EFFICIENCY AND RESILIENCE ASSESSMENT UNDER CASCADING FAILURES IN TRANSIT NETWORKS

ANTONIO CANDELIERI[1,2], ILARIA GIORDANI[1,2], BRUNO G. GALUZZI[1], FRANCESCO ARCHETTI[1,2]

[1] Department of Computer Science, Systems and Communication., University of Milano-Bicocca, Milan, Italy

[2] Consorzio Milano Ricerche, Milan, Italy

ABSTRACT

This paper presents a network analysis approach to assess efficiency and resilience of public transport systems under cascading failures. Results on the two case studies of the RESOLUTE project (i.e., Florence, in Italy, and the Attika region, in Greece) are presented. Failures can be of different types (accidents, infrastructure collapses, attacks, etc.) and can lead to impacts with different severities. The key element of cascading failures is time: as time passes by, more locations or connections of the network can be affected, consecutively, as well as change their own condition. The network analysis approach proposed simulates failures propagation and evaluates the associated impacts on the transport system. Analogously to the analysis of road networks, proposed in literature, the network average efficiency and the relative size of the largest connected component have been considered for the analysis of the two RESOLUTE case studies. The cascade is simulated as follows. The node betweenness – the number of shortest paths through that node (i.e. stop/station) – is the "capacity" of that node. The worst-case of the cascading failure is considered. The node with the highest betweenness is the one triggering the cascade: it is removed and the new betweenness for every remaining node is computed, since it changes with the new shortest paths arrangement. All the nodes with betweenness higher than capacity are removed and the process continues until no more nodes can be deleted, that is the end of the cascade. Finally, efficiency and relative size of the largest connected component are computed along the cascade, to compare network at the begin and the end of the cascade. Analysis is repeated by considering the chance to increase node capacity by a given percentage, allowing to assess which is the implied improvement on resilience and on efficiency supporting a more sustainable allocation of financial resources.

Keywords: network analysis, resilience, networked infrastructure, urban transport system, cascading failure

1 INTRODUCTION

This paper stems from the activities of the European H2020 project RESOLUTE (RESilience management guidelines and Operationalization applied to Urban Transport Environment), whose general aim is the operationalization of the resilience concepts into a set of guidelines and related software for assessing resilience in a wide variety of conditions [1].

The term resilience, from its Latin root "resilire", means - loosely speaking - the capability of a system to resist, rebound or spring back in response to endogenous events (e.g. component failures) or exogenous (natural or man-made) attacks. Resilience means different things to different scientific and professional communities and is being addressed under different names; indeed, it came 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. Thus, different tools are needed to analyse and support decisions for anticipation, prevention, mitigation and restoration, depending on different types of disruptions [2], [3], [4].

This paper is focused on Urban Transport Systems (UTSs) using tools from network science to analyse resilience, namely the levels of flexibility and capacity to ensure the persistence of key functions even in the presence of cascading failures.

Methods of representation and attacks strategies come from several different communities like water distribution systems [5], transportation system [6], optimization [7], internet [8] and engineering design [9]. Coherently with the aim of RESOLUTE to provide a toolbox of wide utilization, a topological approach has been adopted in this paper, based on the description of a UTS, possible targeted attacks and cascading failures, using graph theory. This approach offers the benefit of not requiring a huge amount of data – typically the only information about the interconnections is needed to create the graph associated to the UTS infrastructure and still they can provide fundamental insights about the structural weakness of a transport network.

In this paper, we analyse the urban UTSs of Florence (Italy) and Attika region (Greece), the use cases of the RESOLUTE project. An attack simulation is considered, based on a cascading failure starting from the node with the highest betweenness value.

The structure of this paper is as follows: Section 2 presents the main elements of the graph theory and network-based centrality measures. Section 3 shows how the previous modelling tools can be used to model a UTS and some possible disruption events. Section 4 introduces the issues of modelling the cascading event and computing the ensuing dynamic network evolution. Section 5 describes the two-real life UTSs used in this paper. Finally, Section 6 describes the computational results obtained for the two UTSs in the context of cascading failure analysis.

