of Complex Network Analysis approach to the Power Grid.

Mei *et al.* [41] consider the whole electrical system as a Complex System: a system that is characterized by a self-organized criticality. The study does not only deal with a pure Complex Network Analysis study (as many of the other here surveyed), but it provides more insights into the peculiarity of the electrical Grid and systems. In particular, the book goes further in analyzing and modeling the blackout with particular regard to the dynamical aspects of the power system such as the power flow distribution and load redistribution. In addition to this enhanced perspective the book provides a section dedicated to modeling the growth and evolution process of the High Voltage Power Grid which is a topic which has deserved small attention so far by Complex Network Analysis studies.

In his statistical analysis of reliability of Europeans Grids Casals [42] goes deeper in analyzing the causes reported by transmission operators for the failures experienced. The data shows that only about 5% of the failures are caused by overload of the lines, while about 30% of the failures is classified as "unknown" (the other causes of problems are categorized as external and failures which correspond respectively to almost 30% and almost 40%). As remarked by the author this last cause of accidents need to be better understood and categorized in order to have a better understanding of the phenomena that trigger failures in the Power Grid.

Wang *et al.* [63] take into account also the electrical properties of the network considering the admittance matrix expressed as a combination of network adjacency matrix and line impedances. The finding of this study on line impedances shows for this property a probability distribution with heavy tails (i.e., some long lines resulting in very high impedance) that is best fitted to a double Pareto lognormal distribution that experiences an exponential cut-off in the tail part. This is due to economical aspects of construction and maintenance of long lines. The authors define a model to generate power networks that is based on creating small size sub-networks, then connect them together in a sort of lattice topology and finally generate the impedances following a certain distribution (e.g., double Pareto lognormal).

In [40], we point out an important difference resulting in the comparison of the unweighted and weighted study of the Medium and Low Voltage Power Grid: a general increase in the number of nodes traversed by following the shortest paths between the unweighted and weighted situation. However, the most interesting part of the study deals with an integration of all the parameters investigated with the Complex Network Analysis to provide a measure that associates the topological quantities to the economical aspects that might influence the spread of a distributed energy exchange market. We combine the topological measures to have two quantities: one represents the losses experienced in the network (a function of the weighted paths and nodes traveled) and the other represents the reliability and redundancy of the network (a function of the disruption behavior and of

the redundancy of available paths). With these indicators we show which network samples are more appealing for a distributed energy exchange market.

In [75] a mainly modeling study is conducted on the cascading failure behavior of networks whose connectivity properties are modeled using the principles of percolation theory. The model used does not take into account real or synthetic Grids to apply the findings, while it bases on the assumption that the Power Grid can be modeled by a random geometric graph. The aim of the work is to give the analytical conditions (unlike other works that focus on the empirical application on real or synthetic Power Grid networks) concerning the existence and non-existence of the largest connected component in a network when subject to link failures which are node degree dependent. The author proofs under which conditions a cascading failure exists and when links can be considered *vulnerable*, *reliable* and *unreliable*.

The study of Sun [55] has more the flavor of a survey and a state of the art description than a new analysis.

Pahwa *et al.* [56], in addition to the vulnerability propose three mitigation strategies in order to improve the network behavior against cascading effects. 1) a load reduction (by a given percentage) in the load supported by each node in the network after the initial failure, so that nodes and links operating below maximal capacity can deal with the redirect load; 2) an idea similar to 1) in the sense of load reduction, but this is applied only to the nodes along the tree that has his root in the node connected to the faulted line and reaches the leaves in the path; 3) considers the situation of isolating a cluster of the network in case of a starting cascade event with the assumption that the cluster can be self-sufficient in terms of energy production (e.g., by using renewable sources). The results shown are extremely encouraging in both the number of nodes that remain in the network (that are not involved in the cascade) when the load is preemptively slightly reduced.

