Component Criticality in Failure Cascade Processes of Network Systems

Enrico Zio, Giovanni Sansavini

Component criticality in failure cascade processes of network systems

Enrico Zio[1,2,*] and Giovanni Sansavini[1]

[1] Energy Department, Politecnico di Milano, Milan, Italy

[2] Ecole Centrale Paris and Supelec, Paris, France

[*] Address correspondence to Enrico Zio, via Ponzio 34/3 – 20133 Milano, Italy; tel: 00390223996340; enrico.zio@ecp.fr, enrico.zio@supelec.fr, enrico.zio@polimi.it.

ABSTRACT: In this work, specific indicators are used to characterize the criticality of components in a network system with respect to their contribution to failure cascade processes. A realistic-size network is considered as reference case study. Three different models of cascading failures are analyzed, differing both on the failure load distribution logic and on the cascade triggering event. The criticality indicators are compared to classical measures of topological centrality, for identifying the one most characteristic of the cascade processes considered.

KEYWORDS: Complex Infrastructures Vulnerability, Cascade Failures, Criticality Indicators, Network Models, Centrality Measures

1 INTRODUCTION

Cascading failures are a major threat to distributed, interconnected systems such as power transmission networks[1], communication systems[2], transportation systems[3], social networks[4] or even metabolic networks[5]. These types of failures are usually initiated when a heavily loaded component of the system fails, and its load is redistributed to other components; the redistribution may cause the load on other components to exceed their capacity causing them to fail or protection mechanisms to shut them down to prevent further failures.

Models and simulation tools have been devised to describe cascading failure processes in network systems; yet, substantial challenges are still open in this research field, e.g. related to[1]: i) the origins and implications of the distributions of cascade sizes, ii) the dependencies between the initiating events and the successive cascading events, iii) how to design protections for avoiding the propagation of cascading failures after their outbreak.

In this study, some of these issues are looked at within an abstract modeling paradigm for analyzing the system response to cascading failures, which can be used to guide a successive detailed simulation focused on the most relevant physical processes and network components. The need for such an analysis tool is even stronger for systems in which the cascade dynamics is rapid and modifications are actuated onto the network in order to mitigate the evolution of the cascade. For example, in electrical power transmission networks a cascade of events leading to blackout usually occurs on a time scale of minutes to hours and is completed in less than one day[1].

In this paper, three different models of cascading failures are considered differing for the logic of redistribution of the failure load[6,7,8]. Two scenarios for the cascade triggering event are considered, i.e., either a random failure or a targeted intentional attack. The modeled

cascade spreading process is followed step by step and indicators are evaluated for each component, such as the frequency of its participation to a cascade, the average time before its entrance into a cascade, the average duration and final size of the cascade emerging from its failure. The criticality of the components identified by the indicators of their contribution to the development of cascading failures is associated to classical measures of topological centrality for the three considered scenarios[9,10,11,12,13].

The contents of the paper are organized as follows: the three models of failure propagation in network systems are presented in Section 2; in Section 3, the indicators of component criticality in cascading failures are introduced; in Section 4, they are computed for a realistic-size network of electrical power transmission; in Section 5, the results obtained are compared with topological centrality measures and the most relevant to the cascade process is identified. Conclusions are drawn in Section 6.

2 MODELS OF CASCADING FAILURES IN NETWORK SYSTEMS

In the following, the three models of cascading failure processes considered are briefly presented[7].

2.1 *Local propagation of a fixed amount of load*

Consider a network system of N identical components with random initial loads sampled uniformly between a minimum value L^{min} and a maximum value L^{max}. All components have the same limit of operation L^{fail}, beyond which they are failed. When a component fails, a fixed and positive amount of load P is propagated locally to first-neighbors of the failed component in the network. If there is no working component in the neighborhood of a failed component, the cascade spread in that "direction" is stopped.

To start the cascade, an initial disturbance imposes on each component an additional load D. If the sum of the initial load L_j of component j and the disturbance D is larger than a component load threshold L^{fail}, component j fails. This failure occurrence leads to the redistribution of an additional load P on the neighboring component which may, in turn, get overloaded and fail within a cascade which follows the connection pattern of the network system. As the components become progressively more loaded, the cascade continues.

The algorithm for simulating the cascading failures proceeds in successive stages as follows:

0. At stage $i = 0$, all N components are initially working under independent uniformly random initial loads $L_1, L_2, \ldots, L_N \in [L^{min}, L^{max}]$, with $L^{max} < L^{fail}$.
1. An initial disturbance D is added to the load of each component.
2. Each unfailed component is tested for failure: for j =1, ..., N, if component j is unfailed and its load > L^{fail} then component j fails.
3. The components loads are incremented taking into account the network topology, i.e., the failed component neighborhood: for each failed component, the load of its first-neighbors is incremented by an amount P. If the neighborhood set of the failed component is empty, the associated failure propagation comes to an end.
4. The stage counter i is incremented by 1 and the algorithm is returned to step 2.

The algorithm stops when failures are not propagated further.

The cascade propagation algorithm is embedded in a Monte Carlo simulation framework, in which a large number of cascades, e.g. 10000 in this study, is triggered for the same range of initial load, $[L^{min}, L^{max}]$, in order to obtain statistically significant results for various realizations of the same average loading condition. The damage caused by the cascades for any initial load level, $[L^{min}, L^{max}]$, is quantified in terms of the number of network components which have failed on the average, i.e. the average cascade size, S. It is assumed that each

system operates in such a manner that the initial component loadings vary from L^{min}=0 to L^{max}=L^{fail}=1. Thereafter, the average component loading L = (L^{min}+1)/2 is raised by increasing L^{min} from 0 to 1 at steps of 0.005. The average component loading, L, provides information on the initial working conditions of the network in which the cascade is triggered. Large L values represent highly-loaded systems, where each component is on the average operating close to its limit capacity, L^{fail}=1. The range of loading conditions is normalized from 0 to 1 so that the model for cascading failure is not limited to the propagation of failures in specific applications. The normalized loads describe the propagation of an `abstract' cascade of failures. Then, for specific applications, the range of loading conditions is scaled to describe the dynamics of the actual physical quantities spreading for the network being analyzed. Since varying loading conditions are explored only through the variation of L^{min} from 0 to 1, the average initial load, L, lies in the range from 0.5 to 1. L=0.5 describes the lowest possible loading condition in the network in which the components can assume all the possible loads between 0 and 1 with a uniform probability. Conversely, L=1 describes the highest possible loading conditions in which all of the components operate at their limit capacity and stop functioning in the next time step.

As the simulation is repeated for different ranges of initial load, [L^{min}, L^{max}], with L^{max} = 1 and L^{min} varying from 0 to 1, the pair (L, S) is recorded.