2 BACKGROUND INFO

2.1. Basic concepts in graph theory

Let us denote a graph with $G = (V, E)$, where $V = \{1, 2, \ldots, n\}$ is the set of nodes and E is the set of edges. Each edge of G is represented by a pair of nodes (i, j) with $i \neq j$, and $i, j \in V$ and $i, j = 1, \ldots, n$. If $(i, j) \in E$, i and j are called adjacent, or neighbours. Anyone of the edges having i as one of its nodes is called incident on i. We distinguish between simple-graph, that allows only one edge between two nodes, and multi-graph, that allow multiple edges between the two nodes. However, in case of a multi-graph it is not possible to identify an edge only by its nodes, but it is necessary to use specific attributes, that characterize and distinguish each edge by another edge, represented by the same pair of nodes.

The number of neighbour of a node i, denoted by d_i, is called node degree. We denote with $\delta(G)$ and $\Delta(G)$ the minimum and the maximum degree of the nodes of G, respectively. The adjacency relationship between the nodes of G is represented through a non-negative $n \times n$ matrix A, called Adjacency Matrix of G. The entry $A_{i,j} = 1$ if i and j are adjacent nodes, and 0 otherwise. The adjacency matrix can be used also for multi-graph and graph with loops, by storing the number of edges between two vertices in the corresponding matrix element, and by allowing nonzero diagonal elements.

If all the nodes of G are pairwise adjacent, then G is called complete. A graph G is undirected if (i, j) and (j, i) represent the same edge, and it is simple if it is undirected, without self-loops (edges starting from a node and ending on the same node) and only one edge can exist between each pair of nodes (i, j), with $i \neq j$. Undirected graphs have the properties that $A_{i,j} = A_{j,i} = 1$ if $i \neq j \; \forall \, i, j \in V$. Simple graphs have the properties that $A_{i,i} = 0, \forall \, i \in V$.

A path from i to j is a sequence of distinct adjacent nodes starting from i and ending to j. The shortest path between i and j is the one related to the shortest list of adjacent nodes from

i to j, and it is usually named distance $d(i, j)$. The largest distance among each possible pair of nodes in G is named *diameter*.

A connected graph is a graph where a path exists between each pair of nodes $i, j \in V$, otherwise is called disconnected. The length of a path is the number of edges of that path. If $i, j \in V$, a geodesic between i and j is a path of the shortest length that connect i and j. The length of a geodesic between i and j, is called distance $d_{i,j}$. The maximum distance $D(G)$ between any two vertices in G is called the diameter of G.

2.2. Network-based centrality measures

In computer science and network science, a network can be defined as a graph in which nodes and/or edges have attributes (e.g. names). Networks from different domains share some properties that can be measured by a set of indices, called centrality measures [10], which can take specific ranges of values in correspondence of each specific domain.

Historically the first and conceptually simplest measure is degree centrality based on the idea that important nodes are those with the largest number of links to other nodes in the graph. The degree centrality of a node i is defined as

$$D_c(i) = d_i. \tag{1}$$

The degree can be interpreted in terms of the immediate risk of a node for catching whatever is flowing through the network (such as a virus, or some information).

Another important centrality measures is the betweenness centrality [11] that quantifies the number of times a node acts as a bridge along the shortest path between two other nodes. The betweenness can be represented as

$$B_c(i) = \frac{1}{n(n-1)} \sum_{j \neq i, k \neq i} \frac{\sigma_{jk}(i)}{\sigma_{jk}}, \tag{2}$$

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 that pass through i.

3 MODELLING OF A UTS AND DISRUPTION EVENTS USING GRAPH THEORY

3.1. Modelling of a UTS

The main elements of a UTS can be easily mapped into elements of a graph, basically nodes and edges [12] [13]. Nodes represent locations of interest on the transportation network, such as towns, bus/rail stops, road intersections, etc. while edges represent connections/links between locations, such as roads, rail lines, bus line sections, etc. A simple graph that represents the situation is in Fig. 1, a). This graph represents each station by a node, a link between nodes indicates that there is at least one route that services the two corresponding stations consecutively. No multiple links are allowed.