For the sake of completeness, we also mention studies about CNA which have a minor focus on the Power Grid. Most often, the Power Grid is used as a possible example. In particular, Amaral *et al.* in [76] show a study of the Southern California Power Grid and the model follows an exponential decay for node degree distribution. Watts and Strogatz in [1] show the Small-world phenomenon applied to the Western States Power Grid while Newman, within a more general work [27], shows the exponential node degree distribution for the same Grid as an example. On the other hand, Barabasi *et al.* [77] model the Power Grid as a scale-free network characterized by a Power-law node degree distribution.

4 Discussion and Conclusion

The survey of CNA studies and their comparison shows how important properties of a real system such as the Power Grid can be studied using graph modeling tools and which conclusions about the reliability of the infrastructure can be drawn. Complex Network Analysis proofs to be an excellent set of tools that provide, although without dealing with the details and complexities of the electrical properties in the case of the Power Grid, a comprehensive and general understanding of the properties that characterize a network. We see an interesting trend in the various works analyzed, that is, the research towards more complex representation of the properties of the network than a simple graph. In fact, although Complex Network Analysis can help in understanding the foundational properties of the network of the Power Grid it is always worth to remember that the Power Grid is subject to the law of physics and the principles of electrotechnics. From the initial studies (e.g., [24, 31, 33, 34, 55]) considering the Power Grid just as an indirected graph without any property (i.e., weight) on edges and with no characterization of the nodes, more recent studies take into account the electrical properties of the Power Grid system. Of course the aim of these later studies is always to provide anyway a simplification of the highly complex Power System, but they add those essential parameter to better simulate the Grid behavior: impedance parameters associated to the transmission lines, power limits supported by the substations (i.e., nodes) such as for instance in [41, 59, 60, 54, 49]. The inclusion of a more detailed description of the Power Grid under investigation enables to better understand the dynamics guiding the Power Grid with a mixed approach: both preserving the idea of the Complex Network Analysis of having a general and statistical behavior of the overall behavior of the Power Grid, and, on the other hand, to take into account the physical/electrical properties essential to characterize the Power Grid. Latest results in [60, 59, 58] show a better agreement to real Power Grid behavior of the models that take into account physical parameters, compared to the pure topological analysis, and the observed behaviors and critical points in real power systems thus justifying this *enhanced Complex Network Analysis* approach.

A noteworthy general aspect is the role that Complex Network Analysis has in the Power Grid infrastructure vulnerability analysis: CNA does not want to substitute the traditional approaches to Power Systems resilience and safety analysis since they have proved extremely successful in governing and managing the it with only occasional highly disruptive events. We stress once again that Complex Network Analysis techniques applied to the Power Grid world are high simplification of all the complexities governing the Power Systems, but it can anyway be helpful to give a general vision that can help in identifying quickly and in a simple way critical spots or aspects of the Power Grid which then may be further and deeply analyzed with traditional electrical engineering tools.

A recent emerging function that CNA is acquiring deals with the design aspects of the Power Grid [78]. In fact, CNA can be used in simulating new Grid topologies, assessing the strategies to evolve actual topologies to improve them, especially in the panorama of the forthcoming Smart Grid that might require a reshape of the Grid especially in the Medium and Low Voltage part.

The samples almost entirely belong to the High Voltage network and are part either to the American Grid (or subsamples of it) or to the European one (or part of it) or Chinese Grid. Generally, the node degree distribution tends to follow an exponential distribution with some minor exceptions, allowing a general characterization of the properties of the network by the average degree parameter. Betweenness studies are also interesting to characterize the criticality of certain nodes as essential for the ability of the network to guarantee its navigability. Unfortunately, the computation of betweenness is done only by a small part of the studies under assessment. However, the tendency is to have at least for High Voltage samples a behavior that is closer to a Power-law distribution for the probability characterizing this metric. It means having very few nodes with very high values that are responsible to allow the majority of the shortest paths across the network. On the other hand the results from the different studies contrast regarding the small-world phenomenon in Power Grid networks. Indeed, some of the conditions imposed by Watts and Strogatz [1] are not met by Power Grid samples due to physical and economic reasons. This property must depend on the specific sample analyzed. The geography of the country whose sample is derived is sometimes important (e.g., Italy [39]) while for other studies it has lower impact [37]. A point of agreement between all the studies is about the reliability of the Power Grid networks when facing failures. A general good resilience to random breakdown, while extreme vulnerability is experienced by attacks that target the critical nodes (i.e., high node degree or high betweenness nodes).

