Modelling and Simulation of Railway Networks for Resilience Analysis

Kushal Srivastava[1], Corinna Köpke[1], Katja Faist[1], Johannes Walter[1], John Marschalk Berry[2], Claudio Porretti[3], and Alexander Stolz[1,4]

[1] Fraunhofer Institute for High-Speed Dynamics, Ernst-Mach-Institut, EMI, Am Klingelberg 1, 79588 Efringen-Kirchen, Germany
kushal.srivastava@emi.fraunhofer.de
https://www.emi.fraunhofer.de
[2] Rete Ferroviaria Italiana S.p.A, Piazza della Croce Rossa, 1 - 00161 Rome, Italy
j.berry@rfi.it
https://www.rfi.it/
[3] Leonardo S.p.A., Piazza Monte Grappa n.4, 00195 Rome, Italy
claudio.porretti@leonardo.com
https://www.leonardo.com/en/home
[4] Albert-Ludwigs-Universität Freiburg, Emmy-Noether-StraSSe 2, 79110 Freiburg im Breisgau, Germany.
alexander.stolz@mail.inatech.uni-freiburg.de
https://uni-freiburg.de/

Abstract. The work focuses on the impact of disruptions on a railway transportation network. The modeling of the transportation network with the help of graph theory is presented and criticality/vulnerability assessment and impact propagation in these networks is studied. Furthermore, the work emulates defined mitigation measures in the modeled network and quantifies the resilience of the network. The results are produced from an agent-based simulation tool called CaESAR (Cascading Effects Simulation in Areas for increasing Resilience) that uses network graphs in cooperation with their behavioral characteristics. The tool is under integration with a broader framework (S4RIS platform) designed under the EU H2020 project Safety4Rails aimed at integrating multiple solutions to be made available to operators and first responders for better responses in case of threats and disruptions.

Keywords: Resilience Indicators · Critical Infrastructure · Impact Propagation · Transport Networks · Graph Theory.

1 Introduction

Critical Infrastructures (CI) play a key role in the daily functioning of society and thus are key to overall economy of a country. Furthermore, as they grow, CI get more complex and consequently more sensitive regarding disruptive events and interdependencies. Due to this criticality, CIs have been studied for quite some time interms of risks involved and their resilience. In the project

Safety4Rails, resilience has been defined as "the ability to repel, prepare for, take into account, absorb, recover from and adapt ever more successfully to actual or potential adverse events" [3] [14]. The resilience cycle as followed in the project consists of different but overlapping and interwined phases namely, identification, protection, detection, response and recovery [3]. Several resilience studies in the transportation area use topological models [4]. In [1], an undirected graph is used to model the transportation network. The resilience of every node is computed as weighted average number of reliable independent paths with all other city nodes in the network (betweeness centrality). The overall resilience is then a weighted sum of all node resiliences. [2] examines the interdependent rail networks in rush hours. The work emphasizes that topological shapes of the network play key role in dynamics of the cascades and conclude that in complex networks, cascade effects are more responsible for poor performance than failures itself.

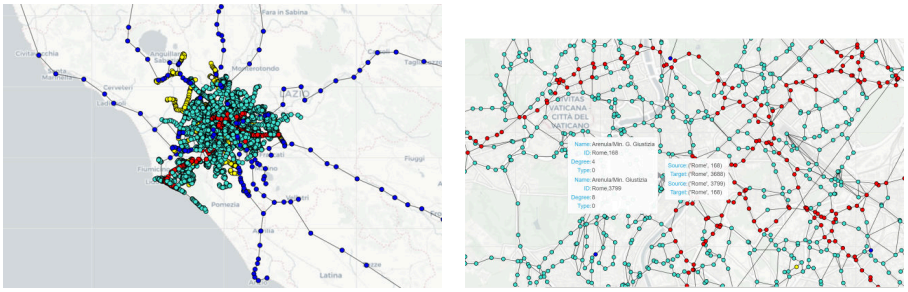
With an increase in digitization of transportation networks (similar to other forms of CIs), there is potential for cyber attacks that can maximize disruptions and cause delays in recovery. One example is the jamming of CCTV monitoring cameras to block visibility, occupancy, and other forms of on-ground information to operators and first responders. This can lead to delays in time critical decisions or mismanagement of resources that can prove to be bottlenecks for trivial rescue operations. These are challenges that are assessed in this paper using the CaESAR tool. The paper is further divided into four sections. Section 2 presents the modelling technique. It briefly discusses aspects of graph theory for criticality assessment and impact propagation. Section 3 discusses the problem definition in the form of a use case. Section 4 presents the results and corresponding visualizations. Finally, section 5 provides a summary and brief outlook of the work.

