Graph Centralities and Cascading Failures in Networks

Francesc Comellas and *Llorenç Sánchez*

Department of Mathematics, Universitat Politècnica de Catalunya,
Esteve Terrades 5, 08750 Castelldefels (Barcelona) Catalonia, Spain

Extended Abstract

Cascading failures. A cascading failure in a complex system is a process in which an initial failure in one or several of its elements leads to a sequence of failures which spread to other elements and in some cases collapses the whole system.

In our interconnected world, these processes are increasingly becoming common and affect our everyday lives. This is the case of the recent power-grid failure of 16 June 2019 in Argentina and neighbor countries and the recurrent failures of Twitter, Facebook, Instagram and other social networks, all of them carrying important disruptions in relevant services.

What causes these cascade processes is not well understood, and thus they are really difficult to predict or contain. In many cases cascading failures arise from changes in flows, as happens in transportation networks, including electric grids or logistic supply chains, but there are not many models or a general graph theoretical framework to help in their analysis.

Background. The publication of a paper by Wang and Rong (WR) in 2009 [12] on cascading failures in the power grid of western US and Canada had a big impact. Their study was based on public data available from the most-cited paper on small-world networks by Watts and Strogatz [13] and received a wide interest after its review in the New York Times, with some curious anecdotes, like that a military analyst, L. M. Wortzel, asked the House of Foreign Affairs Committee to investigate the authors of the article for possible terrorism.

The WR paper had an important repercussion leading to tens of publications on this topic, including the study of other networked systems, interdependent networks, vulnerable sets, etc. (see, for example, [2, 11, 14, 10] and references therein.) This research is based mainly on computer simulations and in many cases lacks a founding mathematical background.

Since 1998 [13] we know that many graphs which are associated to complex systems belong, mainly, to a category that is now known as "small-world scale-free". Their characteristics are a large local concentration of vertices or clustering (the vertices have many common neighbors) and, at the same time, a small average distance and diameter. Other properties present in some real complex systems are that their degrees follow a power law [1] and modularity (sometimes related to the existence of communities or clusters) [9, 4]. Thus, it make sense to consider different graph families when modeling a cascading failure process.

In [12], WR studied the vulnerability of the western US and Canada electric grid, modeled as a graph, from the failure of some sets of vertices. Their model for cascading failures, following the previous work of Motter and Lai[6], is very simple as each vertex has an associated load which is computed from its degree and the degree of its neighbors. Vertices have also a maximum capacity. If a vertex fails, its load is distributed to all its neighbors and if their new load becomes larger than this capacity, they also fail. The most important result is that the collapse of a network is triggered more easily from the failure of a few nodes of low degree than the failure of nodes of high degree. From their model we see that WR actually selects the set of initial failing vertices from values of their degree centrality. Clearly, there is need to go beyond this simple model and consider other centralities which could reflect better a dynamic flow process in a real network.

Our approach. In this research we consider different families of model graphs and also data available for real systems (US and European power grids and airport networks, human interactome etc.), and we compare the classical WR model with new cascading failure models based on other relevant graph centralities. In this way we can study which topological and dynamic parameters of a network are most likely to affect the propagation (cascading) of failures.

The new centralities are based on graph parameters and invariants like usual distance measures, the *betweenness* of a vertex, which reflects how many short paths between all pairs of possible vertices go through a given vertex, *communicability* [3] which measures the set of paths that begin and end in a vertex, and others based on the eigenvectors of the graph Laplacian and adjacency matrices, like PageRank.

All these centralities are a significant improvement with respect to the usual degree centrality from WR, which only accounts for direct links to a vertex. As an example, the PageRank of a vertex depends on the PageRank value of all vertices in the graph; vertices that are several edges apart contribute less to the calculation. Thus, a vertex that is linked (directly or through others) to many vertices which have a high PageRank value receives a high rank itself and it is said to be more central.

With this study, we want to know which sets of nodes can produce a maximum cascading failure for each different graph family and centrality.