Extending this notion, one may either introduce multiple edges between nodes depending on the number of services between them (Fig. 1, b)). This is the way we use to describe UTS, that allows us to model multiple "routes" between two station (nodes) of the graph.
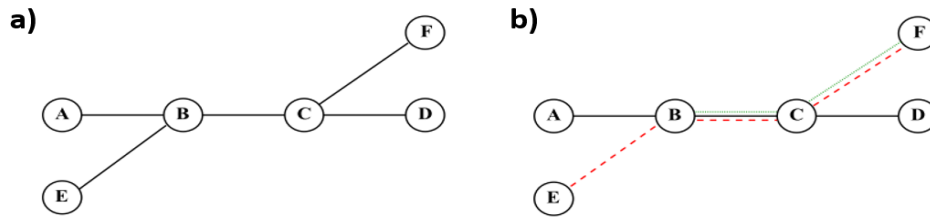
Figure 1: A piece of public transport map, in which Stations A-F are served by two lines (the black solid and red dashed line, respectively). Such map is represented using a) a simple-graph and b) multi-graph.

3.2. Modelling of UTS's disruption events using graph theory

It is important to define the possible mapping of disruption events into graph modifications. Disruptions can be of different types (accidents, infrastructure collapses and attacks, etc.) and their impact can lead to impacts of different severity: injuries, fatalities. Common disruptions, such as a road link blocked, a rail service interruption, a strike, etc., have an impact of lower severity. These events will increase the travel time for passengers and lead to cancelled trips, generating social and/or economic costs (also considering the – possibly relevant – costs for restoring the service level and for repairing or rebuilding the infrastructure).

In the following, a relevant real-world event towards graph modifications is considered: the closure of a station/stop. The simulation of such closure is done by disabling the station/stop (removal of the corresponding node in the graph) but maintaining all the paths passing through it (**Error! Reference source not found.**). However, since it is no more possible to change transportation line/route at that station, overall connectivity and paths in the graph will change. Let $G = <V, E>$ denote the original graph associated to the UTS and $k \in V$ the target node (i.e. the station/stop to close). The resulting graph after this event is $G' = <V', E'>$ where: $V' = V \setminus \{\bar{v}\}$ and

$$E' = E \setminus \{(u, \bar{v}) \vee (\bar{v}, u), u \in V\} \cup \{(u, w) \text{ such that } (u, \bar{v}) \in E \wedge (\bar{v}, w) \in E\}. \tag{3}$$
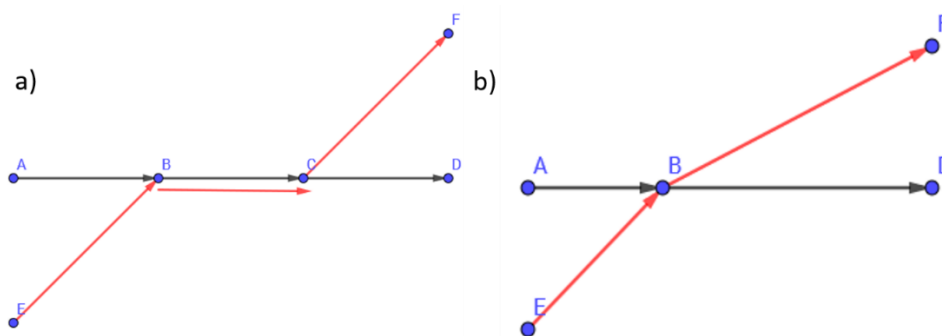

Figure 2: A piece of UTS, before (a) and after (b) the closure of the station C, as described in evqn (1).

4 CASCADING FAILURES CAUSED FROM A STATION CLOSURE

A cascading failure is a domino effect which originates when the failure of a given node triggers subsequent failures of one or several other nodes, which in turn trigger their own failures [7], [14]. Thus, the number of failed or stressed nodes increases, propagating throughout the network. Example of cascading failures in the real world is, for example, wide-scale power outage or the previous current global economic crisis.

A simple model for cascading failures on a UTS assumes that each node i is characterized by certain load L_i, that describes some quantity able to describe the importance of the node in the network. Example of these quantities can be the centrality measures [7]. In particular, a common choice is to use the value of the betweenness centrality B_i (see Section 2.2).