An important result for the accuracy of the Complex Network Analysis studies is the similarity of results that this type of analyses give compared to the traditional electrical engineering results [33, 37, 58, 59]. This really shows how the theoretical study and the measured quantities in the real environment are very close.

Of course to enable Complex Network Analysis to be a more useful tool in the Power Systems domain, especially in order to create models that take into account the physical parameters of the network data and information related to the electrical parameters, it is essential for CNA and Power Grid scientists to have precise and reliable information of the topology and characteristics of the electrical Grid. As remarked by Rosas-Casals [42] in order to better model and more precisely explain the behavior of the Power Grid, high quality and precise information are required about the Grids, their components, the actual flow of energy data the Grids deal with in their operations, the properties of the power lines and also information for those part of the Grid which are held privately in the hands of Grid operators and distribution operators. Security and criticality of this data is a recognized aspect in a world that fears terrorist attacks, but without data feeding the models and research works it is not possible to discover actual vulnerability points and help to build more secure and reliable Power Grids.

Networks are an integral model of phenomena surrounding us, may these be biological networks (e.g., food-webs, protein interactions) or human generated ones (e.g., airline travel routes, computer chip wiring, telephone call graphs) [79]. Having methods and tools to better understand them and their dynamics is beneficial for knowledge advancement and better design of future systems. Complex Network Analysis is such a modeling technique that provides methods and metrics for an analytical comprehension of network behaviors. Public infrastructures are important for today's society, in particular the Power Grid, which is by nature a complex network, has a critical role for the economy in every country. Having an overall view of the Power Grid as a Complex System gives the ability to assess the potential issues the electrical system may face due to topological failures. In this paper, we have shown what are the main studies conducted on different Power Grid networks using metrics and techniques from the emerging CNA field of study. Although the basic methodology of study is the same, indeed the Complex Network Analysis techniques, the results show differences in some properties such as node degree distribution statistics, the presence or absence of the small-world property. On the other hand, the commonality is the behavior of the Power Grid networks when facing failures: a general good resilience to random breakdown, while they show extreme vulnerability when facing attacks that target the critical nodes (i.e., high degree or high betweenness nodes, nodes or lines that manage the highest electricity flow). The morphology of the country definitely influences the the topology of the network and thus its properties and reliability, symbolic is the case of the Italian Grid [39].

To have a more complete idea of the Power Grid networks it is worth investigating other Grids from other countries, not only limiting oneself to networks in Europe and United States and China as the studies analyzed so far focus on. The investigation of Grids belonging to other geographies such as Asia and South America could lead to new topologies. On the other hand, it is important to study more samples belonging to the Medium and Low Voltage Grids as to the best of our knowledge the only study in this direction is our own [40]. This is interesting not only because it highlights some different properties from the High Voltage, but also because it can provide indications useful for the design of the future Smart Grid. The analysis we performed also takes into account the weights representing physical properties of the Grid and is therefore more informative about transport capabilities of the distribution network. In addition, Complex Network Analysis can be used not only as a tool for the analysis of the Grid, but also to consider how the electrical Grid might evolve according to design

principles to be optimized at a topological level [63, 78]. It is also interesting taking into account the influence of the network topology on electricity distribution costs for the future scenarios of Smart Grid solutions [40].

Acknowledgment

References

[1] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, no. 6684, pp. 440–442, June 1998.

[2] S. H. Strogatz, "Exploring complex networks." *Nature*, vol. 410, no. 6825, pp. 268–76, March 2001.

[3] A. L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, p. 509, 1999.

