# Kent Academic Repository <br> Full text document (pdf) 

## Citation for published version

Oliva, Gabriele and Esposito Amideo, Annunziata and Starita, Stefano and Setola, Roberto and Scaparra, Maria Paola (2020) Aggregating Centrality Rankings: A Novel Approach to Detect Critical Infrastructure Vulnerabilities. In: Critical Information Infrastructures Security. Lecture Notes in Computer Science, 11777 . Springer, pp. 57-68. ISBN 978-3-030-37669-7.

## DOI

https://doi.org/10.1007/978-3-030-37670-3_5

## Link to record in KAR

https://kar.kent.ac.uk/80071/

## Document Version

## UNSPECIFIED

## Copyright \& reuse

Content in the Kent Academic Repository is made available for research purposes. Unless otherwise stated all content is protected by copyright and in the absence of an open licence (eg Creative Commons), permissions for further reuse of content should be sought from the publisher, author or other copyright holder.

## Versions of research

The version in the Kent Academic Repository may differ from the final published version. Users are advised to check http://kar.kent.ac.uk for the status of the paper. Users should always cite the published version of record.

## Enquiries

For any further enquiries regarding the licence status of this document, please contact:
researchsupport@kent.ac.uk
If you believe this document infringes copyright then please contact the KAR admin team with the take-down information provided at http://kar.kent.ac.uk/contact.html

# Aggregating Centrality Rankings: A Novel Approach to Detect Critical Infrastructure Vulnerabilities 

Gabriele Oliva ${ }^{1(\boxtimes)}$ (D), Annunziata Esposito Amideo ${ }^{4}$ © ${ }^{(D)}$, Stefano Starita ${ }^{3}$ © Roberto Setola ${ }^{1}$ (D) and Maria Paola Scaparra ${ }^{2}$ (D)<br>${ }^{1}$ University Campus Biomedico of Rome, Rome, Italy g.oliva@unicampus.it<br>${ }^{2}$ Centre for Logistics and Heuristic Optimisation (CLHO), Kent Business School, University of Kent, Canterbury, England<br>${ }^{3}$ Sasin School of Management, Chulalongkorn University, Bangkok, Thailand<br>${ }^{4}$ University College Dublin, Dublin, Ireland


#### Abstract

Assessing critical infrastructure vulnerabilities is paramount to arrange efficient plans for their protection. Critical infrastructures are network-based systems hence, they are composed of nodes and edges. The literature shows that node criticality, which is the focus of this paper, can be addressed from different metric-based perspectives (e.g., degree, maximal flow, shortest path). However, each metric provides a specific insight while neglecting others. This paper attempts to overcome this pitfall through a methodology based on ranking aggregation. Specifically, we consider several numerical topological descriptors of the nodes' importance (e.g., degree, betweenness, closeness, etc.) and we convert such descriptors into ratio matrices; then, we extend the Analytic Hierarchy Process problem to the case of multiple ratio matrices and we resort to a Logarithmic Least Squares formulation to identify an aggregated metric that represents a good tradeoff among the different topological descriptors. The procedure is validated considering the Central London Tube network as a case study.


Keywords: Critical infrastructures • Criticality analysis • Ranking aggregation • Analytic Hierarchy Process • Least squares optimization

## 1 Introduction

Critical infrastructures are prone to disasters, both man-made and natural (e.g., see [1-3] in the case of railway infrastructures). Given the potential consequences of such disasters, it is mandatory to quantify and identify subsystems that are particularly critical, in that their disruption may cause severe consequences on the remaining subsystems. In this view, identifying such vulnerabilities is essential for deciding how to invest resources in order for instance to protect vulnerable subsystems. This is particularly relevant for critical infrastructure networks (e.g.,

[^0]power networks, railway networks, etc.), where the importance/criticality of a subsystem may not depend just on the physical characteristics of such subsystems, but also on the complex web of connections and relations that intertwine such composing elements [4,5]. Assessing critical infrastructure vulnerabilities is paramount to arrange efficient plans for their protection. Critical infrastructures are network-based systems hence, they are composed of nodes and edges. The literature shows that node criticality, which is the focus of this paper, can be addressed from different metric-based perspectives (e.g., degree, maximal flow, shortest path) [6-10]. However, each metric provides a specific insight while neglecting others. This paper attempts to overcome this pitfall through a methodology based on ranking aggregation. Specifically, in this paper we develop a methodology to aggregate topological descriptors based on the Analytic Hierarchy Process (AHP) [11]: first, we convert the numerical topological descriptors into ratio matrices and then we extend the Logarithmic Least Squares (LLS) AHP methodology [12-16] in order to find a least-squares optimal ranking that is a compromise among the considered ones. It should be noted that the problem of aggregating rankings has raised some interest in previous research: in [17] Kendall and Hausdorff distances are used to compare rankings and a medianbased approach is used to identify an overall ranking; in [18] interval ordinal rankings are considered; in [19] (and references therein) the bucket order problem is considered, i.e., finding an agreement based on several ranking matrices with ordinal information. Notice that, in [6], the authors quantify the correlation of centrality measures with risk levels in Dependency Risk Graphs and provide an heuristic algorithm to recursively select a subset of nodes based on the centrality measure with the highest correlation. In this paper we approach such a problem from a different perspective starting from the topological structure of the infrastructure and looking for those nodes that "optimize" a set of metrics which are not limited to the centrality ones. In this way, the aggregated ranking hereby proposed has a number of benefits: (i) being the result of a least-squares minimization problem, it represents the optimal tradeoff among the considered metrics; (ii) it provides a numerical characterization of the criticality of each node; (iii) it is not computationally expensive, as it consists in solving a system of $n$ linear equations with $n$ unknowns, where $n$ is the number of nodes in the network. The remainder of this paper is organized as follows: after some notation, which concludes this section, we present our aggregation methodology in Sect. 2; then, in Sect. 3 we validate the methodology with respect to a case study, namely, the Central London Tube network; finally, we provide some conclusive remarks and future work directions in Sect.4.