The transfer of a fixed amount of load P to other (neighboring) components of the network upon failure of one of its components may be representative of those systems where each node equally contributes to the global system activity and following their progressive failures the same amount of damage is caused to the still working ones. The original version of this probabilistic model[(14)] was developed to capture salient features of the cascading failures in large, fully-connected infrastructure systems. It was initially intended to reproduce loading-dependent cascading failures, such as the ones occurring in large blackouts of electric power

transmission systems[14]. Its scope was extended afterwards to investigate coupled infrastructure systems[15], e.g. power-communication, power-market, communication-transportation, and market-market systems. In this view, *P* should not be thought of as actually distributing the load of a component to the neighboring components; rather, one should think of *P* as an increased "stress" in the neighboring components due to failures in the network. As a biological analogue, it seems interesting to mention that in biological systems the death of a neural cell lead to the release of a toxin and this, in turn, is responsible for the death of many other cells[16].

2.2 *Local redistribution of the failure load*

In some systems and under some operating conditions, the transfer of a fixed amount of load *P* to other (neighboring) components upon a failure may not be the proper mechanism. It may be more realistic that the actual load previously carried by the now failed component is passed onto the other (neighboring) components in the network. This redistribution scheme is more suitable to characterize the redistribution of the electrical load that occurs in a power transmission network when some component is disconnected from the system due to overloads. To model such condition, step 3 of the cascade propagation algorithm in the previous Section 2.1 is modified as follows:

3. The components loads are incremented taking into account the network topology, i.e., the failed component neighborhood: given the generic component *j*, failed under load $L^*_j > L^{fail}$, its load L^*_j is spread uniformly among its neighbors, by incrementing their load of an amount equal to L^*_j divided by the present degree k_j of the failed component, i.e., the number of nodes to which component *j* is currently connected. If the operating neighborhood set of the failed node is empty (i.e., if there are no operating components connected to it), the failure propagation comes to an end.

By this modification, the load arising from the failure of a component is uniformly shared among its neighbors (it still holds that in case of an empty neighborhood, the load is no longer propagated and the cascade is stopped in that "direction").

2.3 *Redistribution of the load based on the shortest paths*

In the previous model, the loading of the components is independent of the connectivity pattern of the system, which only affects the load redistribution following a failure. This loading model may apply to systems like the power distribution networks, in which the load at each substation is independent on the number of overhead transmission lines injecting onto it. For other systems, like information networks, the load on a component, e.g. a router or a hypernode, can be modeled as dependent on the number of links transiting through it.

To model this situation, let us assume that at each time step one unit of the relevant quantity processed by the network, e.g. information, is exchanged along the shortest path connecting every pair of components; the load at a component is then the total number of shortest paths passing through that component[17,18]. At any instant of time, this load is to be compared with the component capacity which is the maximum load that it can process. In man-made networks, the capacity of a component is limited by technological limitations and economic considerations. For modeling purposes, it can be assumed that the capacity C_j of component j is dimensioned proportionally to its nominal load L_j at which it is designed to operate initially,

$$C_j = \alpha \cdot L_j \quad j = 1, 2, ..., N \tag{1}$$

where the constant $\alpha > 0$ is for simplicity assumed equal for all components. When all the components are working, the network operates without problems in so far as $\alpha>0$. On the contrary, the occurrence of component failures leads to a redistribution of the shortest paths

in the network and, consequently, to a change in the loads of the surviving components. If the load on a component increases beyond capacity, the component fails and a new redistribution of the shortest paths and loads follows, which, as a result, can lead to a cascading effect of subsequent failures. This model was employed to investigate cascading failures in many real-world networks, such as the Internet at autonomous system level[8], the electrical power grid of the western United States[8] and the IEEE power transmission test system[19].

The importance of the cascade effect with respect to intentional attacks stems from the fact that a large damage can be caused by the attack on a single component. Obviously, in general more links render a network more resistant against cascading failures, but this increases the cost of the network.

When looking at the potential of a cascading process triggered by the removal of a single component, two situations are expected: if prior to its removal the component is operating at a relatively small load (i.e., if a small number of shortest paths go through it), its removal will not cause major changes in the balance of loads and subsequent overload failures are unlikely; however, when the load of the component is relatively large, its removal is likely to affect significantly the loads of other components and possibly start a sequence of overload failures. Intuitively, the following behavior is expected[8]: global cascades occur if the network exhibits a highly heterogeneous distribution of loads and the removed component is among those with highest loads; otherwise, cascades are not expected.

In the modeling scheme adopted, the distribution of loads is highly correlated with the distribution of links: networks with heterogeneous distribution of links are expected to be heterogeneous with respect to the load, so that on average components with large number of links will have high loads. This behavior confirms the robust-yet-fragile property of heterogeneous networks, which was first observed in Ref. (20), with respect to the attack on several components.

3 INDICATORS OF COMPONENT CRITICALITY IN CASCADING FAILURES

The flow redistribution process is simulated at discrete time steps. At t_0 the network is intact; at t_1 a failure occurs; at t_i, $i \geq 2$ the cascading failure progresses as nodes overload and cause further failures in neighboring elements. The cascading process is followed until the response stabilizes; at this point, indicators of the severity of the cascade are computed. Four features of the cascade process have been analyzed for the different models.

The frequency of participation to a cascade, f_i, of every component i=1, 2,…N has been evaluated normalizing the number of its failures over the number of failure cascades simulated starting from different initial conditions (load disturbance or component attacked, depending on the model):

$$f_i = \frac{\text{\# of failures of component i}}{\text{\# of cascades simulated}} \tag{2}$$

This measure gives information about the importance of a component in the buildup of a cascade.

The average discrete time step in which a component i=1, 2,…N joins the cascade of failures, called the entrance time t_i, has been assessed averaging over the total number of cascades simulated:

$$t_i = \frac{\text{time when component i enters the cascade}}{\text{total \# of cascades simulated}} \tag{3}$$

This indicator is a measure of how early in time a component gets involved in a cascade process.

To catch how the failure of a component i=1, 2,…N causes other components to fail subsequently, the average duration measured as discrete time steps, d_i, and final size, s_i, of a cascade following the failure of component i have been evaluated through the same averaging procedure used for computing the average entrance time, t_i:

$$d_i = \frac{\text{duration of a cascade following the failure of component i}}{\text{total \# of cascades simulated}} \quad (4)$$

$$s_i = \frac{\text{final size of a cascade following the failure of component i}}{\text{total \# of cascades simulated}} \quad (5)$$

It is expected that the failures of more critical components will result in larger sizes of the developing cascades; furthermore, two different behaviors can be anticipated for the duration of the generated cascade: namely, depending on the cascade evolution, the failure of a critical component could lead either to the sudden failure of the remaining working components, with a very short cascade duration or to a long chain of delayed failures, resulting in a long cascade duration. In this sense, the final size of the cascade is considered a direct indicator of the criticality of a component whereas the duration measure by itself does not allow drawing clear-cut conclusions about the critical contribution of components to the cascading failure process.