2 Modelling of the Network

This section describes the network model built for S4R project including the models for resilience quantification, criticality assessment, impact propagation as well as a description of the use case. The Safety4Rails project includes four Simulation Exercises. The networks for the Rome exercise in Safet4Rails were generated from open source data available at OpenMobilityData [15]. This data is based on The General Transit Feed Specification (GTFS) (GTFS Static Overview | Static Transit | Google Developers) [16], which is a common format for public transport networks, including schedules and geographic information. This data is used to generate a representation of the public transport network consisting of nodes and edges. The nodes are generated according to the stops.txt, which describes stops and their geographic locations as well as some further information, like the location type. Since the GTFS data contain several stops with the same name, they are consolidated to one node, where the geo-location is the average of the geo-locations of all stops with the same name.

Each node contains information regarding the transportation types serving it, i.e. by which lines it is served.

Nodes are connected based on trips described in Trips.txt. A trip contains a series of stops, where the vehicle travels along a route at a specific time. The stops located on the corresponding trip are then connected according to it. Added edges are unique, i.e. they represent the opportunity to travel from one station to another one. Figure 1a represents the modelled network using the open source data. This modelling consits of bus (yellow), tram(turquoise), metro (red) and train (blue) stations. The visualizations generated in CaESAR are interactive, and it is possible to hover over the graph nodes and get more information including name of nodes, their degree, geo-location (as shown in figure 1b).



(a) (b)

Fig. 1: The figure shows the modelled network in the exercise. a) The Rome testbed as modelled using open source data available at OpenMobilityData. b) Zoomed-in version of the testbed to show the interactive dialog box and the connections.

The modelling uses bi-directional graph to better represent the flow of traffic throughout the network. The implementation uses multiple python modules including Networkx[21], GeoPandas [23], Shapely [22], Bokeh [24] and NumPy [20]. The threats are modelled as objects with attributes including time of attack, list of nodes to attack and time for repair. The nodes in-turn have attributes including name, geo-location, capacity and repair times. The repair times are used in the recovery strategy of the nodes after impact. The nodes are recovered after their individual repair time has elapsed after damage.

2.1 Numerical Simulation using CaESAR

CaESAR is a python and C++ based tool that can perform offline and online analysis of impacts on networks. Offline analysis is stochastic analysis of the net-

work with single-point/multi-point failures to identify worst performing combinations, while online analysis here is performing analysis to specific threats/failures in network in real-time, based on data from a suitable integrated platform. CaESAR has been used to analyse cascading effects in different types of infrastructures such as water, electricity and mobile phone grids. In project RESISTO [18], CaESAR has been used to model and understand telecommunication grids [12], in SATIE [17], [19] for understanding airport networks and now is being utilized in Safety4Rails for analysis of railway networks [3]. It uses network and threats modelled as described above along with probabilities for failures, delays and repair times with certain variance to quantify and analyze resilience of the networks. In order to understand cascades in the networks, it is also capable of employing different propagation algorithms. As a result of this analysis, CaESAR delivers visualizations of propagation of impact in the network highlighting damage and recovery, resilience curves for the network and time-series of states of the components in the network.