### 1.1 Notation

We denote vectors via boldface letters, while matrices are shown with uppercase letters. We use $A_{i j}$ to address the $(i, j)$-th entry of a matrix $A$ and $x_{i}$ for the $i$-th entry of a vector $\mathbf{x}$. Moreover, we write $\mathbf{1}_{n}$ and $\mathbf{0}_{n}$ to denote a vector with $n$ components, all equal to one and zero, respectively; similarly, we use $1_{n \times m}$ and $0_{n \times m}$ to denote $n \times m$ matrices all equal to one and zero, respectively. We
denote by $I_{n}$ the $n \times n$ identity matrix. We express by $\exp (\cdot)$ and $\ln (\cdot)$ the component-wise exponentiation or logarithm of a vector or matrix.

## 2 Aggregating Heterogeneous Rankings

In this section, we describe the methodology adopted to calculate an aggregated ranking that is representative of several rankings over the same set of alternatives.

### 2.1 The Approach in a Nutshell

Generally, different ranking criteria capture peculiar elements in terms of node criticality. Hence, any one of them provides a useful point of view to better understand the role and the relevance of each node. Consequently, selecting one ranking criterion while discarding another, may lead to misleading prioritizations in the protection strategies. To overcome such a limit we propose to aggregate the different ranking criteria into a single "super-ranking", i.e., an aggregated ranking that potentially collects all the different aspects of traditional metrics. In this view, our main idea is to convert the numerical rankings into square matrices containing the ratios of the importance of pairs of alternatives, and then combine them in a least square sense via the Logarithmic Least Squares Analytic Hierarchy Process (LLS-AHP) methodology [12-16], in order to obtain an aggregated ranking that is a good trade-off among the available ones. This approach has the advantage to allow a fair comparison among the criteria, in that the rankings are compared in terms of ratios of utilities and not in terms of actual utilities, which may have very different scales. Moreover, the least squares approach provides clear information on the degree of conflict among the rankings, in that the smaller the value of the objective function of the least squares problem is, the more data are in accordance, and vice versa.

### 2.2 Formal Definition of the Method

Let us consider a situation where we are given $m$ cardinal (i.e., numerical) rankings $\boldsymbol{r}^{(1)}, \ldots, \boldsymbol{r}^{(m)}$ over the set of $n$ nodes in a given graph. In particular, each ranking $\boldsymbol{r}^{(i)}$ is an $n \times 1$ vector having positive entries, and $r_{j}^{(i)}$ represents the numerical value or utility associated to the $j$-th node according to the $i$-th ranking. In order to obtain an aggregated ranking that is representative for the given $m$ rankings, our approach is composed of two logical steps: (1) converting the rankings into ratio matrices and (2) calculating the overall ranking. During the first step, we convert each ranking $\boldsymbol{r}^{(i)}$ into an $n \times n$ matrix $W^{(i)}$ such that the $(u, v)$-th entry $W_{u v}^{(i)}$ is in the form $W_{u v}^{(i)}=r_{u}^{(i)} / r_{v}^{(i)}$. In other words, $W_{u v}^{(i)} \bmod -$ els the relative utility or importance of the $u$-th alternative over the $j$-th one according to the $i$-th ranking. As a second step, we aim at finding the ranking vector $\boldsymbol{w}^{*}$ that solves the following problem.

Problem 1. Find $\boldsymbol{w}^{*} \in \mathbb{R}^{n}$ that solves

$$
\begin{equation*}
\underset{\boldsymbol{w} \in \mathbb{R}^{n}}{\arg \min } f(\boldsymbol{w})=\sum_{i=1}^{m} \sum_{u=1}^{n} \sum_{v=1}^{n}\left(\ln \left(W_{u v}^{(i)}\right)-\log \left(w_{u}\right)+\log \left(w_{v}\right)\right)^{2} \tag{1}
\end{equation*}
$$

subject to

$$
\left\{w_{u}>0, \quad \forall u \in\{1, \ldots, n\}\right.
$$

The above problem aims at finding the vector $\boldsymbol{w}^{*}$ such that the logarithm of the ratio of its components is the least squares compromise among the logarithms of the corresponding ratios $W_{u v}^{(i)}$. In other words, Problem 1 aims at finding the weight $w_{u}$, to be assigned to each node, such that the ratios $w_{u} / w_{v}$ minimize the deviation from respect to the ratios $W_{u v}^{(i)}$ for the m considered criteria. In order to solve this problem, which is in general non-convex and may have non-unique solution, we aim at finding a vector $\boldsymbol{y}^{*}$ such that $\boldsymbol{w}^{*}=\exp \left(\boldsymbol{y}^{*}\right)$, where $\exp (\cdot)$ is the component-wise exponential; in other words, we aim at solving the following unconstrained problem.
Problem 2. Find $\boldsymbol{y}^{*} \in \mathbb{R}^{n}$ that solves