It is important to stress once more that the two averages in the indicators (4) and (5) are taken with respect to the total number of cascades triggered in the system, to reflect the component average relevance to the cascade process.

4 CASE STUDY

The indicators of component criticality introduced in Section 3 have been computed for the topological network of the 380 kV Italian power transmission network (Fig. 1), considering cascades evolving according to the three failure propagation models of Section 2. The 380 kV Italian power transmission network is a branch of a high voltage level transmission, which can be modeled as a network of N=127 nodes connected by K=171 links[21,22], defined by its $N \times N$ adjacency (connection) matrix $[a_{ij}]$ whose entries are 1 if there is an edge joining node i to node j or 0, otherwise. Its topology is taken as reference but the

failure propagation models applied to it have very little specific to such system; it is only used so to give concrete examples of the findings.

Three models that capture the propagation of failures in interconnected systems have been reviewed in Section 2. They show that initial perturbations, e.g. a uniform overload *D*, even small, or the removal of one system component have the potential to trigger large, system cascading failures. As reported in Section 2, each one of these models has been applied to "abstractly" describe cascading failure processes in spatially-distributed, multi-component infrastructures, such as power distribution systems[8,14,15,19], telecommunication networks[8,15], transportation systems[15] and even market systems[15]. The simulation frameworks provided by these models abstract the physical details of the services delivered by the infrastructures, while at the same time capturing the essential operating features. In this respect, here we abstractly model cascading failures that propagate over the bare topological structure of a network, with little specific to the electrical service it provides. We assess the extent to which the criticalities identified by the different models relate to one another, and to classical measures of topological centralities for networks. We expect that different models identify a number of common criticalities. The bare topological structure of the 380 kV Italian power transmission network will be the sole responsible for these "invariant" criticalities in failure propagations.

In all simulations, the cascading failure evolution has been followed step by step, the relevant information collected and, eventually, the quantities **Error! Reference source not found.** - (5) have been computed.

Figure 1. The 380 kV Italian power transmission network[21,22].

4.1 *Local propagation of a fixed amount of load*

Table I shows the results for the cascade model relative to the local propagation of a fixed amount of load (Section 2.1). The values of the parameters which govern the outbreak and the propagation of the cascade, i.e. D and P, respectively, are set independently of each other and must be determined in the context of the specific failure propagation process to be represented. Different regions of the parameter space may originate completely different behaviors with respect to the propagation of failures. Large D values result in many component failures at the initial stage of the cascade; large P values facilitate the successive spreading of the failure cascade to a large number of components, since a large overload (P) is transferred to the neighborhood of a component, upon its failure. In our work, the initial disturbance D and the load transfer amount P are heuristically set equal to 2% and to 7% of the failure load L^{fail}, respectively ($L^{fail} = L^{max} = 100\%$). This choice of values for the cascade parameters is intended to model systems in which small perturbations, D, have the potential to trigger cascading failures that can sustain themselves and affect the entire system, thanks to the large P overload value. Consequently, the majority of the simulations results in few initial failures at the outbreak of the cascade and successive propagation to a large number of components during the next steps. This occurs in particular when the system operates at high loading conditions. Ultimately, this choice of the cascade parameters allows the clear identification of the effects that the initial loading conditions, L, have on the propagation of the cascading failures.

Three out of the four component criticality indicators of Section 3, namely f_i, d_i and s_i identify the most critical components with respect to the different cascade features they measure. Components 64, 68 and 88 turn out to be the most critical with respect to f_i and d_i while according to s_i component 64 and 88 are less important than other components, e.g. 101 (Villanova in Fig. 1); this is due to the fact that the latter constitutes a bridge between

different loosely-connected subsets of components, namely between the Northern and the Southern Adriatic backbone, and thus functions as a channel for spreading the failure to regions of the system which are far apart.

The ranking agreement among the f_i and d_i indicators is somewhat unexpected since they are related to different cascade features, namely, the frequency of participation and the duration of the cascade.

It is interesting to note that the average entrance time t_i (Table I) is shortest for those components which least participate and contribute to the cascade development (e.g., 127, 117 and 116). According to this propagation model, following a failure, a small extra load is given to the neighboring components and, consequently, the cascade never affects the whole system, in particular sparing the less connected components, e.g., nodes 127, 117 and 116. Thus, the poorly connected nodes only enter the cascade soon after its initiation if either they are themselves triggering it or they reside in the neighborhood of a triggering node: this results in their small average entrance time in the cascade, t_i (Table I).

Table I. Summary of the criticality indicators rankings for the model of local propagation of a fixed amount of load; only the twenty-four most critical nodes are reported.

Node	f_i	Node	t_i	Node	d_i	Node	s_i
64	0.2828	125	1.909	64	4.524	68	29.31
68	0.2776	126	1.918	88	4.504	24	29.20
88	0.2750	124	1.943	68	4.487	115	28.94
67	0.2705	121	1.952	35	4.448	43	28.91
79	0.2686	123	2.007	67	4.443	7	28.88
35	0.2679	52	2.033	79	4.440	101	28.85
60	0.2668	55	2.035	60	4.435	3	28.83
75	0.2653	115	2.052	59	4.433	2	28.79
59	0.2646	3	2.071	98	4.423	21	28.67
81	0.2638	56	2.074	75	4.421	64	28.64
98	0.2630	2	2.080	103	4.406	88	28.61
63	0.2629	7	2.081	97	4.404	103	28.59
62	0.2617	120	2.085	43	4.396	35	28.57
103	0.2616	122	2.086	24	4.385	110	28.49
92	0.2607	24	2.094	81	4.382	79	28.35
97	0.2596	8	2.100	63	4.379	52	28.29
91	0.2591	113	2.101	14	4.376	28	28.27
41	0.2582	68	2.104	40	4.375	92	28.27

Node	f_i	Node	t_i	Node	d_i	Node	s_i
61	0.2581	54	2.104	101	4.370	55	28.26
14	0.2579	114	2.112	61	4.367	120	28.24
71	0.2577	127	2.115	27	4.366	47	28.23
40	0.2576	101	2.117	28	4.365	8	28.18
43	0.2569	119	2.120	7	4.360	125	28.18
78	0.2551	21	2.120	41	4.360	124	28.16

To validate the results against several initial conditions, a sensitivity analysis of the model and of the rankings provided by the four criticality indicators is carried out with respect to the initial disturbance *D*. The choice of this parameter is critical for the model because a large *D* should more easily induce simultaneous overloads in more nodes. Figure 2 portrays the effects of the propagation of failures in terms of the cascade size *S* against the initial loadings *L*, for eight different values of the initial disturbance *D*. The analysis is performed for values of *D* that span the entire feasibility range $D \in [0, 1]$. Larger *D* values result in a cascade with significant size even at low loading conditions.

