Evaluation of cascade effects for transit networks

Antonio Candelieri, Ilaria Giordani, Bruno G. Galuzzi, Francesco Archetti

Abstract This paper presents a network analysis approach to simulate cascading effects in a transit network with the aim to assess its resilience and efficiency. The key element of a cascade is time: as time passes by, more locations or connections of the transit network which are nodes and edges of the associated graph can be affected consecutively as well as change their own condition. Thus, modifications in terms of efficiency and resilience are also dynamically evaluated and analysed along the cascade. Results on the two case studies of the RESOLUTE project (i.e., Florence, in Italy, and the Attika region, in Greece) are presented. Since the two case studies are significantly different, important differences are reflected also on the impacts of the relative cascades, even if they were started in both the two cases from the node with the highest betweenness centrality.

Antonio Candelieri.
Department of Computer Science, Systems and Communications, University of Milano-Bicocca, viale Sarca 336, 20125, Milan, Italy.
Consorzio Milano-Ricerche, via Roberto Cozzi, 53, 20126, Milan, Italy
e-mail: antonio.candelieri@unimib.it

Ilaria Giordani
Department of Computer Science, Systems and Communications, University of Milano-Bicocca, viale Sarca 336, 20125, Milan, Italy.
Consorzio Milano-Ricerche, via Roberto Cozzi, 53, 20126, Milan, Italy
e-mail: ilaria.giordani@disco.unimib.it

Bruno G. Galuzzi
Department of Computer Science, Systems and Communications, University of Milano-Bicocca, viale Sarca 336, 20125, Milan, Italy.
e-mail: brunogaluzzi@gmail.com

Francesco Archetti
Department of Computer Science, Systems and Communications, University of Milano-Bicocca, viale Sarca 336, 20125, Milan, Italy.
Consorzio Milano-Ricerche, via Roberto Cozzi, 53, 20126, Milan, Italy
e-mail: francesco.archetti@unimib.it

Keywords. Network Analysis. Resilience. Cascading effects. Urban Transport System.

1 Introduction

In complex Urban Transport System (UTS), such as urban city or regional transports, the nodes individually experience a load, such as the volume of passengers passing through that node, and in normal circumstances, this load does not exceed the capacity of that node [1]. Cascades failures are initiated when a heavily loaded node is lost for some reason (an accident, an infrastructure collapses or an attack) and a load of that node (i.e. the flow passing through it) must be redistributed to other nodes of the UTS. This redistribution may cause other nodes to exceed their capacity causing them also to fail. Then, the number of failed nodes increase, propagating throughout the network.

The analysis of **cascades** in complex networks [2, 3, 4] is an important topic towards modelling and assessment of network resilience. In this case, the term **resilience** indicates the capability of a UTS to resist to a node failure, and the possible resulting cascade. Resilience has come to define a set of properties of a much broader socio-technical framework to cope with infrastructure threats and disruptions including preparedness, response, recovery and adaptation. All these concepts run across several application domains like ecology, economics and networked infrastructures.

Methods of representation and analysis of resilience come from several different communities like statistical physics, graph theory, optimization, network science and engineering design. An interesting tool of quantitative analysis of recovery capability of a system is reported in [5], which considers the London underground where a new resilience measure is proposed according to the speed with which the passenger count time series return to normal condition. This information is taking as an indicator of how quickly the underground transport system is able to recover from the shock and, after that, resume normal operations.

In the extensive literature about cascades and resilience, a primary distinction must be made between percolation cascades and capacity cascades [6]. In the former, e.g. epidemiological networks, the nodes change their status due to interaction with their neighbourhood. In the latter, e.g. water distribution networks [7, 8] and transit networks, cascades occur when, due to failure in edges/nodes, the flow can no longer be carried out by the edges with their capacities or when some of the nodes fail. The failure in a capacity cascade can jump to nodes that are many hops away from the initial failure also skipping the neighbourhoods.