$$
\begin{equation*}
\underset{\boldsymbol{y} \in \mathbb{R}^{n}}{\arg \min } g(\boldsymbol{y})=\sum_{i=1}^{m} \sum_{u=1}^{n} \sum_{v=1}^{n}\left(\ln \left(W_{u v}^{(i)}\right)-y_{u}+y_{v}\right)^{2} \tag{2}
\end{equation*}
$$

The above problem is easily solved in a closed form. Specifically, being an unconstrained convex problem, the minimum is attained at $\boldsymbol{y}^{*}$ such that, for all $u \in\{1, \ldots, n\}$, it holds $\left.\frac{\partial g(\boldsymbol{y})}{\partial y_{u}}\right|_{\boldsymbol{y}=\boldsymbol{y}^{*}}=0$. By some algebra, it can be shown that the optimal $\boldsymbol{y}^{*}$ satisfies

$$
m\left(n I_{n}-\mathbf{1}_{n} \mathbf{1}_{n}^{T}\right) \boldsymbol{y}^{*}=\sum_{i=1}^{m} \log \left(W^{(i)}\right) \mathbf{1}_{n}
$$

where $\log \left(W^{(i)}\right)$ is the $n \times n$ matrix collecting the logarithm of the corresponding entries of $W^{(i)}$ (note that we assumed the rankings have positive entries hence the logarithm is always finite). Note further that matrix $n I_{n}-\mathbf{1}_{n} \mathbf{1}_{n}^{T}$ is the Laplacian matrix of a complete graph and is singular [20]; hence, in order to find $\boldsymbol{y}^{*}$, one may need to resort to a pseudoinverse, i.e., by setting

$$
\boldsymbol{y}^{*}=\frac{1}{m}\left(n I_{n}-\mathbf{1}_{n} \mathbf{1}_{n}^{T}\right)^{\dagger} \sum_{i=1}^{m} \log \left(W^{(i)}\right) \mathbf{1}_{n}
$$

where $\left(n I_{n}-\mathbf{1}_{n} \mathbf{1}_{n}^{T}\right)^{\dagger}$ denotes the left pseudoinverse of $n I_{n}-\mathbf{1}_{n} \mathbf{1}_{n}^{T}$. An alternative approach is to solve in an approximated way via the differential equation

$$
\dot{\boldsymbol{y}}(t)=m\left(\mathbf{1}_{n} \mathbf{1}_{n}^{T}-n I_{n}\right) \boldsymbol{y}(t)+\sum_{i=1}^{m} \log \left(W^{(i)}\right) \mathbf{1}_{n}
$$

which is known to asymptotically converges to a vector that satisfies the above singular system of equations [21].

## 3 Case Study



Fig. 1. Central London tube map.

In this section, we consider as an example the Central London Tube network (Fig. 1). Specifically, we represent each station by a node (we consider 50 stations) and we model by directed edges (178 in total) the connections among neighboring stations; in particular, we associate to each edge a weight that corresponds to the average travel time (in seconds) between its endpoints. In other words, we consider a graph that is bidirectional (i.e., there is an edge from $i$ to $j$ whenever there is an edge from $j$ to $i$ ) and asymmetric (i.e., the weight associated to the edge from $i$ to $j$ is different from the weight of the edge from $j$ to $i$.) Fig. 2 reports the resulting asymmetric graph, where edges' color corresponds to the average travel time, according to the provided heatmap; notice that the association between the numerical identifier for each station and the corresponding name can be found in Table 1. With respect to the aforementioned graph, we consider some of the most popular centrality measures in the literature. Specifically, we consider (see [22] and references therein for details):

- In-degree: sum of the weights of the edges incoming at each node;
- Out-degree: sum of the weights of the edges outgoing at each node;
- Betweenness: measures how often a node belongs to the shortest paths between any pair of nodes. If the graph is weighted then path lengths depend on the weights. Specifically the betweenness is defined as $b_{u}=$ $\sum_{s, t \neq u} N_{s t}^{(u)} / N_{s t}$, where $N_{s t}^{(u)}$ is the amount of minimum paths between nodes $s$ and $t$ passing via node $u$ and $N_{s t}$ is the total number of minimum paths between nodes $s$ and $t$.


Fig. 2. Central London tube map as a bidirectional asymmetric weighted graph, where weights corresponds to the average travel time (in seconds) between neighboring stations.