2.2 Resilience quantification

Resilience of the network is considered as the quantified performance of the system, which is the percentage of active nodes. The state of the nodes can be considered as binary or varying between $[0, 1]$. For a network with N nodes, let s_i represent the state of the i^{th} node, then performance θ at any time-step (t) can be defined as average of all states of nodes in the network:

$$\theta = \frac{1}{N} \sum_{i=1}^{N} s_i \tag{1}$$

While this is a basic form of performance quantification, some key performance indicators can be derived from the resulting graphs. This includes:

1. The maximum rate of performance degradation, $d(theta)/dt$. An alternate to this is the time taken to reach minimum performance.
2. Total downtime.
3. Total time for recovery.
4. Minimum performance.
5. Maximum state of recovery.

These indicators have been applied on the railway simulation grid and the results have been explained in section 4

2.3 Criticality assessment

The assessment of criticality of nodes in the network is based on properties from graph theory. These properties have also been widely used by researchers when quantifying vulnerabilities in different applications of graph networks. An example is the work by Artemis P. and Eusebi C. [8], wherein authors have used

graph properties to derive connectivity and activity density of transportation nodes over time. Another example is [10], wherein authours use the degree of each node to measure their criticality. This work focuses on identification of critical nodes from the perspective of threats and potential disruptions and applies these metrics (degree and betweenness centrality) to achieve it. In the integrated platform of S4R, in the case of detection of an event, the gathered list of critcal components are communicated to the operators, which helps in an informed decision making. The two metrics "degree of nodes" and "betweeness centrality" which are briefly explained as follows:

Degree of nodes: The degree of nodes is defined as number of connections or edges the node has to other nodes.

Betweeness Centrality (BC): The betweeness centrality defines how much a given node is in-between others [13]. It is measured with number of shortest paths between any two nodes in the graph, which passes through the considered node, for which the BC is being computed [13]. A target node will have higher BC if it appears in many shortest paths. These nodes would correspond to central intersections from a topological perspective [6]. These nodes would be expected to have higher traffic than those with low BC values. With N as the total number of nodes in the graph, the BC is often normalized by a factor of $N(N-1)$. Mathematically the BC of a node v_k is defined as [6]:

$$BC(v_k) = \sum_{v_i \neq v_k \neq v_j} \frac{\sigma_{v_i v_j}(v_k)}{\sigma_{v_i v_j}} \tag{2}$$

where $\sigma_{v_i v_j}$ is the total number of shortest paths from node v_i to node v_j, and $\sigma_{v_i v_j}(v_k)$ is the number of those paths that cross v_k [6]. In order to demonstrate the effectiveness of these assessments, in addition to the use case, result section also discusses the impact of failures in these nodes with the help of resilience curves.

2.4 Impact propagation

To compute the resilience with respect to disruptive events, CaESAR uses impact propagation to estimate the consequences on the network. In the current study, connectivity of stations/nodes is used. Implicitly, failure in a station/track would first reflect in the immedeately connected station and from there propagate into the network. With this as a motivation, a propagation algorithm is designed to reflect this behaviour. A custom delay is added before propagation to next station. Since transportation network (individual or different networks modelled as dependent) are fully connected, meaning there is no island/disconnected node, if run long enough the simulation will show an impact on every node of the network. In terms of impact propagation, impact does not necessarily mean physical damage. In this this context, impact can for example represents

number of passengers reaching their destinations. Figure 3 represents the stages of a connectivity based impact propagation in a directed graph. p_i represents the probability of propagation to the corresponding connected node. To ensure flexibility, the delay is parameterized and represented here as m.

At time-step T, impacted node is *1*, at $t + m$, impacted nodes are *1, 2, 3* and after $t + 2m$ time-steps, impact propagates to the next stage of connected nodes *1, 2, 3, 4, 5, 6*. The recovery of nodes is independent of each other and depends on their individual repair times. With connectivity based propagation, this needs to be further tuned in co-ordination with ground staff and expert knowledge to introduce limits based on the type of impact, type of compoents etc.