[4] R. Albert, H. Jeong, and A. L. Barabási, "Error and attack tolerance of complex networks," *Nature*, vol. 406, no. 6794, pp. 378–382, Jul. 2000.

[5] P. Erdős and A. Rényi, "On random graphs. I," *Publ. Math. Debrecen*, vol. 6, pp. 290–297, 1959.

[6] ——, "On the evolution of random graphs," in *publication of the mathematical institute of the Hungarian academy of sciences*, 1960, pp. 17–61.

[7] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, and A. L. Barabási, "The large-scale organization of metabolic networks." *Nature*, vol. 407, no. 6804, pp. 651–4, Oct. 2000.

[8] J. Doye, "Network Topology of a Potential Energy Landscape: A Static Scale-Free Network," *Physical Review Letters*, vol. 88, no. 23, pp. 1–4, May 2002.

[9] J. Travers and S. Milgram, "An experimental study of the small world problem," *Sociometry*, vol. 32, no. 4, pp. 425–443, 1969.

[10] W. Aiello, F. Chung, and L. Lu, "A random graph model for massive graphs," *Proceedings of the thirty-second annual ACM symposium on Theory of computing - STOC '00*, pp. 171–180, 2000.

[11] M. Faloutsos, P. Faloutsos, and C. Faloutsos, "On power-law relationships of the internet topology," in *Proceedings of the conference on Applications, technologies, archit ectures, and protocols for computer communication.* ACM, 1999, p. 262.

[12] R. Albert, H. Jeong, and A.-L. Barabási, "Internet: Diameter of the World-Wide Web," *Nature*, vol. 401, no. 6749, pp. 130–131, September 1999.

[13] D. Donato, L. Laura, S. Leonardi, and S. Millozzi, "The Web as a graph," *ACM Transactions on Internet Technology*, vol. 7, no. 1, pp. 4–es, 2007.

[14] J. Kephart and S. White, "Directed-graph epidemiological models of computer viruses," in *Research in Security and Privacy, 1991. Proceedings., 1991 IEEE Computer Society Symposium on*, May 1991, pp. 343 –359.

[15] V. Colizza, A. Barrat, M. Barthélemy, and A. Vespignani, "Predictability and epidemic pathways in global outbreaks of infectious diseases: the sars case study." *BMC Med*, vol. 5, p. 34, 2007.

[16] A. Gautreau, A. Barrat, and M. Barthélemy, "Global disease spread: statistics and estimation of arrival times." *Journal of theoretical biology*, vol. 251, no. 3, pp. 509–22, Apr. 2008.

[17] V. Latora and M. Marchiori, "Is the Boston subway a small-world network?" *Physica A: Statistical Mechanics and its Applications*, vol. 314, no. 1-4, pp. 109–113, 2002.

[18] R. Guimerà and L. a. N. Amaral, "Modeling the world-wide airport network," *The European Physical Journal B - Condensed Matter*, vol. 38, no. 2, pp. 381–385, Mar. 2004.

[19] V. Colizza, A. Barrat, M. Barthélemy, and A. Vespignani, "The role of the airline transportation network in the prediction and predictability of global epidemics." *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 7, pp. 2015–20, Mar. 2006.

[20] M. Boss, H. Elsinger, M. Summer, and S. Thurner, "The network topology of the interbank market," *Quantitative Finance*, vol. 4, pp. 677–684, 2004.

[21] M. Anghel, K. A. Werley, and A. E. Motter, "Stochastic model for power grid dynamics," *Hawaii International Conference on System Sciences*, vol. 0, p. 113, 2007.

[22] B. Carreras, V. Lynch, M. Sachtjen, I. Dobson, and D. Newman, "Modeling blackout dynamics in power transmission networks with simple structure," in *HICSS '01: Proceedings of the 34th Annual Hawaii International Conference on System Sciences (HICSS-34)-Volume 2.* IEEE Computer Society, 2001, p. 2018.