This paper is focused on the capacity cascades in UTSs. Recent works about it can be found, for example, in [9] and [10]. The main goal is to characterize the level of efficiency and resilience that ensure the persistence of key functions even in the presence of cascading failures. This is coherent with the activities of the EU project RESOLUTE (www.RESOLUTE-eu.org), whose general aim is the operationalization of the resilience guidelines and development of software tools for resilience

assessment and support to a quick recovery of the service.
Two main lines of analysis have been pursued and are represented in the literature. When demand and supply data are available, the interaction between demand and supply is simulated through mathematical modelling and/or software simulation and generates the flow within the network. Simulation allows to capture the operational and economic aspects, such as the flow-induced costs, and the behaviour of users both prior and post any disruptions, as presented in [11].
When, as in most cases, data are not available, one is left to work with the network topology; traffic flows are not explicitly modelled, and the number of shortest paths between any two points passing through a node/link is taken as a proxy of the traffic demand in that node/link. This approach is adopted, for instance, in [12] and [13]. The topological structure of a network provides critical information and enables the computation of efficiency/resilience measure, which has been applied to study the Boston subway network and the transit networks of major cities worldwide [14]. Typically, only information about the interconnections is needed to create the graph associated to the transit infrastructure and still they can provide fundamental insights about the structural weakness of a transport network. Therefore, to operationalize resilience management in a wide set of conditions, the topological approach has been adopted in this paper, based on the description of a transit network and a possible capacity cascade using the graph theory.

2 Graph-based modelling of a transit network

The main elements of a networked infrastructure, such as a UTS, can be easily mapped into elements of a directed graph $G = (V, E)$, where V is a set of n nodes and E is a set of edges, which are ordered 2-element subsets (i, j), with i and j elements of the V set. In case of a UTS, the nodes represent the locations of interest on the transportation network, such as towns, bus/rail stops, road intersections, etc. whereas the edges represent connections/links between locations, such as roads, rail lines, bus line sections, etc. Furthermore, another relevant concept is the route, that can be mapped into a series of connected edges of the graph, and with a specific label, to distinguish by other routes.
Considering connectivity information at node levels, a measure of the network organization is the **betweenness centrality** [15], computed for each node i as:

$$g(i) = \sum_{j,k \in V, j \neq i, k \neq i} \frac{\sigma_{jk}(i)}{\sigma_{jk}}, \quad (1)$$

where σ_{jk} is the total number of shortest paths from the node j to node k and $\sigma_{jk}(i)$ is the number of those paths passing through node i. Two different parameters are considered to measure the efficiency and the resilience of a network, during and after a cascade. The first is the relative size of the **largest connected component** (S) [12, 13], defined as

$$S = \frac{N^{'}}{N}, \tag{2}$$

where $N^{'}$ and N are, respectively, the number of nodes in the **largest connected component** after and before the cascade. The second is the **network efficiency** (E) [16], defined as

$$E = \frac{1}{n(n-1)} \sum_{i,j \in V, i \neq j} \frac{1}{d_{ij}}, \tag{3}$$

where $d_{ij} = d(i,j)$ represent the **length of shortest path** between nodes i and j, named **distance**. Normalization by $n(n-1)$ ensures that $E \leq 1$, where 1 is obtained for a complete graph. These two quantities can be computed before, during and after the cascade event.

3 Capacity cascades caused from a station closure

Disruptions of a transportation network can be of different types (accidents, infrastructure collapses, attacks, etc.) and can lead to impacts with different severities: injuries, fatalities. Common disruptions, such as a road link blocked, a rail service interruption, a strike, etc., have an impact with low severity. In this work, we simulate a capacity cascade caused by the **closure of a station or stop**, that means the access to that station/stop is disabled but transport lines/routes passing through it are not interrupted. In this case, the node k of a graph, corresponding to the station of the UTS, is removed from the graph but with the possibility that all the paths passing through it are still maintained. The resulting graph after this event is $G^{'} = (V^{'}, E^{'})$ where: $V^{'} = V - \{k\}$ and

$$E^{'} = E - \{(i,k) \in E, (k,j) \in E\} \cup \{(i,j) : \exists (i,k) \in E \wedge \exists (k,j) \in E\} \tag{4}$$