A cascading failure general model. Each node starts with an initial load L_j which depends on the centrality that we are studying, and an initial maximum capacity, C_j. This load and capacity are defined as $L_j = [c_j(\sum_m c_m]^\alpha$ and $C_j = T \cdot L_j$, where c_j is the centrality of vertex j, m runs on the set of adjacent vertices, α is a tunable parameter which controls the strength of the initial load and T is the tolerance parameter ($T \geq 1$). Once these parameters are fixed, the next steps are:

- Sort all verticess by their load L_j and select (initial failing sets) N_{IFS} vertices with the highest load, N_{IFS} with the lowest load and N_{IFS} with a load value around the median.
- For a given load group consider, independently, the failure of each vertex and distribute its load to active neighbours. If i is the failing vertex and j an active adjacent vertex then: $L_j^{new} = L_j + \Delta L_{ji}$ where $\Delta L_{ji} = L_j \dfrac{L_j}{\sum_{m\in\Gamma_i} L_m}$ is the contribution of extra load received from the set of failing neighbours Γ_i. Therefore an adjacent vertex with a higher load will receive a higher shared load from the the failing vertex.
- When the load has been distributed to its neighbouring vertices, check if any of the vertices affected exceeds its maximum capacity C_j. If this is the case, these vertices will fail and their load redistributed as explained. This procedure is repeated until the remaining active vertices stabilize or all vertices fail.
- When the cascading process is over, we count the vertices that have failed due to the failure of the initial vertex i, CF_i, and repeat the process with the remaining vertices in the same load group.
- To measure the robustness of the whole graph, we calculate CF_{IFS} as $CF_{\text{IFS}} = \frac{\sum_{i\in\text{IFS}} CF_i}{N_{\text{IFS}}(N-1)}$ where IFS and N_{IFS} represent the initial failing set of vertices and its cardinality, respectively.

Note that the original cascading failure model of Wang and Rong, corresponds to consider for c_i the degree centralities.

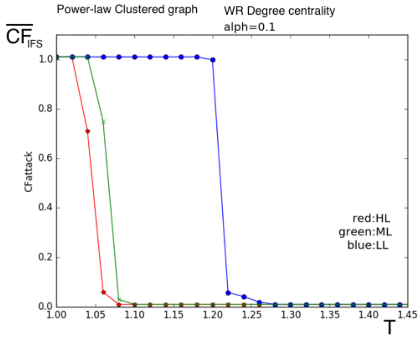
Methodology. We have considered ten topologically different graph families: Geographical Threshold 2D, Geographical Threshold 3D, Watts-Strogatz (high clustering), Watts-Strogatz (low clustering), Barabasi-Albert, Erdős - Rényi, Power-law Clustered, Random Geometric, Random Partition, Random Regular. See the NetworkX documentation [7] for details on each family and associated bibliography.

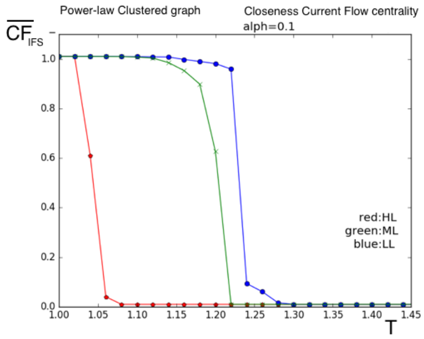
The centralities used in our study are: Degree, Betweenness, Communicability, Closeness Current Flow, PageRank and Eigenvector [7, 5].

We have coded all cascading failure methods in Python 2.7 while using the NetworkX package 1.10 [7] to generate all graphs and call its preprogrammed centrality functions.

For each combination of a graph family, centrality and α value (0.1 and 0.5), we run a set of 20 simulations for 20 different graphs with the same order $|V| = 100$ and size $|E| \approx 400$ from this graph family (these graphs are saved and used for each other centrality). Each run involves increasing the load tolerance parameter from 1 to 1.5 in steps of 0.025. The results are averaged for these 20 simulations.

As an example of the results obtained, the following two figures compare, for Power-law Clustered graphs, cascading failure processes based on the WR model of degree centrality (left) with the closeness current flow centrality model (right), both with parameter $\alpha = 0.1$.





Both figures represent $\overline{CF}_{\text{IFS}}$ versus T. We note a clear different behavior with respect the two centralities when the initial failing vertices have a load around the median.