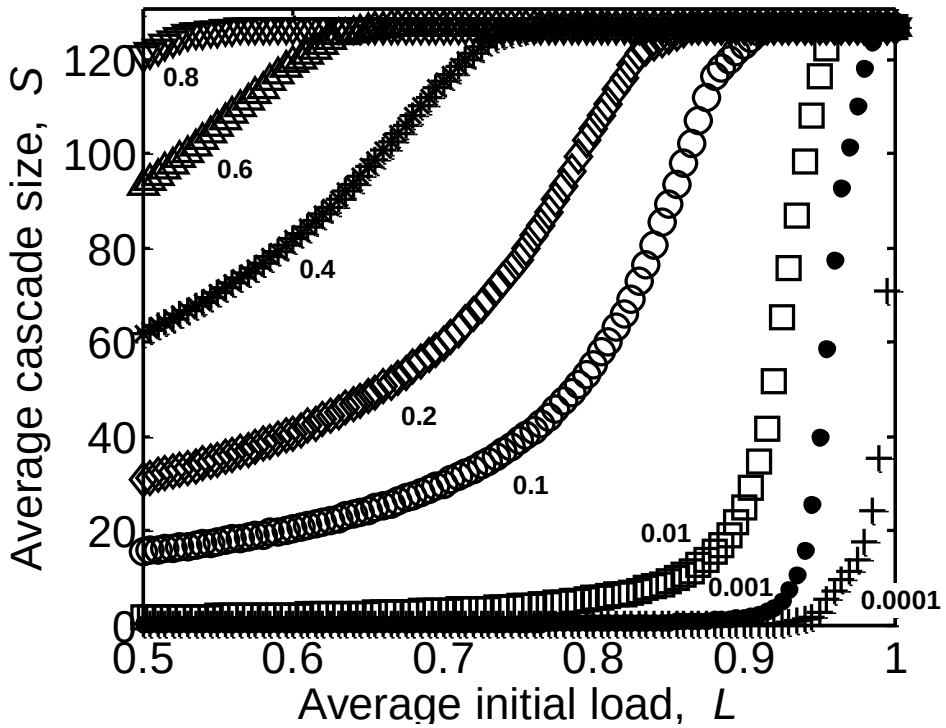

Figure 2. The average cascade size S vs. the average initial load L, for eight different values of the initial disturbance $D \in [0.8, 0.6, 0.4, 0.2, 0.1, 0.01. 0.001, 0.0001]$. Each point in the diagrams is averaged over 10000 cascades triggered for the same range of initial load $[L^{min}, L^{max}]$.

Rankings provided by the four criticality indicators are evaluated for every scenario involving different D. The similarities in the classification of the nodes criticalities are assessed to test the consistency of the indicators with respect to different conditions that initiate a cascade of failures. To this aim, the 127 items of the rankings are grouped in five batches [10, 35, 37, 35, 10] of decreasing criticality, i.e., the first batch encompasses the ten most critical nodes, the second batch encompasses the following thirty-five most critical nodes, and so on, until the ten least critical nodes, that are supplied in a risk-informed perspective. Then, the rankings provided by homologous indicators are compared looking for similarities in the batches of equal criticality and a degree of consistency is evaluated. For each one of the five groups of batches, the occurrences of the same node are enumerated (the first occurrence is not accounted for since the interest is on multiple occurrences of the same node in the rankings obtained for various D). The occurrences for different items are added up to build a synthetic index which then is scaled in the interval [0%, 100%] through division by the maximum number of repeated occurrences, i.e., the product of the size of the batch times the number of compared rankings decreased by one (to account for the first occurrences of the items): 100% similarity indicates that the same components are included in the batches for various D; 0% similarity indicates that for different D the batches of equal criticality contain diverse components. Table II shows the degree of consistency in the rankings provided by the four proposed indicators for nine scenarios with different D. f_i and d_i consistently rank the most and the least critical nodes in propagating the cascade of failures for the different scenarios. In particular, all the nine rankings for f_i consistently identify nodes

64 and 68 as the most critical ones. On the other hand, the rankings provided by t_i and s_i are more sensitive with respect to the values of the initial disturbance D. In particular, two different patterns are identified in the rankings provided by t_i and s_i, namely, for small D, i.e., $D \in [0.02, 0.01. 0.001, 0.0001]$ and for large D, i.e., $D \in [0.8, 0.6, 0.4, 0.2, 0.1]$.

Table II. Degree of consistency in the rankings provided by the four proposed indicators for nine scenarios with different D, i.e., $D \in [0.8, 0.6, 0.4, 0.2, 0.1, 0.02, 0.01. 0.001, 0.0001]$.

	Degree of consistency			
Batch	f_i	t_i	d_i	s_i
10	87.50	57.50	82.50	60.00
35	92.86	73.21	85.36	73.57
37	91.89	74.66	79.73	78.38
35	92.86	76.43	84.64	72.14
10	88.75	60.00	83.75	52.50

4.2 *Local redistribution of the failure load*

Table III, shows the results for the cascade model which redistributes the failure load onto the neighborhood of the failed node for the values of initial disturbance D = 0.02. All the component criticality indicators agree that 64, 68, 35, 59, 60 and 88 are most critical to the cascading process. In particular, nodes 59 (Piacenza) and 60 (Caorso) form a bridge between two densely connected areas in Northern Italy. In this failure propagation model, the amount of load transferred to the neighboring components after a failure is typically larger than in the previous case; this gives rise to a stronger coupling among components so that when a cascade is initiated it is more likely to fully develop and affect the whole system: thus, the components more prone to failure are the ones which are most connected.

It can be also noticed that the nodes which least contribute to the cascade process, nodes 69, 70, 87 and 26 according to f_i, are ranked as having the highest s_i; this is due to the fact

that if they get involved in a failure cascade this happens early in time, e.g. node 70 and 69 according to t_i, before the cascade has spread over a large portion of the system.

Table III. Summary of the criticality indicators rankings for the model of local redistribution of the failure load; only the twenty-four most critical nodes are reported.