To select the station to close, the g value is computed for every node of the network and the node with the highest g value is selected as target node. The intentional capacity cascade in the network is simulated starting from the removal of the target node. As in [1], we definite $L_i(t)$ the total number of shortest paths, or load, passing through node i at a certain iteration t of the cascade, whereas ψ_i is the maximum load that the node i can handle, or capacity. This value corresponds to the initial load at iteration $t = 0$, multiplied by a tolerance parameter $\alpha \geq 1$:

$$\psi_i = \alpha L_i(0) \;\; i = 1, 2, ..., n \tag{5}$$

To start the cascading effect, the node with higher (initial) load is removed, and a new graph is re-computed using Eq. 4. Then, the g value for each of the remained nodes is re-computed: if any node has a g value exceeding its own capacity, then that node is removed from the network. The process iterates until no more node must be removed from the network, that is the termination of the cascade. Note that the load for each node is updated along the iterations of the cascade simulation, whereas the capacities are set at the beginning. Logically, capacity could be increased with the aim to contrast cascades of failures in the network.

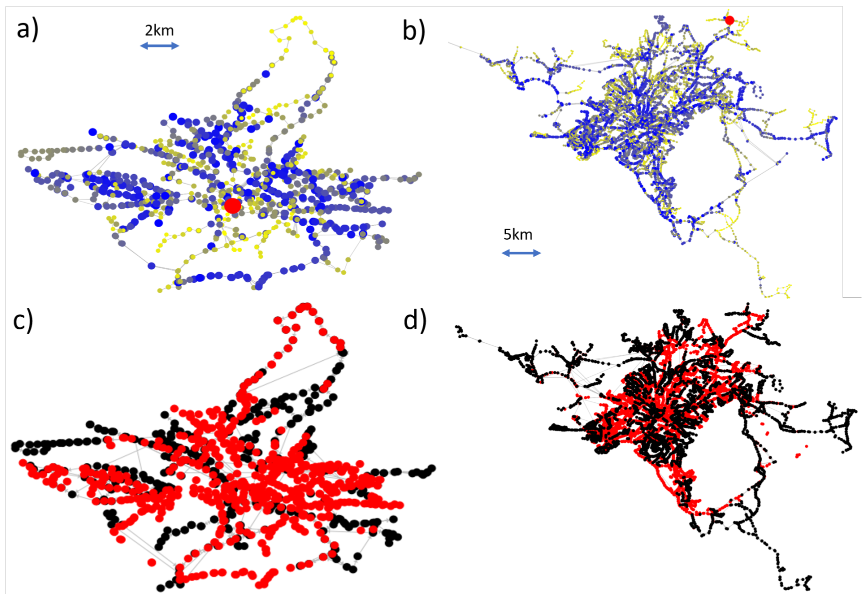

Fig. 1 a) Graph of the bus transport network in Florence before the cascade. b) Graph of the transport network in Attika region before the cascade. c) Graph of the bus transport network in Florence after the cascade. d) Graph of the transport network in Attika after the cascade effect. In a) and b) the red points represent the nodes from which the cascades begin (the ones with the highest g-value), whereas in c) and d) the red points represent the node failed along the cascade.

4 Experimental setting and results

In this section, we report the results obtained from network analysis applied to two real UTSs. These systems are modelled through a directed multi-graph, because more than one route/line may connect two stations; moreover the graph is direct to model direction of each line from one stop to the next.
The first one consists of the public bus transportation in Florence. The number of bus stops is 999, whereas the number of directed edge is 3226. Fig. 1, a) shows the associated graph of the network. To improve the visualization, we did not draw multiple edges. The different colors and sizes of the nodes indicate the different values of their $g-$ value. Passing from the yellow color to the blue one, and from the lower size to the greater size of the nodes, we have an increase of the g-value. The bigger red point indicates the node with the highest g-value (35517.7).