[23] I. A. Hiskens and M. Akke, "Analysis of the nordel power grid disturbance of january 1, 1997 using trajectory sensitivities," *IEEE Transactions on Power Systems*, vol. 14, no. 3, Aug 1999.

[24] D. J. Watts, *Small Worlds: The Dynamics of Networks between Order and Randomness.* Princeton, NJ, USA: Princeton University Press, 2003.

[25] B. Bollobas, *Random Graphs*, W. Fulton, A. Katok, F. Kirwan, P. Sarnak, B. Simon, and B. Totaro, Eds. Cambridge University Press, 2001.

[26] R. Durrett, *Random Graph Dynamics (Cambridge Series in Statistical and Probabilistic Mathematics).* New York, NY, USA: Cambridge University Press, 2006.

[27] M. E. J. Newman, "The Structure and Function of Complex Networks," *SIAM Review*, vol. 45, no. 2, p. 167, 2003.

[28] M. Kochen, *The Small World.* Greenwood Publishing Group, 1989.

[29] X. Shi, M. Bonner, L. Adamic, and A. Gilbert, "The very small world of the well-connected," in *Proceedings of the nineteenth ACM conference on Hypertext and hypermedia.* ACM, 2008, pp. 61–70.

[30] R. Cancho, C. Janssen, and R. Solé, "Topology of technology graphs: Small world patterns in electronic circuits," *Physical Review E*, vol. 64, no. 4, pp. 1–5, Sep. 2001.

[31] R. Albert, I. Albert, and G. Nakarado, "Structural vulnerability of the north american power grid," *Physical Review E*, vol. 69, 2004.

[32] P. Crucitti, V. Latora, and M. Marchiori, "A topological analysis of the italian electric power grid," *Physica A: Statistical Mechanics and its Applications*, vol. 338, no. 1-2, pp. 92 – 97, 2004, proceedings of the conference A Nonlinear World: the Real World, 2nd International Conference on Frontier Science.

[33] D. P. Chassin and C. Posse, "Evaluating North American electric grid reliability using the Barabási Albert network model," *Physica A Statistical Mechanics and its Applications*, vol. 355, pp. 667–677, Sep. 2005.

[34] A. J. Holmgren, "Using Graph Models to Analyze the Vulnerability of Electric Power Networks," *Risk Analysis*, vol. 26, no. 4, pp. 955–969, August 2006.

[35] M. Rosas-Casals and B. Corominas-Murtra, "Assessing European power grid reliability by means of topological measures," *Transactions ob Ecology and the Environment*, no. 121, pp. 515–525, June 2009.

[36] M. Rosas-Casals, S. Valverde, and R. V. Solé, "Topological vulnerability of the European power grid under errors and attacks," *International Journal of Bifurcation and Chaos*, vol. 17, no. 07, p. 2465, 2007.

[37] R. V. Solé, M. Rosas-Casals, B. Corominas-Murtra, and S. Valverde, "Robustness of the European power grids under intentional attack," *Physical Review E*, vol. 77, no. 2, pp. 1–7, Feb. 2008.

[38] P. Crucitti, V. Latora, and M. Marchiori, "Locating critical lines in high-voltage electrical power grids," *Fluctuation and Noise Letters*, vol. 5, no. 2, pp. L201–L208, 2005.

[39] V. Rosato, S. Bologna, and F. Tiriticco, "Topological properties of high-voltage electrical transmission networks," *Electric Power Systems Research*, vol. 77, no. 2, pp. 99–105, 2007.

[40] G. A. Pagani and M. Aiello, "Towards decentralization: A topological investigation of the medium and low voltage grids," *Smart Grid, IEEE Transactions on*, vol. 2, no. 3, pp. 538 –547, sept. 2011.

[41] S. Mei, X. Zhang, and M. Cao, *Power Grid Complexity*. Springer, 2011.

[42] M. Rosas-Casals, "Power grids as complex networks: Topology and fragility," in *Complexity in Engineering, 2010. COMPENG '10.*, feb. 2010, pp. 21 –26.