Each node i is characterized by a given capacity ψ_i, which is the maximum load that can be handled by that node. A natural assumption is that the capacity assigned to a node is proportional to the load that it is expected to handle, since cost constraints prohibit indiscriminately increasing a node's capacity

$$\psi_i = (1 + \alpha) \cdot L_i(0) \quad \forall i \in V^0 \tag{4}$$

where $\alpha \geq 0$ is a tolerance parameter which quantifies the excess load that a given node can handle, and V^0 is the set of the nodes of the original graph.

Cascading failures are initiated when a heavily loaded node is lost for some reason, and the load on that node must be redistributed to other nodes in the network. The removal of the node simulates the loss of the node. The redistribution of the load to the other nodes in the network require the re-computation of the load value $L_i(t + 1)$ for each node i. This re-computation can radically alter the values of the loads on the network. Certain nodes can have a load smaller than their previous value, and some others can have a higher one. In particular, if a node i has a new load $L_i(t + 1) > \psi_i$, then this node also fails. These failures can, in turn, trigger more failures, thus leading to a cascade. The process iterates until no more nodes must be removed from the network, that means when existing an iteration $\bar{t}$, for which

$$L_i(\bar{t} + 1) \leq \psi_i \quad \forall i \in V^{\bar{t}} \tag{5}$$

where V^t is the set of the nodes of the network, at iteration $\bar{t}$, with $V^{\bar{t}} \subseteq \ldots, \subseteq V^0$.

Two different parameters are considered to measure the effectiveness of a network and its resiliance at the begin and at the end of the cascade. The first is the size of the largest connected component (S) [12] [14] defined as

$$S = \frac{N'}{N}, \tag{6}$$

where N' and N are, respectively, the number of nodes in the largest connected component at the end and at the begin of cascade. The second is the network efficiency (E) [15], defined as

$$E = \frac{1}{n(n - 1)} \sum_{i,j \in V, i \neq j} \frac{1}{d_{ij}}, \tag{7}$$

where $d_{i,j} = d(i, j)$ represent the shortest path between a node i and j, named distance. Normalization by *n(n-1)* ensures that $E \leq 1$, and 1 represent a complete graph.

Logically, capacity could be increased with the aim to contrast cascades of failures in the network. Thus, the parameter α, can be used in the algorithm to increase the capacity of every node from its "baseline" value to twice, according to a given step. This fact allows for computing E and S for different configurations of the node capacities.

5 CASE STUDIES

Two real-world case studies have been considered in this paper, that are the test cases considered in RESOLUTE.

The first one is the public bus transportation in Florence. This city is one of the biggest towns in central Italy, with a population of about 400.000 inhabitants and a surface of about 100 km^2. **Error! Reference source not found.**, a) shows the corresponding graph, consisting of 999 nodes and 3226 edges. To improve the visualization, we did not draw multiple edges between two nodes.

The second UTS considered is bigger than the first one and consists of the public transportation network of the Attika region. This region is located on the eastern edge of Central Greece and covers about 3.808 km^2. In addition to Athens, it contains within its area the cities of Piraeus, Eleusis, Megara, Laurium, and Marathon, as well as a small part of the Peloponnese peninsula and some islands. About 3.750.000 people live in the region, of whom more than 95% are inhabitants of the Athens metropolitan area.

The public transportation network of the Attika region includes bus, tramway and subway. **Error! Reference source not found.**, b) shows the corresponding graph, consisting of 7681 nodes and 18128 edges. Again, to improve the visualization, we did not draw multiple edges between two nodes.

These two UTS are modelled through a directed multi-graph, because more than one route/line may connect two stations.
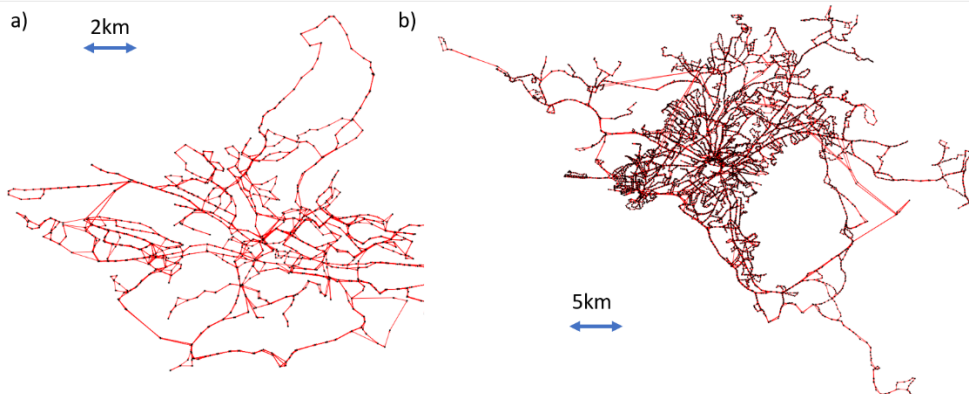