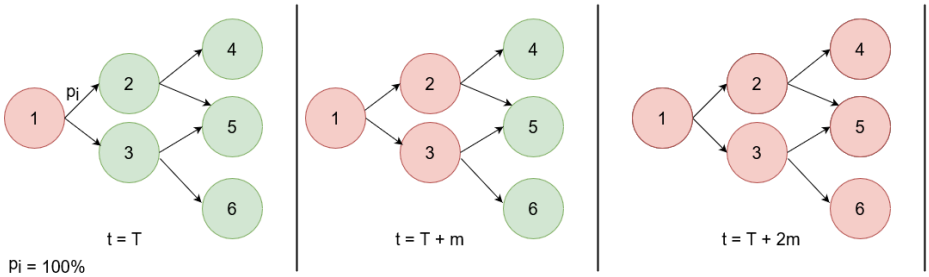


Fig. 2: Demonstration of connectivity based impact propagation in the network with 100% propagation probability and m time steps between cascade to next stages.

3 Use Case

The exercise considers a hypothetical scenario where there is a physical attack at the Rome Termini station coordinated with a cyber attack (DoS, Denial of Service) on the vulnerable CCTV system installed at the stations. With this information together, a threat is designed for the simulator. The analysis is performed in two phases of the resilience cycle (see Section 1), the prevention phase and the response phase. In the prevention phase, offline analysis is performed using aspects of graph theory and what-if analysis. Vulnerable nodes are computed based on their degree of connectivity and BC. These nodes are then plugged in the simulator as impact soure to evaluate the impact on the overall resilience of the network.

In the response phase, online analysis is performed, where in CaESAR depends on the state information from other detectors in the system to simulate impact. The obtained failures in the systems/components are mapped to threats

in the network and simulated. Further analysis is performed based on the resilience curves and impact propagation visualizations. The scenario is based on the open-source network generated for Rome, where there is a co-ordinated disruption of a railway station with very high degree of connectivity using a physical attack and blocking access to key surveillance components using DOS attack. Furthermore, the time of disruption is chosen to be inline with the rush-hour to maximize impact. The physical attack is identified using Ganimede tool which focuses detection of objects and people in each frame and their movement, while the cyber attack is detected using the platform CuriX. The simulation is setup with organized attacks at simulation time-step 5. The mitigation measures are mentioned in table 1 and discussed further in the results section.

The tools were integrated using Kafka platform on Distributed Messaging System (DMS). Information is populated in DMS using different channels referred to as topics. Tools can subscribe to these topics in forms of consumer groups and get a constant feed of the messages being sent to the topic. The format of the message is pre-defined. In this exercise, JSON format is used. With the help of some specific fields like data source, event category, asset ID and event description, the receivers can program scripts to automatically filter information of use and post-process them for suitable analysis and visualization.

For this exercise, to present mitigation measures, four different scenarios are considered as follows:

1. **No measure**: Impact propagates throughout the network.
2. **Placement of security guards**: Reduction in extent of impact and repair time.
3. **Design and implementation of suitable evacuation routes**: Reduction in extent of impact and repair time.
4. **Re-routing of rail traffic**: Reduced repair time for the impacted station and five directly connected stations

These scenarios are parametrically summarized in the following table.

Table 1: Table of scenarios, with and without mitigation measures employed in the exercise along with the KPIs.