Results. Our results show, as expected, a different behavior among graph families and centralities. However in many cases cascading failure sizes depend on the load of the initial vertices in a similar way to the results obtained by Wang and Rong, which we have reproduced validating our methods. But, when checking, for a given graph family, the effects of selecting the initial failing nodes according to different centralities show a pattern: With cascading failure models based on centralities like Betweenness and Communicability, all three load options for the initial failing set lead to no cascading failures for small values of the tolerance parameter. Models based on the Closeness Current Flow and Pagerank centralities, on the other hand, show graph global failures at higher values of the tolerance. This reflects the fact that, for the former centralities, a high load for a vertex means that it is well connected through shortest paths to nearby vertices (and all other vertices) and this facilitates a redistribution of loads. For the latter centralities a high load means a more diffuse connection to all vertices.

We have also found for some models a clearly different behavior with respect to the WR results for initial failing sets of median loads, see the figure above. All these results suggest that, to increase a graph resilience, and thus to protect a real life network associated to it, there is need to decide which centrality can model better flows in the network and check all vertices according to it to select those vulnerable following all former criteria and not just by considering their degrees as it happens in most of the current models for cascading failures.

Acknowledgement. This research was supported by the Ministerio de Ciencia e Innovación/Agencia Estatal de Investigación, Spain, and the European Regional Development Fund under project PGC2018-095471-B-I00.

References

[1] A.-L. Barabasi , R. Albert. Emergence of scaling in random networks. In **Science** 286, (1999) 509-512. `doi:10.1126/science.286.5439.509`.

[2] S.V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley, S. Havlin. Catastrophic cascade of failures in interdependent networks. In **Nature** 464 (2010) 1025-1028. `doi:10.1038/nature08932`.

[3] E. Estrada. The communicability distance in graphs. In **Linear Algebra Appl.** 436 (2012) 4317-4328. `doi: 10.1016/j.laa.2012.01.017`.

[4] G. Lin, Z. Di, Y, Fan. Cascading failures in complex networks with community structure. In **nt. J. Mod. Phys.** 25, (2014), 1440005. `doi: 10.1142/S0129183114400051`.

[5] E. Estrada, P. A. Knight A First Course in Network Theory. Oxford University Press, Oxford, UK, 2015

[6] A.E. Motter, Y.C. Lai. Cascade-based attacks on complex networks. In **Phys. Rev. E.** 66 (2002) 065102. doi:10.1103/PhysRevE.66.065102.

[7] NetworkX documentation. https://networkx.github.io/documentation/networkx-1.10/. Accessed 29 June 2019.

[8] L. Sánchez Marin. Cascading Failures in Complex Networks. TFG EETAC-UPC. June 13, 2018.

[9] R.V. Solé, S. Valverde. Information theory of complex networks: on evolution and architectural constraints. In **Lecture Notes in Phys.** 650 (2004), 189 – 207. doi:10.1007/978-3-540-44485-5_9+.

[10] R. Spiewak, S. Soltan, Y. Forman, S.V. Buldyrev, G. Zussman. A study of cascading failures in real and synthetic power grid topologies. In **Network Science** 1-21. Pub. online: 23-Aug-2018. doi:10.1017/nws.2018.14.

[11] V. Turau, C. Weyer. Cascading failures caused by node overloading in complex networks. 2016 Joint Workshop on Cyber- Physical Security and Resilience in Smart Grids (CPSR-SG) doi:10.1109/CPSRSG.2016.7684104.

[12] J.W. Wang, L.L. Rong. Cascade-based attack vulnerability on the US power grid. In **Safety Science** 47 (2009) 1332-1336. doi:10.1016/j.ssci.2009.02.002.

[13] D.J. Watts, S. H. Strogatz. Collective dynamics of small-world networks. In **Nature** 93 (1998) 440-442. doi:10.1038/30918.

[14] Y. Yang, T. Nishihawa, A.E. Motter. Small vulnerable sets determine large network cascades in power grids. In **Science** 358 (2017) eaan3184. 10.1126/science.aan3184.

Bordeaux Graph Workshop 2019. Bordeaux, October 28-31, 2019.
Bordeaux University, Talence, France.
BGW 2019 booklet, pag. 118-121 https://bgw.labri.fr/2019/booklet.pdf