Figure 3: a) The graph associated to the PTN of Florence. The black points correspond to the node(stops) of the network, whereas the red line represent the edges (links). b) The graph associated to the PTN of the Attika region.

Both the network result connected, and therefore their value of S is 1. The value of E for the Florence's UTS is   whereas the value of E is 0.044, whereas for the Attika's UTS is 0.023.

6 EXPERIMENTAL RESULTS: SIMULATION OF CASCADING FAILURE

To simulate a possible capacity cascade for the two networks, we simulate the closure of the station with the highest value of betweenness. So, we remove this node according to the

mechanism described in Sect.3.2. The re-computation of the betweenness 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, as described in Sect. 4.

Fig. 4 summarizes the value of E and S computed during the cascade, setting the capacity of the nodes as described in eqn (4), with $\alpha = 0$. The black and red curves represent the Florence's UTS, and the Attika's UTS, respectively.

In the first case (Florence's UTS) we note a decrease of the value of S and E, that pass from 1 no 0.35, and from 0.044 to 0.007, respectively, and the length of the cascade is 17. Also, in the second case (Attika's UTS) we note a decrease of the value of S and E, that pass from 1 to 0.32 and from 0.023 to 0.002, respectively, and the length of the cascade is 37.
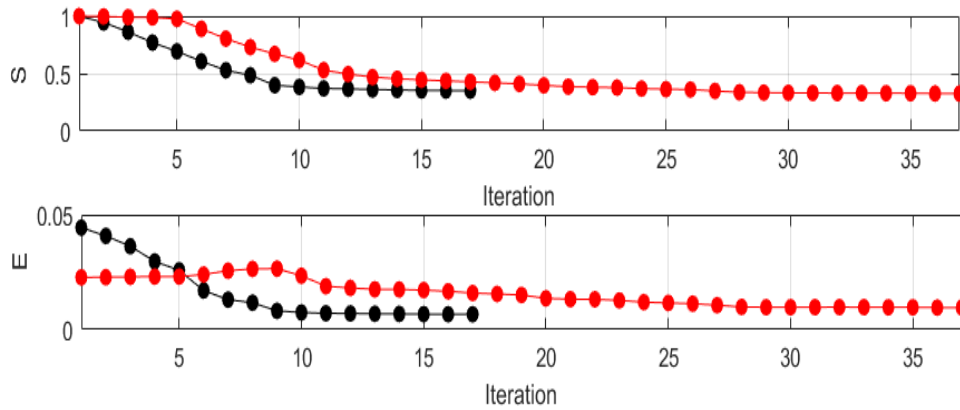

Figure 4: Values of S (a), E (b), during the cascade, with $\alpha = 0$. The black curves refer to the Florence's UTS. The black curves refer to the Attika's UTS.

Fig. 5, a) and b) show the value of S and E, respectively, obtained at the end of the cascate, for different values of α, between 0 and 1, for the Florence's UTS. The baseline values (α = 0) of S and E are 0.23 and 0.0046, respectively. These values remain substantially constant between α = 0 to α = 0.4. Then we have a significant increase of their values passing from α = 0.4 to α = 0.5. Finally, they became stable when α is 0.6 or higher. In this range of α - values, we can observe that the values of E and S result substantially unaltered at the end of the cascade. In conclusion, the increase of E and S result correlated each other, and with respect to the increase of α for the Florence UTS: a specific value of α exists after that the damages of the cascading failure (decrease of E and S) result to be completely mitigated.