- Pagerank: it is a measure of importance of the nodes that results from a random walk on the network. Specifically, the random walk is performed with probabilities that depend on the edges' weights. If at some point a node has no outgoing edges, a new random node is chosen. The pagerank measure is the average time spent at each node during the walk.
- Hubs \& Authorities: such metrics are defined together in a recursive way. The 'hubs-score' of a node is the sum of the 'authorities-score' of its neighbors, and vice-versa. Such values can be regarded as the left (hubs) and right (authorities) singular vectors that correspond to the largest singular value of the adjacency matrix of the graph.
- Closeness: this metric is based on the inverse sum of the distances from a node to all other nodes in the graph. Specifically, the closeness is defined as $c_{u}=A_{u} /\left(C_{u}(n-1)\right)$, where $A_{u}$ is the number of reachable nodes from node $u$ (not counting $u$ ), $n$ is the number of nodes in the graph, and $C_{u}$ is the sum of the distances from node $u$ to all reachable nodes (if the node is isolated then $c_{u}=0$ ).
- Eigenvector Centrality: this metric uses the eigenvector corresponding to the largest eigenvalue of the graph adjacency matrix. The scores are normalized such that the sum of all values is equal to 1 .

Overall, we obtain $m=8$ (numerical) ranking vectors $\boldsymbol{r}^{(i)}$. In Table 1, we report the numerical data for each topological descriptor and for the proposed aggregated metric, while in Table 2 we report the ranking of the stations based, again, on the topological descriptors and on the proposed aggregated metric. In order to provide an immediate understanding of the above data, we show in Fig. 3 the criticality of each node in the network based on the different metrics via a red-blue heat-map, i.e., the more the color of the nodes is red the more the value of the corresponding metric is closer to the maximum value. According to the figure, the different topological indicators identify very different nodes as the most important, and that the proposed aggregated metric represents, indeed, a compromise among the original metrics.


Fig. 3. Visual representation of the nodes' criticality according to the different topological descriptors and to the proposed aggregated measure. (Color figure online)


Fig. 4. Kendall's correlation between the ranking obtained based on the proposed aggregated metric and the rankings obtained according to the considered topological descriptors, considering all stations (Fig. 4a) and considering only the 20 most important stations according to the aggregated metric (Fig. 4b).

Table 1. Nodes featured in the case study with the numerical values of the considered topological descriptors and of the proposed aggregated centrality.