Scenario	Performance during impact	Reduction of repair time	Minimum Performance
1	0%	0%	40.98
2	50%	10%	53.74
3	75%	50%	82.21
4	0%	15%	59.04

4 Results

In order to respect the confidentiality of the exercise, the result of the assessment is presented using anonymized stations. In the offline analysis, five different stations from the network are selected with decreasing BC and degrees (refer 2). The threats are designed accordingly and plugged into the simulator. The repair time of the nodes in the network is considered to be dependent on the degree of the node, and is defined as $20+5*D_n$, where D_n is the degree of the corresponding node. The propagation delay is considered to be 2 time steps between stages of the graph. Here, 20 is an assumed base value. With a minimun degree of 2, the lowest repair time of such a node in the network will be 30 time steps. As this is a theoretical study and repair times are sensitive information, the overall recovery has been presented in an abstract manner using time-steps. Specific infomation from the end-users can be used to scale the overall recovery to reflect the reality. In reference to the figure 3, here the time delay between cascades is, $m = 2$. Another critical assumption used in the study is that the time taken to recover a node is independent of the extent of impact (state of the node). The state information is used only to quantize the resilience with respect to a threat.

Table 2: Table of nodes considered with their properties, and measured KPIs for the system dervied from the resilience curves

Station	Degree	BC	Maximum gradient	Time of outage (timesteps)	Minimum performance	Time to minimum performance
Station 1	12	0.2975	40.98	208	71.28%	63
Station 2	10	0.2531	53.74	218	71.869%	63
Station 3	14	0.0190	82.21	232	72.35%	81
Station 4	4	0.0037	59.04	284	76.20%	121
Station 5	2	0.0	59.04	294	77.82%	175

Table 2 gives the criticality analysis of the network, with both the attributes (degree and BC) and derived KPIs. As seen in figure 4a, for nodes with higher BC, there is an early decline in the performance. As these are highly connected nodes, impact quickly propagates in all directions of the network. Due to this, large number of nodes are impacted very quickly and so performance rapidly drops. For nodes with lower values of degree and BC, due to a lower connectivity, the impact spreads to smaller number of nodes in the beginning before reaching more dense sections of the network. This is reflected with lower gradient at the start and higher time to maximum impact (that is minimum performance). This gives more time to operators and first responders to organize suitable mitigation measures to curtail the impact. The maximum value of gradient alone does not reflect the right criticality, as it does not depend on the start of impact rather the overall impact propagation in network. Hence, even if an impact starts in a secluded node with low BC, when run long enough it reaches the center of the

network. From this point, the network has same impact with cascade as it would if the damage originated in one of the dense nodes. Another key aspect is the time to minimum performance, for nodes with low BC, the performance impact is low in the begining until the cascades. This promotes mitigation measures in respect to isolations, where such sections can be removed from the overall network to limit the damage. In terms of minimum performance, the higher the connectivity of the nodes, the lower the minimum performance is. As impact propagation uses connectivity, by the time the recovery in the network begins, higher number of nodes (that is all connected nodes) are damaged as compared to impacts originating from lower connectivity nodes.

In the online analysis, different detector tools generate events. CaESAR polls for these messages and triggers on suitable messages. In the result, an event of physical attack (explosion) at one of the central stations is considered and analyzed. The repair and propagation times are as mentioned in table 1. Three mitigation measures are considered as defined in table 1. Figure 4b presents the resilience curves for the scenario with and without mitigation measures. As expected, the case of no mitigation measure has the worst performance. In absence of any mitigation measures, crowd formation is expected and disruptions in multiple lines, as the station impacted is highly connected. Due to a high BC value, the impacted station lies along the route to large section of stations, disruptions are visible in multiple lines. This is represented in the form of a cascade using connectivity. Figure 3 demonstrates this cascade for two different selection of stations. Station 1 (left) and station 2(right) represents two stations placed centrally and remotely. The cascade is demonstrated using with snapshots of the state of the network at different stages (timesteps) in the simulation. The initial stage is at $t = 10$, and the final at $t = 210$. The color coding in the impact propagation gif is programmed with red representing completely failed nodes and then a spectrum to blue representing the extent of recovery. Once the node completely recovers, its color is overridden to green to distinguish undamaged nodes in the network.