The second network consists of the public transportation network (bus, tramway,...), in the Attika region. The number of stops is 7681, whereas the number of directed edges is 18.128. Fig. 1, b) shows the associated graph of the network. Again, passing from the green color to the blue one we have an increase of the g-value, and the bigger red point indicates the node with the highest g-value (438585.6). The peculiar location of the node with highest betweenness in Attika depends on different factors. First of all, the Attika's UTS has more branches towards peripheral regions. This leads to having some sub-graphs (i.e. clusters) associated with the peripheral areas. Nodes - as well as links - connecting clusters in a graph, are characterized by high values of betweenness since all the paths between two clusters pass through them. Thus, several nodes with high betweenness values are located on the branches connecting peripherals. Secondly, there are a lot of nodes with very high betweenness also in the center of the network, however the rule we have adopted to choose the "triggering" node is just to select the one with highest node betweenness value, even if the difference with the second or the third ones- which could be less peripheral - is very small.

To simulate a possible capacity cascade for the two networks we simulate the removal of the node with the largest $g-$value of the graphs. The re-computation of the g-value for each node permits to identify the new failing nodes in the cascade (i.e. nodes with the capacity lower than the current load). These nodes are removed, and the process iterated until no more nodes fail. Fig. 1, c) and d) show the two final

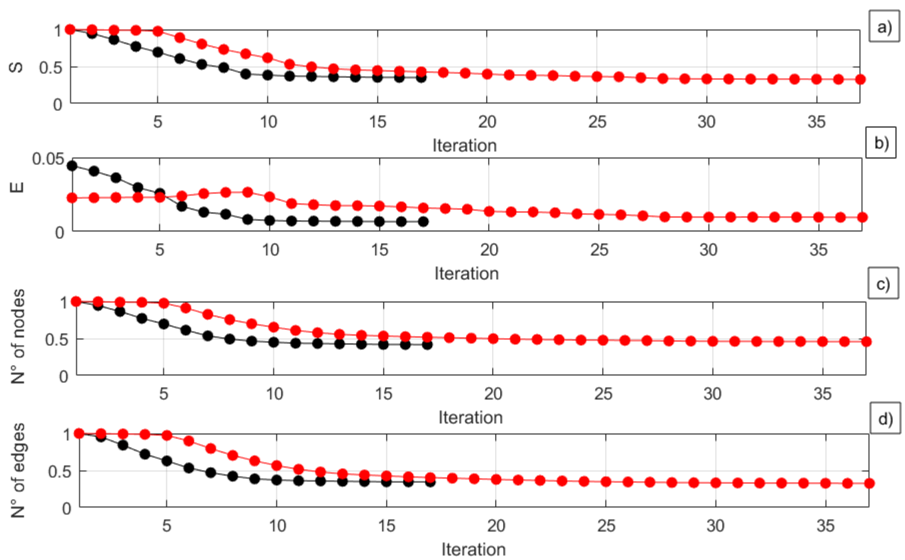

Fig. 2 Values of S (a), E (b), the relative number of remaining nodes (c), and the relative number of remaining edges (d), along the cascade. Black curves refer to the Florence UTS; red curves refer to the Attika UTS.

networks at the end of the cascade, respectively, in which the black points represent the nodes removed along the cascade. Fig. 2 a) and b) show the values of E and S computed during the cascade, for Florence (black) and Attika (red), respectively. In Fig.2, c) and d) we represent also the number of remaining nodes and edges of the two networks, during cascades. Both these quantities are divided by their corresponding values before the cascade.
For both the cases we note that the removal of a such critical nodes generate an important disruption of the network, causing the decrease of both E and S. In particular, in the Florence UTS S and E decrease from 1 to 0.35, and from 0.044 to 0.007, respectively, with a cascade consisting of 17 iterations. In the second Attika UTS case we also note a decrease of S and E, decreasing from 1 to 0.32 and from 0.023 to 0.002, respectively, with a longer cascade consisting of 37 iterations.
The decrease of such quantities is coherent with the decrease of the number of nodes and edges of the two graphs, during the cascade. However, the Attika's UTS shows a greater resilience to the cascading failure compared to the Florence's one. Indeed, it requires a largest number of iteration to reach similar values of E and S during the cascade, with no significant changes between the first and the fifth iteration.