| Id | Name | In-degree | Out-degree | Betweeness | Pagerank | Hubs | Authorities | Closeness | Eigenvector | Aggregated |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | Angel | 461 | 502 | 73.33 | 0.021 | 0.018 | 0.004 | 0.004 | 0.003 | 0.013 |
| 2 | Baker Street | 880 | 872 | 455.6 | 0.037 | 0.023 | 0.026 | 0.005 | 0.037 | 0.042 |
| 3 | BankMonument | 1089 | 1003 | 468.3 | 0.046 | 0.035 | 0.225 | 0.005 | 0.010 | 0.052 |
| 4 | Barbican | 260 | 252 | 54 | 0.013 | 0.003 | 0.020 | 0.004 | 0.002 | 0.009 |
| 5 | Bayswater | 342 | 310 | 66.16 | 0.018 | 0.001 | 0.001 | 0.003 | 0.001 | 0.006 |
| 6 | Blackfriars | 262 | 289 | 39.5 | 0.016 | 0.003 | 0.004 | 0.004 | 0.003 | 0.008 |
| 7 | Bond Street | 595 | 618 | 566.21 | 0.025 | 0.033 | 0.040 | 0.006 | 0.064 | 0.044 |
| 8 | Cannon Street | 186 | 190 | 79.16 | 0.012 | 0.033 | 0.006 | 0.004 | 0.003 | 0.011 |
| 9 | Chancery Lane | 239 | 237 | 87.86 | 0.013 | 0.003 | 0.016 | 0.005 | 0.007 | 0.012 |
| 10 | Charing Cross | 342 | 357 | 88.73 | 0.016 | 0.017 | 0.014 | 0.005 | 0.042 | 0.021 |
| 11 | Covent Garden | 230 | 217 | 22.73 | 0.012 | 0.003 | 0.003 | 0.005 | 0.019 | 0.009 |
| 12 | Edgware Road | 529 | 467 | 196.83 | 0.025 | 0.008 | 0.011 | 0.005 | 0.015 | 0.020 |
| 13 | Embankment | 500 | 501 | 236.90 | 0.022 | 0.030 | 0.063 | 0.005 | 0.032 | 0.035 |
| 14 | Euston | 264 | 267 | 120.56 | 0.013 | 0.006 | 0.005 | 0.005 | 0.011 | 0.013 |
| 15 | Euston Square | 290 | 305 | 140.90 | 0.014 | 0.007 | 0.004 | 0.005 | 0.005 | 0.018 |
| 16 | Farringdon | 418 | 434 | 73.33 | 0.019 | 0.013 | 0.003 | 0.004 | 0.003 | 0.012 |
| 17 | Gloucester Road | 317 | 351 | 90.66 | 0.016 | 0.001 | 0.001 | 0.003 | 0.002 | 0.007 |
| 18 | Goodge Street | 235 | 239 | 7.13 | 0.012 | 0.004 | 0.006 | 0.005 | 0.023 | 0.009 |
| 19 | Great Portland Street | 328 | 348 | 162.00 | 0.015 | 0.009 | 0.010 | 0.005 | 0.011 | 0.017 |
| 20 | Green Park | 1024 | 1037 | 819.89 | 0.039 | 0.071 | 0.052 | 0.007 | 0.090 | 0.069 |
|  | High Street Kensington | 381 | 348 | 51.33 | 0.019 | 0.000 | 0.001 | 0.003 | 0.001 | 0.004 |
| 22 | Holborn | 525 | 533 | 244.52 | 0.025 | 0.006 | 0.006 | 0.005 | 0.023 | 0.021 |
| 23 | Hyde Park Corner | 283 | 285 | 158.83 | 0.013 | 0.014 | 0.023 | 0.005 | 0.026 | 0.021 |
| 24 | St Pancras | 1027 | 980 | 322.40 | 0.042 | 0.005 | 0.020 | 0.005 | 0.008 | 0.027 |
| 25 | Knightsbridge | 360 | 333 | 89.50 | 0.016 | 0.006 | 0.005 | 0.004 | 0.009 | 0.013 |
| 26 | Lancaster Gate | 297 | 357 | 141.16 | 0.017 | 0.002 | 0.002 | 0.004 | 0.005 | 0.010 |
| 27 | Leicester Square | 409 | 407 | 101.64 | 0.019 | 0.010 | 0.013 | 0.005 | 0.050 | 0.021 |
| 28 | London Bridge | 155 | 153 | 0.00 | 0.009 | 0.051 | 0.010 | 0.004 | 0.003 | 0.000 |
| 29 | Mansion House | 245 | 220 | 26.66 | 0.015 | 0.002 | 0.006 | 0.004 | 0.002 | 0.007 |
| 30 | Marble Arch | 294 | 262 | 213.16 | 0.014 | 0.006 | 0.007 | 0.005 | 0.019 | 0.016 |
| 31 | Marylebone | 220 | 225 | 0.00 | 0.012 | 0.006 | 0.006 | 0.004 | 0.014 | 0.000 |
| 32 | Moorgate | 446 | 473 | 178.33 | 0.022 | 0.073 | 0.013 | 0.004 | 0.004 | 0.022 |
| 33 | Notting Hill Gate | 433 | 452 | 87.00 | 0.024 | 0.001 | 0.000 | 0.003 | 0.001 | 0.006 |
| 34 | Old Street | 432 | 380 | 54.00 | 0.019 | 0.004 | 0.028 | 0.004 | 0.002 | 0.012 |
| 35 | Oxford Circus | 968 | 899 | 529.20 | 0.037 | 0.039 | 0.056 | 0.006 | 0.093 | 0.058 |
| 36 | Paddington | 371 | 398 | 127.50 | 0.021 | 0.004 | 0.003 | 0.004 | 0.004 | 0.011 |
| 37 | Piccadilly Circus | 572 | 572 | 128.55 | 0.024 | 0.031 | 0.034 | 0.006 | 0.074 | 0.035 |
| 38 | Queensway | 282 | 253 | 76.16 | 0.015 | 0.000 | 0.001 | 0.004 | 0.002 | 0.005 |
| 39 | Regents Park | 340 | 350 | 58.72 | 0.015 | 0.022 | 0.019 | 0.005 | 0.035 | 0.021 |
| 40 | Russell Square | 322 | 327 | 80.00 | 0.015 | 0.007 | 0.003 | 0.005 | 0.008 | 0.012 |
| 41 | Sloane Square | 373 | 380 | 124.50 | 0.016 | 0.011 | 0.008 | 0.004 | 0.010 | 0.016 |
| 42 | South Kensington | 551 | 583 | 158.66 | 0.024 | 0.004 | 0.006 | 0.004 | 0.005 | 0.015 |
| 43 | St James Park | 294 | 283 | 51.66 | 0.013 | 0.025 | 0.015 | 0.005 | 0.021 | 0.018 |
| 44 | St Pauls | 315 | 314 | 75.00 | 0.017 | 0.057 | 0.011 | 0.004 | 0.005 | 0.017 |
| 45 | Temple | 274 | 283 | 104.83 | 0.015 | 0.014 | 0.008 | 0.004 | 0.009 | 0.015 |
| 46 | Tottenham Court Road | 471 | 491 | 245.21 | 0.022 | 0.013 | 0.011 | 0.006 | 0.052 | 0.026 |
| 47 | Victoria | 501 | 512 | 219.16 | 0.022 | 0.020 | 0.033 | 0.005 | 0.032 | 0.030 |
| 48 | Warren Street | 392 | 411 | 186.70 | 0.018 | 0.016 | 0.013 | 0.005 | 0.034 | 0.024 |
| 49 | Waterloo | 695 | 770 | 359.63 | 0.029 | 0.192 | 0.043 | 0.005 | 0.023 | 0.049 |
| 50 | Westminster | 620 | 609 | 508.06 | 0.025 | 0.039 | 0.086 | 0.006 | 0.044 | 0.047 |

Table 2. Rankings of the stations according to the considered topological descriptors, and according to the proposed aggregated centrality.