Node	f_i	Node	t_i	Node	d_i	Node	s_i
64	0.7481	64	3.754	64	8.414	69	35.00
68	0.7358	70	3.876	35	8.277	70	35.00
35	0.7345	68	3.878	59	8.241	87	31.43
59	0.7295	35	3.879	60	8.197	26	30.15
60	0.7275	69	3.882	88	8.177	74	29.54
88	0.7250	88	3.905	68	8.174	1	29.28
24	0.7173	79	3.951	79	8.111	50	29.20
79	0.7170	59	3.952	14	8.110	4	29.00
43	0.7164	43	3.987	61	8.086	117	28.80
14	0.7159	60	3.999	43	8.053	57	28.06
61	0.7133	14	4.015	21	8.008	77	27.59
28	0.7122	110	4.038	62	8.004	116	27.41
67	0.7113	21	4.055	67	8.004	37	27.06
21	0.7111	87	4.062	63	7.991	19	27.00
27	0.7110	98	4.074	40	7.987	72	26.88
63	0.7100	67	4.084	98	7.972	93	26.61
62	0.7086	101	4.087	110	7.949	51	26.59
103	0.7041	75	4.110	75	7.937	54	26.56
75	0.7039	97	4.115	97	7.932	94	26.54
97	0.7034	61	4.123	24	7.921	124	26.49
98	0.7031	24	4.131	28	7.920	44	26.40
40	0.7030	40	4.132	101	7.899	125	26.33
81	0.7009	103	4.140	27	7.895	126	26.23
101	0.6991	81	4.160	103	7.877	121	26.15

4.3 *Redistribution of the load based on the shortest paths*

With no loss of generality, in this analyzed scenario, interest is on cascade onset and propagation over the bare topological structure of the power transmission system; no reference is made to the specific electrical properties which characterize this electrical infrastructure.

The scenario considered regards the malevolent targeted attack aiming at disconnecting node 88, which handles the largest load in the system, i.e., through which pass the largest number of generator-distributor shortest paths. Once the triggering event occurs, flow redistribution takes place as a mechanism to equilibrate supply and demand constraints. The flow redistribution process is followed by introducing an artificial cascade discrete time step t_i: at t_0 the network is intact, at t_1 the initial induced failure occurs; and at $t_{\geq 2}$ the cascading failure progresses as nodes overload and cause further failures in neighboring elements. The cascading process is followed until the response stabilizes and indicators of the severity of the cascade are computed.

In Fig. 3, the final value of the cascade size, *S*, once the system response has stabilized, is plotted versus the tolerance parameter, α. Cascades of failures have been simulated for nodes capacities in the range $\alpha \in [0, 2]$ (Section 2.3). When $\alpha = 0$ the nodes capacities are equal to the initial loads. When $\alpha = 2$ the nodes capacities amount to three times as much as the initial loads. With this choice of α, various operational conditions are assessed. As expected, increasing the flow-carrying capacity of the network elements reduces the extent of the cascades because flow redistribution can be handled at the local scale. Yet, we observe jumps to larger values of *S*. In order to gain a deeper understanding of this, the transition taking place at $0.43 \leq \alpha \leq 0.44$ is analyzed in detail. This behavior is related to the so-called "islanding" effect. For $\alpha = 0.43$, the failures of 'weak' nodes occurring at the second time step split the network into isolated islanding sections (namely, the northern and the southern parts of the network), disconnecting many generator-distributor paths and thus reducing the demand and stabilizing the power transmission system. Conversely, for $\alpha = 0.44$, nodes {71, 83, 84} along the Adriatic backbone are not failed at the second time step, allowing flow redistribution to weaker nodes, which fail subsequently at the third time step disrupting the power transmission network. This behavior suggests the inclusion of 'weak' nodes in the

system design, for early disconnection or islanding and cascade-controlled operation of complex infrastructures. Finally, the sharp transition occurring at $\alpha = 0.51$ is due to the fact that nodes which are neighbor of high load nodes are able to handle the redistribution of flow thanks to the α increase. This is the case also for the sharp transitions at $\alpha = 0.16$, $\alpha = 1.05$ and $\alpha = 1.33$ in Fig. 3.

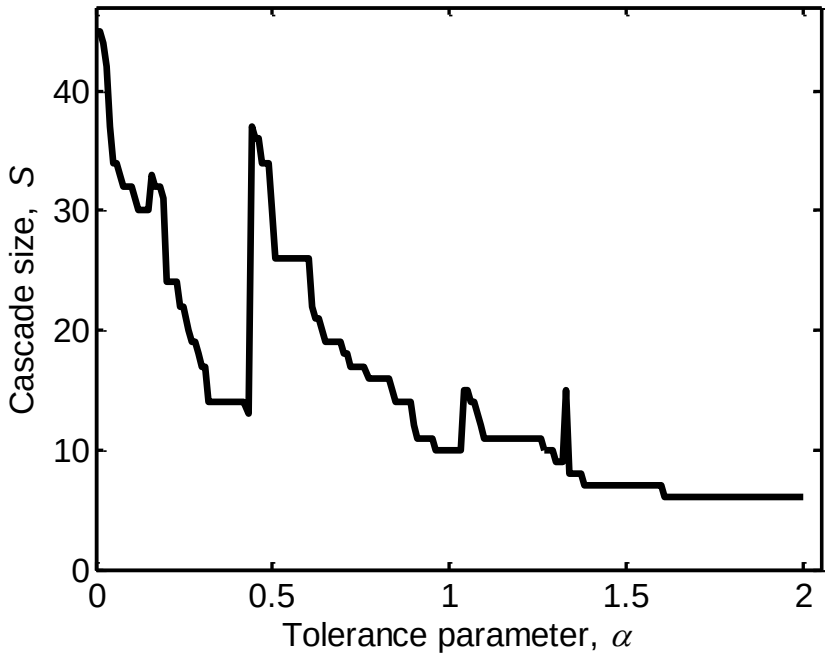

Figure 3. The final value of the cascade size, S, vs. the tolerance parameter, α, when the system response has stabilized. The cascades are triggered by the removal of the most congested node {88}.

Table IV shows the results for the model relative to cascading failures due to the redistribution of the load based on the shortest paths. This model describes a situation completely different from the previous ones since the load at a component is the total number of shortest paths passing through that component. Components 35, 14, 79, 12 and 76 turn out to be the most critical with respect to s_i and d_i whereas t_i gives an opposite ranking for the reason explained before.

The components ranked as the most critical according to f_i are now the ones with a small capacity, since initially they do not have many shortest paths passing through them but still are linked to highly connected components or lie along a direct path linking highly connected components (30 and 35 for critical component 31, LOM9; 7 and 2 for critical components 8, Casanova, and 6, Laboratorio Cesi; 24, 7 and 21 for critical components 17, LOM1, and 15, Musignano): while other components are failing and the highly connected components are still operating, the evolving shortest paths are directed through these critical components which subsequently fail due to their low capacity.

Also, as expected, components of degree one do not participate to any cascade since they involve no shortest path transit. The degree, k_i, of each node i is reported in Table V.