Quite different results have been obtained for the UTS of the Attika region (Figure 6). A general increase of the value of E and S can be observed with α increasing. The baseline values of S and E are 0.34 and 0.0021, respectively. For α > 0 the value of E and S, obtained at the end of the cascade, result greater than their baseline values. Both these quantities increase significantly when the capacity of each node pass from α = 0.3 to α = 0.4. However, increasing the value of α does not imply always an increase of E and S. More in details, we can observe a significant decrease of the value of S and E between α = 0.6 to α = 0.7, and there is any value of α ∈ [0,1], after that the damages of the cascading failure (decrease of E and S) result to be completely mitigated.

Therefore, there is in any case a substantial difference between the two UTSs: for the Florence's UTS it is possible almost completely mitigate the damage of the cascade increasing the capacity of the nodes of about 50% ($\alpha = 0.5$), whereas for the Attika's it is only possible to quite mitigate the damage of about 40% ($\alpha = 0.4$). For this last UTS an increase of the capacity of the nodes does not imply lower damages of the network caused by the cascading failure, necessarily
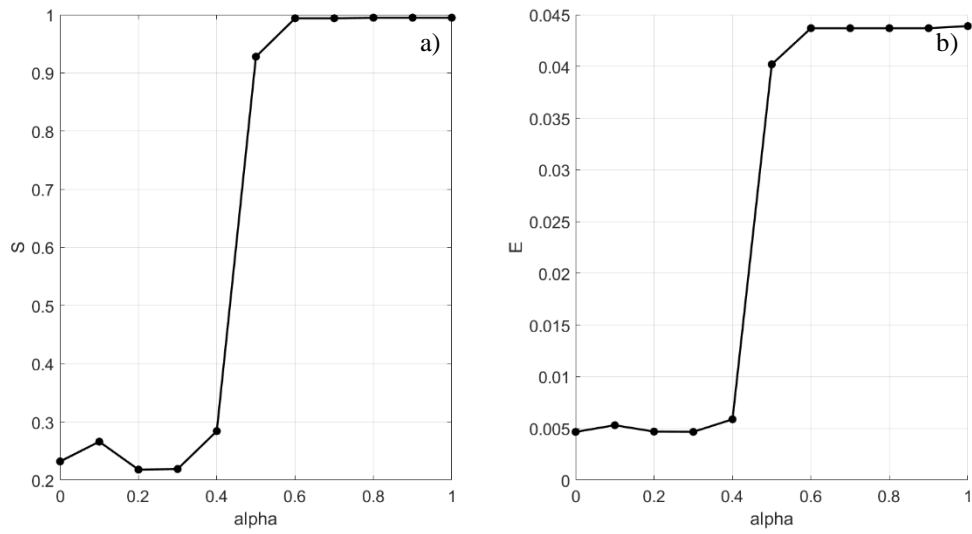

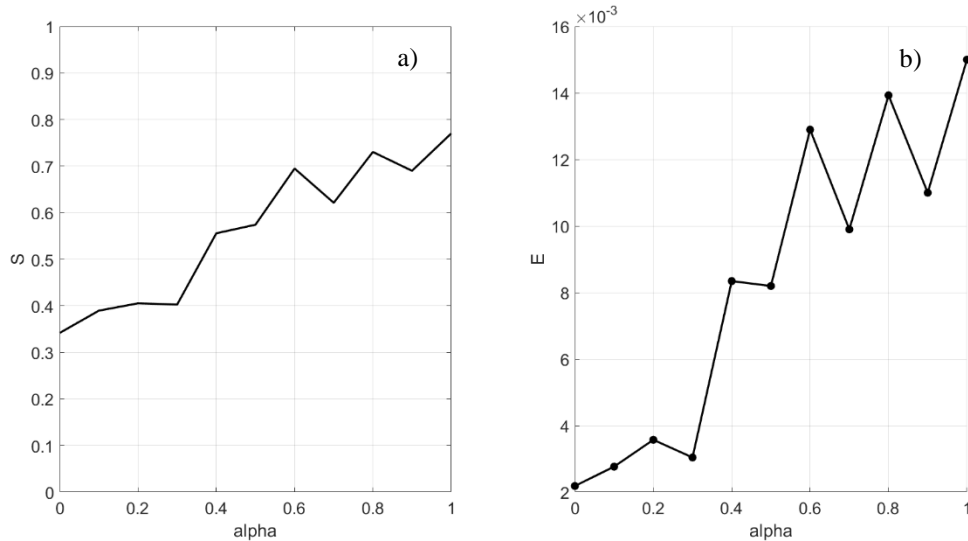Figure 5: Value of S (a) and E (b) at the end of the cascade for different values of alpha, for the Florence's UTS.