Employing guards on the stations can facilitate organized and faster evacuation. This is modelled with lower recovery time, which presents as lower drop in performance of the network when compared to no mitigation. Similarly, for rerouting traffic, it is represented with reduction in repair times. Final consideration is the presence of planned evacuation routes. With the help of event inspectors on ground and co-operation with law enforcement agencies (LEA), a larger portion of crowds can be evacuated with minimum damage. This is supposed to provide LEA and first responders with sufficient opportunity to clear the scene and resume operations. With faster control over the system, certain sections of the stations can still be operational, hence the state of the impacted stations is not reduced to zero. The state can be further tweaked based on severity of impact to better quantify the performance.

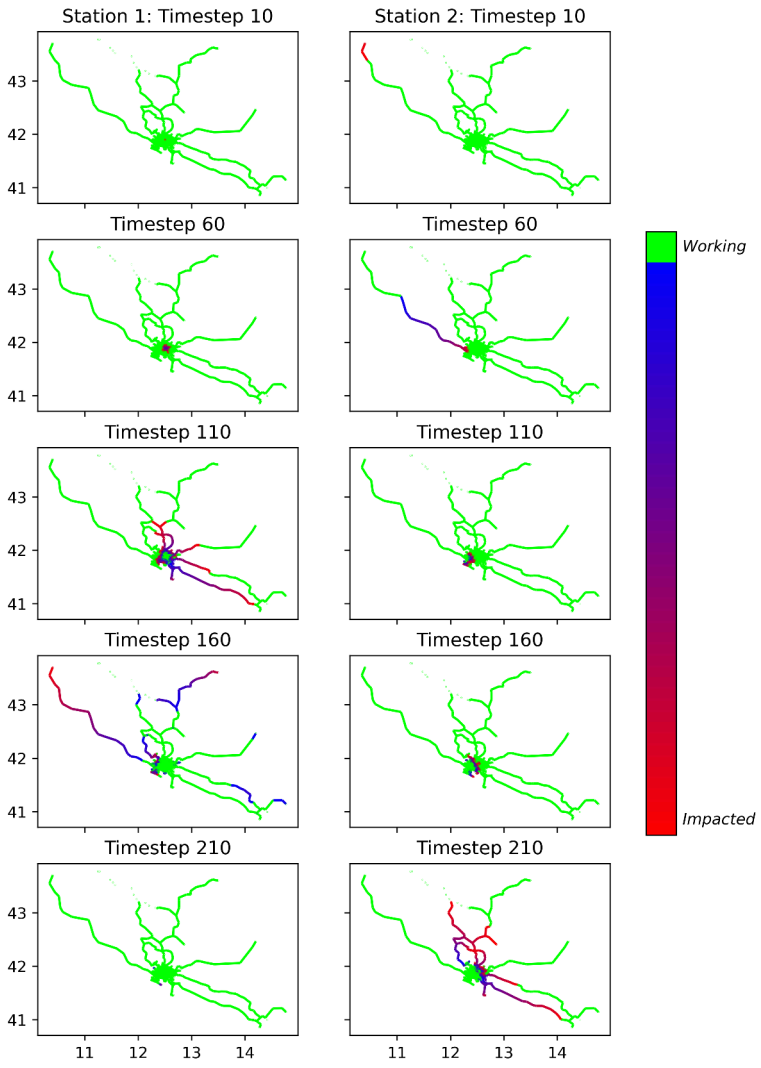
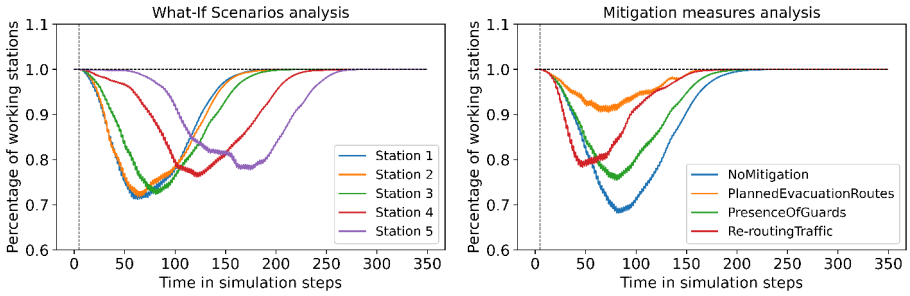