| In-degree | Out-degree | Betweeness | Pagerank | Hubs | Authorities | Closeness | Eigenvector | Aggregated |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| BankMonument | Green Park | Green Park | BankMonument | Waterloo | BankMonument | Green Park | Oxford Circus | Green Park |
| St Pancras | BankMonument | Bond Street | St Pancras | Moorgate | Westminster | Oxford Circus | Green Park | Oxford Circus |
| Green Park | St Pancras | Oxford Circus | Green Park | Green Park | Embankment | Bond Street | Piccadilly Circus | BankMonument |
| Oxford Circus | Oxford Circus | Westminster | Oxford Circus | St Pauls | Oxford Circus | Westminster | Bond Street | Waterloo |
| Baker Street | Baker Street | BankMonument | Baker Street | London Bridge | Green Park | Piccadilly Circus | Tottenham Court Road | Westminster |
| Waterloo | Waterloo | Baker Street | Waterloo | Westminster | Waterloo | Tottenham Court Road | Leicester Square | Bond Street |
| Westminster | Bond Street | Waterloo | Holborn | Oxford Circus | Bond Street | Baker Street | Westminster | Baker Street |
| Bond Street | Westminster | St Pancras | Edgware Road | BankMonument | Piccadilly Circus | Waterloo | Charing Cross | Piccadilly Circus |
| Piccadilly Circus | South Kensington | Tottenham Court Road | Bond Street | nnon Street | Victoria | Victoria | Baker Street | Embankment |
| South Kensington | Piccadilly Circus | Holborn | Westminster | Bond Street | Old Street | Warren Street | Regents Park | Victoria |
| Edgware Road | Holborn | Embankment | South Kensington | Piccadilly Circ | Baker Street | Embankment | Warren Street | St Pancras |
| Holborn | Victoria | Victoria | Notting Hill Gate | Embankment | yde Park Corn | Hyde Park Corner | Victoria | Tottenham Court Road |
| Victoria | Angel | Marble Arch | Piccadilly Circus | St James Park | Barbican | Regents Park | Embankment | Warren Street |
| Embankment | mbankment | Edgware Road | Tottenham Court Road | Baker Street | St Pancras | Charing Cross | Hyde Park Corner | Moorgate |
| Tottenham Court | enham Court | Warren Street | Moorgate | Regents Park | Regents Park | Holborn | Holborn | Leicester Square |
| Angel | Moorgate | Moorgate | Embankment | Victoria | Chancery Lane | Leicester Square | Waterloo | Charing Cross |
| Moorgate | Edgware Road | Great Portland Street | Angel | Angel | St James Park | BankMonument | Goodge Street | Regents Park |
| Notting Hill Gate | Notting Hill Gate | Hyde Park Corner | Victoria | Charing Cross | Charing Cross | Marble Arch | St James Park | Holborn |
| Old Street | Farringdon | South Kensington | Paddington | Warren Street | Moorgate | St James Park | Covent Garden | Hyde Park Corner |
| Farringdon | Warren Street | Lancaster Gate | Farringdon | Hyde Park Corner | Warren Street | St Pancras | Marble Arch | Edgware Road |
| Leicester Square | Leicester Square | Euston Square | High Street Kensington | Temple | Leicester Square | Euston | Edgware Road | St James Park |
| Warren Street | Paddington | Piccadilly Circus | Old Street | Tottenham Court Road | St Pauls | Great Portland Street | Marylebone | St Pauls |
| High Street Kensington | Old Street | Paddington | Leicester Square | Farringdon | Tottenham Court Road | Goodge Street | Great Portland Street | Great Portland Street |
| Sloane Square | Sloane Square | Sloane Square | Bayswater | Sloane Square | Edgware Road | Russell Square | Euston | Sloane Square |
| Paddington | Charing Cross | Euston | Warren Street | Leicester Square | Great Portland Street | Covent Garden | Sloane Square | Marble Arch |
| Knightsbridge | Lancaster | Temple | Lancaster Ga | Great Portland Street | London Bridge | Chancery Lane | BankMonument | South Kensington |
| Bayswater | Gloucester Ro | Leicester Squ | St Pauls | Edgware Road | Temple | Edgware Road | Temple | Temple |
| Charing Cross | Regents Park | Gloucester R | Blackf | uston Square | Sloane Squa | Euston Square | Knightsbridge | Angel |
| Regents Park | High Street Kensington | Knightsbridge | Knightsbridge | Russell Squ | Marble Arch | Sloane Square | Russell Square | Euston |
| Great Portland Street | Great Portland Street | Charing Cross | Gloucester Road | Euston | Holborn | Knightsbridge | St Pancras | Knightsbridge |
| Russell Square | Knightsbridge | Chancery Lane | Sloane Square | Marble Arch | South Kensington | Marylebone | Chancery Lane | Farringdon |
| Gloucester Road | Russell Square | Notting Hill Gate | Charing Cross | Marylebone | Mansion House | St Pauls | Lancaster Gate | Chancery Lane |
| St Pauls | St Pauls | Russell Square | Mansion House | Knightsbridge | Marylebone | Moorgate | South Kensington | Old Street |
| Lancaster Gate | Bayswater | Cannon Street | Russell Square | Holborn | Goodge Street | Temple | Euston Square | Russell Square |
| Marble Arch | Euston Square | Queensway | Queensway | St Pancras | Cannon Street | Angel | St Pauls | Euston Square |
| St James Park | Blackfriars | St Pauls | Regents Park | South Kensington | Euston | Farringdon | Paddington | Cannon Street |
| Euston Square | Hyde Park Corner | Angel | Great Portland Street | Goodge Street | Knightsbridge | Lancaster Gate | Moorgate | Paddington |
| Hyde Park Corner | St James Park | Farringdon | Temple | Paddington | Blackfriars | Cannon Street | Cannon Street | Lancaster Gate |
| Queensway | Temple | Bayswater | Marble Arch | Old Street | Angel | Old Street | Blackfriars | Barbican |
| Temple | Euston | Regents Park | Euston Square | Chancery Lane | Euston Square | Barbican | London Bridge | Goodge Street |
| Euston | Marble Arch | Old Street | Chancery Lane | Barbican | Farringdon | London Bridge | Angel | Covent Garden |
| Blackfriars | Queensway | Barbican | St James Park | Blackfriars | Russell Square | South Kensington | Farringdon | Blackfriars |
| Barbican | Barbican | St James Park | Hyde Park Corner | Covent Garden | Paddington | Paddington | Queensway | Gloucester Road |
| Mansion House | Goodge Street | High Street Kensington | Barbican | Lancaster Gate | Covent Garden | Blackfriars | Old Street | Mansion House |
| Chancery Lane | Chancery Lane | Blackfriars | Euston | Mansion House | Lancaster Gate | Queensway | Barbican | Bayswater |
| Goodge Street | Marylebone | Mansion House | Cannon Street | Gloucester Road | Bayswater | Mansion House | Gloucester Road | Notting Hill Gate |
| Covent Garden | Mansion House | Covent Garden | Covent Garden | Bayswater | Gloucester Road | Gloucester Road | Mansion House | Queensway |
| Marylebone | Covent Garden | Goodge Street | Marylebone | Notting Hill Gate | High Street Kensington | Bayswater | Bayswater | High Street Kensington |
| Cannon Street | Cannon Street | Marylebone | Goodge Street | Queensway | Queensway | Notting Hill Gate | Notting Hill Gate | Marylebone |
| London Bridge | London Bridge | London Bridge | London Bridge | High Street Kensington | Notting Hill Gate | High Street Kensington | High Street Kensington | London Bridge |