5 Conclusions

The experimental results have demonstrated that a set of analytical functionalities can be used simulate a cascading failure and assess, dynamically along the cascade, the re-organization of flows into the network. A software tool has been developed to dynamically analyse the graph associated to the UTS, even under changing conditions, to identify the new failing components along the cascade and, therefore, the critical components which could potentially be empowered to block or at least mitigate the impact of the cascade. This analytical tool is important for assessing the resilience, as well as the efficiency, of a UTS even with respect to disruptive events starting from different locations (i.e. selecting any node as the target) and to support decisions about the capacity increase on critical nodes (i.e. those which may guarantee a lower impact).
Finally, the analytical software tool has been validated on two real UTSs, the bus transport network in Florence and the transport system in the Attika region, respectively. The analysis allowed to identify important differences between the two UTSs with respect to the impact of the corresponding cascades, even if, in both the two cases, the cascade was started by considering, as starting node, the node with highest betweenness centrality in the two networks.

Acknowledgements

This work has been supported by the RESOLUTE project (www.RESOLUTE-eu.org) and has been funded within the European Commissions H2020 Programme under contract number 653460. This paper expresses the opinions of the authors and not necessarily those of the European Commission. The European Commission is not liable for any use that may be made of the information contained in this paper.

References

1. Ash J., Newth D.: Optimizing complex networks for resilience against cascading failure. Physica A: Statistical Mechanics and its Applications. (2007) doi: $10.1016/j.physa.2006.12.058$
2. Buldyrev S., Parshani R., Paul G., Stanley H., Havlin S.: Catastrophic cascade of failures in interdependent networks. Nature. (2010) doi: $10.1038/nature08932$
3. Motter A.: Cascade control and defense in complex networks. Physical Review Letters 93. (2004) doi: $10.1103/PhysRevLett.93.098701$
4. Phadke A., Thorp J.: Expose hidden failures to prevent cascading outages [in power systems]. Computer Applications in Power, IEEE. (1996) doi: $10.1109/67.526849$
5. D'Lima M., Medda F.: A new measure of resilience: An application to the London Underground. Transportation Research Part A: Policy and Practice. (2015) doi: $10.1016/j.tra.2015.05.017$
6. Gutfraind A.: Optimizing Network Topology for Cascade Resiliance. Handbook of Optimization in Complex Networks. (2012) doi: $10.1007/978-1-4614-0857-42$
7. Candelieri A., Archetti F.: Detecting events and sentiment on twitter for improving urban mobility. in ESSEM AAMAS. Vol. 1351. pp. 106-115 (2015)
8. Soldi D., Candelieri A., Archetti F.: Resilience and vulnerability in urban water distribution networks through network theory and hydraulic simulation. Procedia Engineering. (2015) vol. 119, pp. 1259-1268, 2015. doi: $10.1016/j.proeng.2015.08.990$
9. Crucitti P., Latora V., Marchiori M.: Model for cascading failures in complex networks. Phys. Rev. E. (2004) vol. 69, n. 4, pp. 045-104, 2004. doi: $10.1103/PhysRevE.69.045104$
10. Lai Y., Motter A., Nishikawa T. Attacks and cascades in complex networks. Complex Networks: Lecture Notes in Physics. (2004) doi: $10.1007/978-3-540-44485-514$
11. Nagurney A., Qiang Q. A network efficiency measure for congested networks. EPL (Europhysics Letters) (2007) doi: $10.1209/0295-5075/79/38005$
12. Zou Z., Xiao Y., Gao J.: Robustness analysis of urban transit network based on complex networks theory. Kybernetes. (2013) vol. 42, n. 3, pp. 383-399, 2013. doi: $10.1108/03684921311323644$
13. Von Ferber C., Holovatch T., Holovatch Y., Palchykov Y.: Public transport networks: empirical analysis and modeling. The European Physical Journal B. (2009) doi: $10.1140/\mathrm{epjb}/e2009-00090-x$
14. Freeman L. C. A set of measures of centrality based on betweenness. Sociometry. (1977) doi: $10.2307/3033543$
15. Latora V., Marchiori M.: Efficient behavior of small-world networks. Phys. Rev. Lett. (2001) doi: $10.1103/PhysRevLett.87.198701$
16. Latora V., Marchiori M. Is the Boston subway a small-world network?. Physica A: Statistical Mechanics and its Applications. (2002) doi: $10.1016/S0378-4371(02)01089-0$