Fig. 3: Demonstration of connectivity based impact propagation in the network with 100% propagation probability and 2 time steps between cascade to next stages.



(a) (b)

Fig. 4: a) Resilience of the network with five cases of impacted stations with decreasing BC. b) Resilience curves for original use case along with different mitigation measures. The vertical black line (at timestep = 5) represents the time of attack on the network.

5 Conclusion

This work has focused on modelling of transportation networks with interconnections for criticality and resilience analysis. Special focus was on implementation of mitigation measures and rating them to understand their effectiveness. With an increase in dependency on IoT, network components play a critical role in both operations and recovery of the network, hence, an organized cyber-physical attack has been considered. The analysis was performed using CaESAR in the Safety4Rails integrated platform with detection messages being received from other partners using DMS. Two different properties of graph theory have been used to identify critical components in the network, which are then further utilized to create threats and simulated to estimate the resilience of networks. With the quantification of area under the curve, it can be verified that the impact is more severe with nodes of higher degree and betweeness centrality. This finding is also in line with other researches in this direction. In order to understand the cascading effects of its impacts, connectivity-based propagation has been used. It has been demonstrated that disruptions in any part of the network will have impacts in the whole network, unless repairs are fast. The severity of the impact of these disruptions is also dependent on the topological location of the nodes. This finding is also in line with the criticality analysis. In terms of mitigation measures, three different approaches have been examined, with placement of security guards to handle the crowd, planned evacuation routes and remodulation of rail traffic. After discussions with end-users, the measures were translated to simulation environment and it has been demonstrated that the correctly implemented evacuation routes will be very effective. In the future, more cooperation with end-users and suitable partners further mitigation measures can be studied to generate a comprehensive survey.

Acknowledgements This project has received funding from the European Union's Horizon 2020 Framework Programme for Secure societies - Protecting freedom and security of Europe and its citizens and innovation programme under grant agreement No 883532. This output reflects the views only of the author(s), and the European Union cannot be held responsible for any use which may be made of the information contained therein. For more information on the project see: https://safety4rails.eu/.