The fifth and seventh columns of Table I report a ranking with respect to the indicators d_i and s_i which is completely different compared to f_i. The most critical components form a path connecting the Northern area to the Tyrrhenian backbone, i.e. Vignole B. − La Spezia – Marginone – Poggio a Caiano. Whenever this path is broken, i.e., either component 14 or 79 or 12 or 76 fails, the connectivity capability is shifted somewhere else in the network, i.e., in the Po river area, leading to an accruement of the cascade with further failures.

Table IV. Summary of the criticality indicators rankings for the model of redistribution of the load based on the shortest paths; only the twenty-four most critical nodes are reported.

Node	f_i	Node	t_i	Node	d_i	Node	s_i
31	0.1622	49	1.000	35	3.800	14	20.19
8	0.1480	73	1.000	14	3.238	79	15.57
17	0.1465	119	1.000	79	3.107	76	13.92
15	0.1417	121	1.000	12	3.000	35	13.10
6	0.1370	122	1.000	76	3.000	12	10.64
23	0.1307	123	1.000	68	2.923	61	10.50
39	0.1276	124	1.000	43	2.917	86	9.928
93	0.1244	125	1.000	36	2.909	11	9.541
22	0.1244	126	1.000	86	2.812	88	9.423
34	0.1165	43	1.083	41	2.775	75	8.880
54	0.1134	107	1.083	61	2.750	68	8.385
120	0.1134	109	1.111	40	2.700	78	7.899
30	0.1134	56	1.136	11	2.676	48	7.852

Node	f_i	Node	t_i	Node	d_i	Node	s_i
94	0.1134	99	1.161	47	2.650	59	7.676
55	0.1118	115	1.213	46	2.628	81	7.521
81	0.1118	111	1.217	66	2.552	110	7.226
86	0.1087	110	1.226	67	2.500	23	7.205
78	0.1087	102	1.233	7	2.455	36	7.091
21	0.1071	11	1.243	48	2.444	7	7.000
62	0.1071	67	1.250	78	2.420	9	6.932
66	0.1055	60	1.258	62	2.397	58	6.803
96	0.1055	103	1.267	81	2.394	67	6.800
58	0.1039	106	1.278	60	2.387	47	6.750
51	0.0992	118	1.280	88	2.385	62	6.706

Table V. Degree of each node displayed in descending order for the network shown in Fig. 1.

Node	k_i	Node	k_i	Node	k_i	Node	k_i	Node	k_i	Node	k_i
68	7	81	4	62	3	13	2	65	2	126	2
64	6	91	4	63	3	15	2	66	2	1	1
24	5	97	4	71	3	16	2	80	2	4	1
35	5	98	4	73	3	17	2	82	2	19	1
43	5	110	4	76	3	18	2	83	2	26	1
79	5	115	4	78	3	22	2	85	2	37	1
88	5	8	3	84	3	23	2	93	2	50	1
92	5	10	3	86	3	25	2	94	2	57	1
101	5	11	3	89	3	29	2	96	2	69	1
103	5	12	3	90	3	31	2	99	2	70	1
2	4	20	3	95	3	32	2	100	2	72	1
3	4	30	3	104	3	33	2	102	2	74	1
7	4	36	3	106	3	34	2	105	2	77	1
14	4	38	3	107	3	39	2	109	2	87	1
21	4	40	3	108	3	42	2	111	2	116	1
27	4	41	3	113	3	44	2	112	2	117	1
28	4	46	3	114	3	45	2	118	2	127	1
47	4	48	3	119	3	49	2	121	2		
59	4	52	3	120	3	51	2	122	2		
60	4	55	3	5	2	53	2	123	2		
67	4	56	3	6	2	54	2	124	2		
75	4	61	3	9	2	58	2	125	2		

4.4 *Intra-comparison*

Considering each individual model, the ranking results of Tables I, III and IV are consistent with a physical analysis of the network system, indeed highlighting the

components which most affect the failure spreading. In the following, the results for the three presented scenarios are compared to identify similarities in the characterization of the relevance of each network element for the different propagation modes.

The logic of propagation of a fixed amount of load and of redistribution of the failure load give consistent results across the criticality indicators: in both cases, the most critical components according to f_i, d_i and s_i are those with highest degree (68, Ravenna Canala and 64, Martignone) and which constitute a bridge between different loosely-connected subsets of components, whose failure effect spreads to regions far apart in the system (59-60, 88, Montalto and 79, Poggio a Caiano). Conversely, it is not always true that most connected nodes are the most critical as it can be seen from node 24 (Milano Centro), which is not among the most critical in the fixed amount of load redistribution model. In the failure propagation model with redistribution, the amount of load transferred to the neighboring components after a failure is typically larger than in the previous case of propagation of a fixed amount of load to all survivor network nodes; this explains the differences in the ranking among critical components with respect to the previous model and the t_i ranking; note that for a small load transfer, as in the first failure propagation model, the poorly connected nodes only enter the cascade soon after its initiation if either they are themselves triggering it or they reside in the neighborhood of a triggering node, resulting in their small t_i.

In the case of the redistribution of the load based on the shortest paths, the components ranked as the most critical according to f_i are those of small capacity, since initially they do not have many shortest paths passing through them but still are linked to highly connected components.

5 CENTRALITY MEASURES FOR CASCADES

Regarding the role that an element plays in a network, various measures of the importance of a network node, i.e. of the relevance of its location in the network with respect to a given network performance, have been introduced. In social networks, for example, the so-called centrality measures are introduced as importance measures to qualify the role played by an element in the complex interaction and communication occurring in the network. Classical topological centrality measures are the degree centrality[10,11], the closeness centrality[9,11], the betweenness centrality[11] and the information centrality[12,13,23]. They specifically rely only on topological information to qualify the importance of a network element.

The topological degree centrality, C^D, gives the highest score of importance to the node with the largest number of first neighbors. This agrees with the intuitive way of estimating the influence of a node in a graph from the size of its immediate environment. Quantitatively, the topological degree centrality is defined as the degree of a node, normalized over the maximum number of neighbors this node could have: thus, in a network of N nodes, the topological degree centrality of node i, C_i^D, is defined as[10,11]:

$$C_i^D = \frac{k_i}{N-1} = \frac{\sum_{j \in G} a_{ij}}{N-1} \quad 0 \le C_i^D \le 1 \tag{6}$$

where k_i is the degree of node i and $N-1$ is the normalization factor introduced to account for the fact that a given node i can at most be adjacent to $N-1$ other nodes. The running time required for computing C^D for all nodes is $O(N)$.