In order to validate the above intuition, we calculate the Kendall's correlation coefficient ${ }^{1}$ between the ranking obtained based on the proposed aggregated metric and the rankings obtained according to the considered topological descriptors, as shown in Fig. 4; specifically, we show in Fig. 4a the correlation over the entire set of nodes, while Figs. 4b displays the correlations obtained considering the 20 most important nodes according to the aggregated metric. As shown by the figures, it can be noted that the correlations obtained over the whole set of nodes are all less than 0.1 in magnitude, while limiting to a subset of the 20 most important nodes the correlations with most metrics further reduce, except for the eigenvector centrality, which reaches a correlation of 0.2 . Overall, the above results suggest that the proposed index, by aggregating different metrics, assigns a criticality to the nodes that can not be exhaustively explained by any of the original metrics. In fact, by looking at Fig. 3, it can be noted that the most influential nodes according to the proposed aggregated metric are indeed represented by the union of the most influent nodes according to all the different topological descriptors (although we observe that the high importance assigned to some peripheral nodes based on the closeness, in-degree and out-degree criteria is reduced in the aggregated metric.

## 4 Conclusions and Future Work

In this paper we provide a novel methodology to aggregate heterogeneous criticality indices for critical infrastructure networks in order to obtain an overall aggregated ranking that represents a good trade-off among the different metrics. Such an index can be the basis for implementing protection strategies that are not driven by a single factor but consider at the same time multiple facets of node criticality. The main idea is to convert the metrics in ratio matrices and then compute an aggregated metric by means of a generalization of the Logarithmic Least Squares Analytic Hierarchy Process technique to the case of multiple ratio matrices. The experimental results show that the proposed approach assigns

[^1]large relevance to the most influential nodes according to the single indices; yet, the resulting criticality cannot be exhaustively explained by any of the original metrics thus requiring further investigation. Future work will follow three main directions: (i) we will consider different graphs over the same set of nodes (e.g., structural graph, flow graph,. etc.) in order to take into account, at the same time, both structural and dynamical characteristics of the network; (ii) we will extend the framework by implementing a multi-criteria decision procedure to weight differently the different topological descriptors, in order to obtain a synthetic metric that reflects the preferences of stakeholders or decision-makers; (iii) we will inspect the possibility to prioritize ordinal information over cardinal information, extending the framework in [24] to the case of multiple ratio matrices.

Acknowledgements. This work was supported by INAIL via the European Saf€ra project "Integrated Management of Safety and Security Synergies in Seveso Plants" (4STER).