References

1. W. H. Ip and D. Wang: Resilience Evaluation Approach of Transportation Networks, In: International Joint Conference on Computational Sciences and Optimization, 2009, pp. 618-622. https://doi.org/ 10.1109/CSO.2009.294
2. Pagani Alessio, Mosquera Guillem, Alturki Aseel, Johnson Samuel, Jarvis Stephen, Wilson Alan, Guo Weisi and Varga Liz: Resilience or robustness: identifying topological vulnerabilities in rail networks, 2019, R. Soc. open sci.6181301181301. http://doi.org/10.1098/rsos.181301
3. Castanier, Bruno, Cepin, Marko, Bigaud, David, Berenguer, Christophe, Miller, Natalie, Satsrisakul, Yupak, Faist, Katja, Fehling-Kaschek, Mirjam, Crabbe, Stephen, Poliotti, Mauro, Naderpajouh, Nader, Setunge, Sujeeva, Ergün, Salih, Kanak, Alper, Tanrseven, Sercan, Lekidis, Alexios, Matsika, Emmanuel, Sick, Philipp, Cazzato, Eros. A Risk and Resilience Assessment Approach for Railway Networks, In: 31st European Safety and Reliability Conference, 2021. https://doi.org/10.3850/978-981-18-2016-8_402-cd.
4. BEINOVI, Nikola. Resilience in railway transport systems: a literature review and research agenda. Transport Reviews, 2020, 40. Jg., Nr. 4, S. 457-478
5. W. H. Ip and D. Wang, "Resilience and Friability of Transportation Networks: Evaluation, Analysis and Optimization," in IEEE Systems Journal, vol. 5, no. 2, pp. 189-198, June 2011, https://doi.org/10.1109/JSYST.2010.2096670
6. A. Furno, N. El Faouzi, R. Sharma, V. Cammarota and E. Zimeo, A Graph-Based Framework for Real-Time Vulnerability Assessment of Road Networks, In: IEEE International Conference on Smart Computing (SMARTCOMP), 2018, pp. 234-241, https://doi.org/10.1109/SMARTCOMP.2018.00096.
7. Pavlopoulos, G.A., Secrier, M., Moschopoulos, C.N. et al. Using graph theory to analyze biological networks. BioData Mining 4, 10 (2011). https://doi.org/https://doi.org/10.1186/1756-0381-4-10
8. Psaltoglou, Artemis and Calle, Eusebi, Enhanced connectivity index A new measure for identifying critical points in urban public transportation networks. International Journal of Critical Infrastructure Protection. 2018. https://doi.org/10.1016/j.ijcip.2018.02.003
9. Corinna Köpke, Kushal Srivastava, Natalie Miller, Elena Branchini: Resilience Quantification for Critical Infrastructure: Exemplified for Airport Operations 2022
10. Bigdeli, A. et al. 2009, Comparison of Network Criticality, Algebraic Connectivity, and Other Graph Metrics. Proceedings of the 1st Annual Workshop on Simplifying Complex Network for Practitioners. 2009.
11. Köpke, Corinna, Srivastava, Kushal, König, Louis, Miller, Natalie, Fehling-Kaschek, Mirjam, Burke, Kelly, Mangini, Matteo, Praça, Isabel, Canito, Alda, Carvalho, Olga, Apolinário, Filipe, Escravana, Nelson, Carstengerdes, Nils, Stelkens-Kobsch, Tim: Impact Propagation in Airport Systems, In:Cyber-Physical Security

for Critical Infrastructures Protection, (pp.191-206), Guildford, UK, September 18, 2020 https://doi.org/10.1007/978-3-030-69781-5_13.
12. Fehling-Kaschek, Mirjam, Miller, Natalie, Haab, Gael, Faist, Katja, Stolz, Alexander, Häring, Ivo, Neri, Alberto, Celozzi, Giuseppe, Sanchez, Jose, Valera, Javier, Makri, Rodoula: Risk and Resilience Assessment and Improvement in the Telecommunication Industry. (pp. 247-254). In: (2020)https://doi.org/10.3850/978-981-14-8593-0_3995-cd.
13. Robert Layton and Paul A. Watters, T.: Automating Open Source Intelligence: Algorithms for OSINT, 2015
14. Scharte, B, D Hiller, T Leismann, and K Thoma.: Automating Open Source Intelligence: In Resilien Tech. Resilience by design: a strategy for the technology issues of the future (acatech STUDY), Munich:Herbert Utz Verlag, 2014.
15. OpenMobilityData, https://transitfeeds.com/p/roma-servizi-per-la-mobilita/542/latest, Last accessed 28 June 2022
16. Google Transit APIs, https://developers.google.com/transit/gtfs, Last accessed 28 June 2022
17. Safety4Rails, https://safety4rails.eu/
18. Resisto, https://www.resistoproject.eu/
19. Satie, https://satie-h2020.eu/
20. Numpy, https://numpy.org/
21. Networkx, https://networkx.org/
22. Shapely, https://shapely.readthedocs.io/en/stable/manual.html
23. GeoPandas, https://geopandas.org/en/stable/
24. Bokeh, http://docs.bokeh.org/en/latest/