The topological closeness centrality, C^C, captures the idea of speed of communication between nodes in a way that the node which is "closest" to all others receives the highest score. In other words, this measure allows identifying the nodes which on average need fewer steps to communicate with the other nodes, not only with the first neighbors. Because this

measure is defined as “closeness”, quantitatively the inverse of the node's mean distance from all the others is used. If d_{ij} is the topological shortest path length between nodes i and j, i.e., the minimum number of edges traversed to get from i to j, the topological closeness centrality of node i is[9,11]:

$$C_i^C = \frac{N-1}{\sum_{j \in G} d_{ij}} \quad 0 \le C_i^C \le 1 \tag{7}$$

Note that also this measure is normalized to assume values in the interval $[0,1]$.The running time required for computing C^C for all nodes by means of Floyd algorithm is[24] $O(N^3)$.

The topological betweenness centrality, C^B, is based on the idea that a node is central if it lies between many other nodes, in the sense that is traversed by many of the shortest paths connecting pairs of nodes. The topological betweenness centrality of a given node i is quantitatively defined as[11]:

$$C_i^B = \frac{1}{(N-1)(N-2)} \sum_{j,k \in G, j \neq k \neq i} \frac{n_{jk}(i)}{n_{jk}} \quad 0 \le C_i^B \le 1 \tag{8}$$

where n_{jk} is the number of topological shortest paths between nodes j and k, and $n_{jk}(i)$ is the number of topological shortest paths between nodes j and k which contain node i. Similarly to the other topological centrality measures, C_i^B assumes values between 0 and 1 and reaches its maximum when node i falls on all geodesics (paths of minimal length between two nodes). The running time required for computing C^B for all nodes by means of the Floyd algorithm is $O(N^3)$.

The topological information centrality, C^I, relates a node importance to the ability of the network to respond to the deactivation of the node. In this view, the network performance is measured by the network topological efficiency $E[G]$ defined as[23]:

$$E[G] = \frac{1}{N(N-1)} \sum_{i,j \in G, i \neq j} \varepsilon_{ij} \tag{9}$$

where $\varepsilon_{ij} = 1/d_{ij}$ is the efficiency of the connection between nodes i and j, measured as the inverse of the shortest path distance linking them.

The topological information centrality of node i is defined as the relative drop in the network topological efficiency caused by the removal of the edges incident in i[23]:

$$C_i^I = \frac{\Delta E(i)}{E} = \frac{E[G] - E[G'(i)]}{E[G]} \quad 0 \leq C_i^I \leq 1 \tag{10}$$

where $G'(i)$ is the graph with N nodes and $K - k_i$ edges obtained by removing from the original graph G the edges incident in node i. An advantage of using the efficiency to measure the performance of a graph is that $E[G]$ is finite even for disconnected graphs. Also C^I is normalized by definition in the interval [0, 1].

The running time required for computing C^I for all nodes by means of the Floyd algorithm is $O(N^4)$ [25].

Table VI reports the ranking of the individual network components according to the information (C^I), degree (C^D), closeness (C^C) and betweennes (C^B) centrality measures.

Table VI. Information, degree, closeness and betweennes centrality measure ranking for the network of Fig. 1; only the twenty-four most central nodes are reported.

Rank	Node	C^I	Node	C^D	Node	C^C	Node	C^B
1	68	0.6901	68	0.0556	64	0.1691	88	0.2941
2	14	0.6900	64	0.0476	75	0.1649	14	0.2775
3	88	0.6897	24, 35, 43, 79, 88, 92, 101, 103	0.0397	79	0.1645	75	0.2721
4	119	0.6892	2, 3, 7, 14, 21, 27,	0.0317	81	0.1643	64	0.2523

Rank	Node	C^I	Node	C^D	Node	C^C	Node	C^B
			28, 47, 59, 60, 67, 75, 81, 91					
5	64	0.6876			14	0.1617	79	0.2333
6	75	0.6867			78	0.1615	101	0.2105
7	122	0.6856			67	0.1603	76	0.2093
8	79	0.6852			62	0.1601	59	0.1840
9	12	0.6841			61	0.1595	12	0.1817
10	78, 110	0.6840			63	0.1583	110	0.1781
11					76	0.1579	61	0.1732
12	101	0.6838			88	0.1573	102	0.1721
13	59	0.6836			41	0.1537	98	0.1667
14	123	0.6827			71	0.1535	68	0.1584
15	76	0.6826			60	0.1533	71	0.1520
16	47	0.6822			65	0.1520	83	0.1476
17	43	0.6819			59	0.1511	84	0.1452
18	24	0.6814			68	0.1509	40	0.1437
19	35	0.6811			73	0.1484	67	0.1338
20	61	0.6808			82	0.1482	35	0.1326
21	121	0.6802			86	0.1475	78	0.1271
22	71, 81, 98	0.6800			83	0.147	60	0.1232
23					80	0.1469	81	0.1225
24					40	0.1467	107	0.1113

The comparison of the results in Table VI with those obtained with the criticality indicators applied to the models of local propagation of a fixed amount of load (Table I) and of redistribution of the failure load (Table III), shows that the degree centrality measure perfectly matches the criticalities found by the four introduced indicators, i.e., components 68, 64, 35 and 24. Thus, when the cascading propagation process can be modeled in either of these two ways, the attention must be focused on the components with highest degree of connectivity. This is because according to these failure propagation mechanisms, the failure spreading is affected more by the individual component connectivity than by the global network connectivity accounted for by the other centrality measures. It must be noticed that the other centrality measures not only partially reflect the criticalities found by C^D but also complement them, i.e., components 75, 79 and 88.

In case of stronger coupling among components as in the model of redistribution of the failure load, also the betweenness centrality measure can serve the purpose of indicating components criticality with respect to propagating failures, e.g. components 88, 75, 79 and 101.

From the comparison between the results in Table VI and those for the model relative to cascading failures propagated by the redistribution of the shortest paths (Table IV), it can be said that the betweenness and information centrality measures only partially account for the criticalities highlighted by the indicators d_i and s_i (node 14), since not only the centrality of a component is relevant but also the fact that it is connecting central components (as do critical nodes 12 and 76 connected to central component 14 and critical node 79 connected to central component 75 in the network).

With respect to the indicators f_i, it can be said that the most critical components are those less connected, which lie along a direct path linking components with the highest degree centrality (35 and 28 for critical component 31; 3 and 7 for critical component 8; 24 and 7 for critical components 17 and 15; 2 and 7 for critical component 6).

6 CONCLUSIONS

In this paper, the feasibility of evaluating component criticality indicators for a realistic-size network has been investigated. Three different models of cascading failures have been considered differing for the logic of redistribution of the failure load. Cascade events triggered by both random failures and targeted attacks have been analyzed.

When applied to models of local propagation of a fixed amount of load and of redistribution of the failure load, three of the proposed criticality indicators, namely f_i, d_i and s_i, are consistent in their criticality ranking of the components. When applied to the model of

cascading failures due to the redistribution of the shortest paths, s_i and d_i are consistent in ranking the components according to their criticality.