[43] E. Bompard, D. Wu, and F. Xue, "The concept of betweenness in the analysis of power grid vulnerability," in *Complexity in Engineering, 2010. COMPENG '10.*, feb. 2010, pp. 52 –54.

[44] C. J. Kim and O. B. Obah, "Vulnerability assessment of power grid using graph topological indices," *International Journal of Emerging Electric Power Systems*, vol. 8, no. 6, 2007.

[45] P. Hines and S. Blumsack, "A centrality measure for electrical networks," in *Hawaii International Conference on System Sciences, Proceedings of the 41st Annual*, jan. 2008, p. 185.

[46] Z. Wang, A. Scaglione, and R. Thomas, "The node degree distribution in power grid and its topology robustness under random and selective node removals," in *Communications Workshops (ICC), 2010 IEEE International Conference on*, may 2010, pp. 1 –5.

[47] A. Nasiruzzaman and H. Pota, "Critical node identification of smart power system using complex network framework based centrality approach," in *North American Power Symposium (NAPS), 2011*, aug. 2011, pp. 1 –6.

[48] P. Han and M. Ding, "Analysis of Cascading Failures in Small-world Power Grid," *International Journal of Energy Science*, vol. 1, no. 2, pp. 99–104, 2011.

[49] Z. Wang, A. Scaglione, and R. Thomas, "Electrical centrality measures for electric power grid vulnerability analysis," in *Decision and Control (CDC), 2010 49th IEEE Conference on*, dec. 2010, pp. 5792 –5797.

[50] J.-W. Wang and L.-L. Rong, "Cascade-based attack vulnerability on the us power grid," *Safety Science*, vol. 47, no. 10, pp. 1332–1336, 2009.

[51] J. Ding, X. Bai, W. Zhao, Z. Fang, Z. Li, and M. Liu, "The improvement of the small-world network model and its application research in bulk power system," in *Power System Technology, 2006. PowerCon 2006. International Conference on*, oct. 2006, pp. 1 –5.

[52] A. Dwivedi, X. Yu, and P. Sokolowski, "Analyzing power network vulnerability with maximum flow based centrality approach," in *Industrial Informatics (INDIN), 2010 8th IEEE International Conference on*, july 2010, pp. 336 –341.

[53] ——, "Identifying vulnerable lines in a power network using complex network theory," in *Industrial Electronics, 2009. ISIE 2009. IEEE International Symposium on*, july 2009, pp. 18 –23.

[54] Z. Guohua, W. Ce, Z. Jianhua, Y. Jingyan, Z. Yin, and D. Manyin, "Vulnerability assessment of bulk power grid based on complex network theory," in *Electric Utility Deregulation and Restructuring and Power Technologies, 2008. DRPT 2008. Third International Conference on*, april 2008, pp. 1554 –1558.

[55] K. Sun, "Complex networks theory: A new method of research in power grid," in *Transmission and Distribution Conference and Exhibition: Asia and Pacific, 2005 IEEE/PES*, 2005, pp. 1 –6.

[56] S. Pahwa, A. Hodges, C. Scoglio, and S. Wood, "Topological analysis of the power grid and mitigation strategies against cascading failures," in *Systems Conference, 2010 4th Annual IEEE*, april 2010, pp. 272 –276.

[57] R. Kinney, P. Crucitti, R. Albert, and V. Latora, "Modeling cascading failures in the north american power grid," *Eur. Phys. J. B*, vol. 46, no. 1, pp. 101–107, 2005.

[58] S. Arianos, E. Bompard, A. Carbone, and F. Xue, "Power grid vulnerability: A complex network approach," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 19, no. 1, p. 013119, 2009.

[59] E. Bompard, R. Napoli, and F. Xue, "Analysis of structural vulnerabilities in power transmission grids," *International Journal of Critical Infrastructure Protection*, vol. 2, no. 1-2, pp. 5–12, May 2009.