Figure 6: Value of S (a) and E (b) at the end of the cascade for different values of alpha, for the Attika's UTS.

7 CONCLUSIONS

With respect to network analysis, the experimental results obtained on the two public transportation networks of the RESOLUTE project have demonstrated that a set of analytical functionalities can be used to identify critical components (nodes as well as edges) and evaluate, dynamically, the new setting induced by a disruption event as well as the simulation of cascading effects.

The capability to interact with and analyse graphs, dynamically and efficiently, is a core contribution to support decision-making activities in the UTS resilience management.

Finally, a computational framework of modelling and analysing cascading failures has been developed and preliminarily validated on the two UTS. To avoid the inherent computational intractability of exact computation of graph-based resilience measures two heuristic resilience metrics have been validated.

The main output is a series of two measures, namely the average network efficiency (E) and the relative size of the largest connected component (S), with respect to the capacity of nodes (i.e. stops/stations) and when the triggering failure is associated to the removal of the node with the highest betweenness.

The re-computation of the load for each node (i.e. the betweenness) 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. The functionality has been evaluated following the aforementioned schema and proved to be useful in supporting more effective planning strategies. More specifically, this analysis permits to better understand how the UTS behaves under a cascading effect situation and may also support in planning capacity of nodes (stop/stations) to better respond to the cascade. However, it may also be used to evaluate cascading effects in "real" situations, where the triggering event can be the removal of one – or even more – node different from the one having highest betweenness and where the load is computed according to some weights on

the edges of the graph, for instance, the number of lines, vehicles or passengers passing through them.

REFERENCES

[1]   Bellini E, Nesi P, Coconea L, et al. Towards resilience operationalization in urban transport system: The RESOLUTE project approach. *26th Eur Saf Reliab Conf ESREL 2016* 2017; 345.

[2]   Ferreira P, Simoes A. *Conceptual framework*, Resolute-Eu.Org (2016).

[3]   Gaitanidou E, Tsami M. *Ermg Adaptation To Uts*, Resolute-Eu.Org (2016).

[4]   Archetti F, Candelieri A, Giordani I, et al. *Application Framework*, Resolute-Eu.Org (2017).

[5]   Soldi D, Candelieri A, Archetti F. Resilience and vulnerability in urban water distribution networks through network theory and hydraulic simulation. *Procedia Eng* 2015; 119: 1259–1268.

[6]   Berdica K. An introduction to road vulnerability: What has been done, is done and should be done. *Transp Policy* 2002; 9: 117–127.

[7]   Ash J, Newth D. Optimizing complex networks for resilience against cascading failure. *Physica A* 2007; 380: 673–683.

[8]   Cohen R, Erez K, Ben-Avraham D, et al. Resilience of the Internet to random breakdowns. *Phys Rev Lett* 2000; 85: 4626–4628.

[9]   Agarwal J, Blockley D, Woodman N. Vulnerability of structural systems. *Struct Saf* 2003; 25: 263–286.

[10]   Albert R, Barabási A-L. Statistical mechanics of complex networks. *Rev Mod Phys* 2002; 74: 47–97.

[11]   Freeman LC. A Set of Measures of Centrality Based on Betweenness. *Sociometry* 1977; 40: 35–41.

[12]   Berche B, von Ferber C, Holovatch T, et al. Public Transport Networks under Random Failure and Directed Attack. *Dyn Socio-Economic Syst* 2010; 2: 42–54.

[13]   von Ferber C, Holovatch T, Holovatch Y, et al. Public transport networks: empirical analysis and modeling. *Eur Phys J B* 2009; 68: 261–275.

[14]   Zou Z, Xiao Y, Gao J. Robustness analysis of urban transit network based on complex networks theory. *Kybernetes* 2013; 42: 383–399.

[15]   Latora V, Marchiori M. A measure of centrality based on network efficiency. *New J Phys*; 9. Epub ahead of print 2007. DOI: 10.1088/1367-2630/9/6/188.