In general, the frequency of participation of every component to a cascade, f_i, appears to be the most relevant indicator since it highlights the direct contributions of each component in the cascading failure process irrespective of the different propagation logic. In the considered realistic-size network, for the first two model considered, f_i identifies as most critical those nodes with highest degree, while for third model, it ranks as most critical those nodes which have few connections but which are linked to highly connected components or lie along a direct path linking highly connected components. Moreover, f_i and d_i are consistent in their criticality ranking of the components in cascades triggered by several different initial conditions.

Complementary criticality information is provided by the d_i and s_i indicators, which capture the components failure contribution in promoting successive failures. In the present reference case study, for the model of local propagation of a fixed amount of load, s_i particularly highlights those nodes bridging different loosely-connected subsets of components, while for the model of redistribution of the shortest paths, s_i identifies the criticality of those nodes connected to nodes having high centrality values. Conversely, the ranking provided by the t_i indicator is dependent on the coupling strength among components: when the components are weakly coupled, it gives homogeneous results with the other indicators, whereas it gives opposite results if the components are strongly coupled. In this respect, the t_i indicator could be useful in identifying the degree of coupling among components in interconnected systems with respect to propagating failures.

The rankings obtained with the different cascade criticality indicators have been compared with classical centrality measures. The degree and betweenness centralities, which account for the number of connections pointing to a component and for the number of shortest paths

passing through a component, respectively, appear to play a major role in identifying those network components which most contribute to the failure propagation process.

For the models of local propagation of a fixed amount of load and of redistribution of the failure load, the degree centrality measure appears to be the most characteristic for the cascade process, with the betweenness centrality measure providing complementing information in case of intense coupling among components.

For the model of cascading failures propagated by the redistribution of the shortest paths, the betweenness centrality measure only partially highlights those components which most contribute in determining large-sized failure cascades. Further investigations would be worth to identify or devise a general centrality measure characteristic of the model of cascading failures propagated by the redistribution of the shortest paths.

ACKNOWLEDGMENTS

This work has been partially funded by the Foundation pour une Culture de Securite Industrielle of Toulouse, France, under the research contract AO2006-01. The authors wish to thank the anonymous referees for the numerous constructive comments and suggestions which have led to a significantly improved work.

REFERENCES

1. Dobson I, Carreras BA, Lynch VE, Newman DE. Complex systems analysis of series of blackouts: Cascading failure, critical points, and self-organization. Chaos: An Interdisciplinary Journal of Nonlinear Science, 2007; 17(2):026103.

2. Cohen R, Erez K, ben-Avraham D, Havlin S. Resilience of the Internet to random breakdowns. Physical Review Letters, 2000; 85(21):4626-4628.

3. Zheng J-F, Gao Z-Y, Zhao X-M. Clustering and congestion effects on cascading failures of scale-free networks. EPL, 2007; 79(5):58002.

4. Pastor-Satorras R, Vespignani A. Epidemic Spreading in Scale-Free Networks. Physical Review Letters, 2001; 86(14):3200.

5. Smart AG, Amaral LAN, Ottino JM. Cascading failure and robustness in metabolic networks. Proceedings of the National Academy of Sciences, 2008; 105(36):13223-13228.

6. Dobson I, Carreras BA, Newman DE. A loading-dependent model of probabilistic cascading failure. Probability in the Engineering and Informational Sciences, 2005; 19(1):15-32.

7. Zio E, Sansavini G. Modelling failure cascade in network systems due to distributed random disturbances. In: Martorell S, Soares CG, Barnett J, editors. Safety, Reliability and Risk Analysis: Theory, Methods and Applications. Proceedings of the joint ESREL 2008 (European Safety and Reliability) and 17th SRA-Europe (Society for Risk Analysis Europe) Conference; 2008 Sep 22–29; Valencia, Spain. London (UK): CRC Press, Taylor & Francis Group; 2009. p. 1861-6.

8. Motter AE, Lai Y-C. Cascade-based attacks on complex networks. Physical Review E, 2002; 66(6):065102.

9. Sabidussi G. The centrality index of graphs. Psychometrika, 1966; 31(4):581-603.

10. Nieminen J. On the centrality in a graph. Scandinavian Journal of Psychology, 1974; 15(1):332-336.

11. Freeman LC. Centrality in social networks conceptual clarification. Social Networks, 1978; 1(3):215-239.

12. Freeman LC, Borgatti SP, White DR. Centrality in valued graphs: A measure of betweenness based on network flow. Social Networks, 1991; 13(2):141-154.

13. Little RG. Controlling Cascading Failure: Understanding the Vulnerabilities of Interconnected Infrastructures. Journal of Urban Technology, 2002; 9(1):109-123.

14. Dobson I, Carreras BA, Newman DE. A probabilistic loading-dependent model of cascade failure and possible implications for blackouts. Proc. 36th Annual Hawaii International Conference on System Sciences, 2003.

15. Newman DE, Nkei B, Carreras BA, Dobson I, Lynch VE, Gradney P. Risk assessment in complex interacting infrastructure systems. Proc. 38th Annual Hawaii International Conference on System Sciences, January 3-6, 2005. Computer Society Press.

16. Maiese K, Wagner J, Boccone L. Nitric oxide: a downstream mediator of calcium toxicity in the ischemic cascade. Neuroscience Letters, 1994; 166(1):43-47.

17. Newman MEJ, Girvan M. Finding and evaluating community structure in networks. Physical Review E, 2004; 69(2):026113.

18. Batagelj V. Semirings for social networks analysis. Journal of Mathematical Sociology, 1994; 19(1):53-68.

19. Dueñas-Osorio L., Vemuru SM. Cascading failures in complex infrastructure systems. Structural Safety, 2009; 31:157–167.

20. Albert R, Jeong H, Barabási A-L. Error and attack tolerance of complex networks. Nature, 2000; 406:378-382.

21. Terna S.p.A. - Rete Elettrica Nazionale. Dati statistici sull'energia elettrica in Italia. [Internet]. 2002. Available from: http://www.terna.it/LinkClick.aspx?fileticket=PUvAU57MlBY%3d&tabid=418&mid=2501 Italian.

22. Rosato V, Bologna S, Tiriticco F. Topological properties of high-voltage electrical transmission networks. Electric Power Systems Research, 2007; 77:99-105.

23. Latora V, Marchiori M. A Measure of Centrality Based on the Network Efficiency, New Journal of Physics, 2007; 9:188.

24. Floyd RW. Algorithm 97: shortest path. Communications of the ACM, 1962; 5(6):345.

25. Fortunato S, Latora V, Marchiori M. Method to find community structures based on information centrality. Physical Review E, 2004; 70:056104.