[60] P. Hines, E. Cotilla-Sanchez, and S. Blumsack, "Do topological models provide good information about electricity infrastructure vulnerability?" *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 20, no. 3, p. 033122, 2010.

[61] D. Pepyne, "Topology and cascading line outages in power grids," *Journal of Systems Science and Systems Engineering*, vol. 16, pp. 202–221, 2007.

[62] C. D. Brummitt, R. M. DSouza, and E. A. Leicht, "Suppressing cascades of load in interdependent networks," *Proceedings of the National Academy of Sciences*, 2012.

[63] Z. Wang, A. Scaglione, and R. J. Thomas, "Generating Statistically Correct Random Topologies for Testing Smart Grid Communication and Control Networks," *IEEE Transactions on Smart Grid*, vol. 1, no. 1, pp. 28–39, Jun. 2010.

[64] V. Chaitanya, M. Reddy, and D. Mohanta, "Topological analysis of eastern region of indian power grid," in *Environment and Electrical Engineering (EEEIC), 2011 10th International Conference on*, may 2011, pp. 1 –4.

[65] M. E. J. Newman, "Analysis of weighted networks," *Phys. Rev. E*, vol. 70, no. 5, p. 056131, Nov 2004.

[66] R. Albert, I. Albert, and G. L. Nakarado, "Structural vulnerability of the North American Power Grid," *Phys. Rev. E*, vol. 69, no. 2, p. 025103, Feb 2004.

[67] P. Crucitti, V. Latora, and M. Marchiori, "Model for cascading failures in complex networks," *Physical Review E*, vol. 69, no. 4, pp. 3–6, 2004.

[68] S. Milgram, "The small world problem," *Psychology Today*, vol. 2, pp. 60–67, 1967.

[69] J. Kleinberg, "The small-world phenomenon: an algorithm perspective," in *Proceedings of the thirty-second annual ACM symposium on Theory of computing*, ser. STOC '00. New York, NY, USA: ACM, 2000, pp. 163–170.

[70] R. Brown, "Impact of smart grid on distribution system design," in *Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century, 2008 IEEE*, july 2008, pp. 1 –4.

[71] Y.-C. Lai, A. Motter, and T. Nishikawa, "Attacks and cascades in complex networks," in *Complex Networks*, ser. Lecture Notes in Physics, E. Ben-Naim, H. Frauenfelder, and Z. Toroczkai, Eds. Springer Berlin / Heidelberg, 2004, vol. 650, pp. 299–310.

[72] A. E. Motter and Y.-C. Lai, "Cascade-based attacks on complex networks," *PHYS.REV.E*, vol. 66, p. 065102, 2002.

[73] P. Bak, C. Tang, and K. Wiesenfeld, "Self-organized criticality: An explanation of the $1/f$ noise," *Phys. Rev. Lett.*, vol. 59, pp. 381–384, Jul 1987.

[74] ——, "Self-organized criticality," *Phys. Rev. A*, vol. 38, pp. 364–374, Jul 1988.

[75] H. Xiao and E. Yeh, "Cascading link failure in the power grid: A percolation-based analysis," in *Communications Workshops (ICC), 2011 IEEE International Conference on*, june 2011, pp. 1 –6.

[76] L. A. N. Amaral, A. Scala, M. Barthélémy, and H. E. Stanley, "Classes of small-world networks," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, no. 21, pp. 11 149–11 152, October 2000.

[77] A. L. Barabási, R. Albert, and H. Jeong, "Mean-field theory for scale-free random networks," *Physica A: Statistical Mechanics and its Applications*, vol. 272, no. 1-2, pp. 173–187, Oct. 1999.

[78] G. A. Pagani and M. Aiello, "Power grid network evolutions for local energy trading," JBI, University of Groningen, Tech. Rep. ArXiv preprint arXiv:1201.0962, 2012.

[79] A. L. Barabási, "Linked: The new science of networks," *American Journal of Physics*, vol. 71, no. 4, pp. 409–410, 2004.