## References

1. Anderson, C.W., Santos, J.R., Haimes, Y.Y.: A risk-based input-output methodology for measuring the effects of the August 2003 northeast blackout. Econ. Syst. Res. 19(2), 183-204 (2007)
2. Popova, O.P., et al.: Chelyabinsk airburst, damage assessment, meteorite recovery, and characterization. Science 342(6162), 1069-1073 (2013)
3. Setola, R., Sforza, A., Vittorini, V., Pragliola, C. (eds.): Railway Infrastructure Security. TSRRQ, vol. 27. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-04426-2
4. Setola, R.: How to measure the degree of interdependencies among critical infrastructures. Int. J. Syst. Syst. Eng. 2(1), 38-59 (2010)
5. Faramondi, L., et al.: Network structural vulnerability: a multiobjective attacker perspective. IEEE Trans. Syst. Man Cybern. Syst. (99), 1-14 (2018)
6. Stergiopoulos, G., Kotzanikolaou, P., Theocharidou, M., Gritzalis, D.: Risk mitigation strategies for critical infrastructures based on graph centrality analysis. Int. J. Crit. Infrastruct. Prot. 10, 34-44 (2015)
7. Chen, X.: Critical nodes identification in complex systems. Complex Intell. Syst. 1(1-4), 37-56 (2015)
8. Lü, L., Chen, D., Ren, X.-L., Zhang, Q.-M., Zhang, Y.-C., Zhou, T.: Vital nodes identification in complex networks. Phys. Rep. 650, 1-63 (2016)
9. Rueda, D.F., Calle, F., Marzo, J.L.: Robustness comparison of 15 real telecommunication networks: structural and centrality measurements. J. Netw. Syst. Manag. 25(2), 269-289 (2017)
10. Starita, S., Esposito Amideo, A., Scaparra, M.P.: Assessing urban rail transit systems vulnerability: metrics vs. interdiction models. In: D'Agostino, G., Scala, A. (eds.) CRITIS 2017. LNCS, vol. 10707, pp. 144-155. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-99843-5_13
11. Saaty, T.L.: A scaling method for priorities in hierarchical structures. J. Math. Psychol. 15(3), 234-281 (1977)
12. Crawford, G.B.: The geometric mean procedure for estimating the scale of a judgement matrix. Math. Model. 9(3-5), 327-334 (1987)
13. Barzilai, J., Cook, W.D., Golany, B.: Consistent weights for judgements matrices of the relative importance of alternatives. Oper. Res. Lett. 6(3), 131-134 (1987)
14. Bozóki, S., Fülöp, J., Rónyai, L.: On optimal completion of incomplete pairwise comparison matrices. Math. Comput. Model. 52(1-2), 318-333 (2010)
15. Oliva, G., Setola, R., Scala, A.: Sparse and distributed analytic hierarchy process. Automatica 85, 211-220 (2017)
16. Bozóki, S., Tsyganok, V.: The (logarithmic) least squares optimality of the arithmetic (geometric) mean of weight vectors calculated from all spanning trees for incomplete additive (multiplicative) pairwise comparison matrices. Int. J. Gen. Syst. (2019, to appear)
17. Fagin, R., Kumar, R., Mahdian, M., Sivakumar, D., Vee, E.: Comparing and aggregating rankings with ties. In: Proceedings of the Twenty-Third ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems, pp. 47-58. ACM (2004)
18. Dopazo, E., Martínez-Céspedes, M.L.: Rank aggregation methods dealing with ordinal uncertain preferences. Expert Syst. Appl. 78, 103-109 (2017)
19. Aledo, J.A., Gámez, J.A., Rosete, A.: Utopia in the solution of the bucket order problem. Decis. Support Syst. 97, 69-80 (2017)
20. Godsil, C., Royle, G.: Algebraic Graph Theory. Graduate Text in Mathematics. Springer, New York (2001)
21. Olfati-Saber, R., Alex Fax, J., Murray, R.M.: Consensus and cooperation in networked multi-agent systems. Proc. IEEE 95(1), 215-233 (2007)
22. Borodin, A., Roberts, G.O., Rosenthal, J.S., Tsaparas, P.: Link analysis ranking: algorithms, theory, and experiments. ACM Trans. Internet Technol. (TOIT) 5(1), 231-297 (2005)
23. Kendall, M.G.: A new measure of rank correlation. Biometrika 30(1/2), 81-93 (1938)
24. Oliva, G., Bozóki, S.: Incomplete analytic hierarchy process with minimum ordinal violations. arXiv preprint: arXiv:1904.04701 (2019)

[^0]:    (C) Springer Nature Switzerland AG 2020
    S. Nadjm-Tehrani (Ed.): CRITIS 2019, LNCS 11777, pp. 57-68, 2020.
    https://doi.org/10.1007/978-3-030-37670-3_5

[^1]:    ${ }^{1}$ Given two pairs of values $\left(a_{i}, b_{i}\right)$ and ( $a_{j}, b_{j}$ ), we say they are concordant if both $a_{i}>a_{j}$ and $b_{i}>b_{j}$ or if both $a_{i}<a_{j}$ and $b_{i}<b_{j}$; similarly the pairs are discordant if $a_{i}>a_{j}$ and $b_{i}<b_{j}$ or if $a_{i}<a_{j}$ and $b_{i}>b_{j}$. If $a_{i}=a_{j}$ or $b_{i}=b_{j}$ the pairs are neither concordant nor discordant. Given two vectors $\mathbf{a} \in \mathbb{R}^{n}$ and $\mathbf{b} \in \mathbb{R}^{n}$, the Kendall's correlation index [23] $\tau$ is defined as

    $$
    \tau=\frac{\mathcal{C}-\mathcal{P}}{n(n-1) / 2}
    $$

    where $\mathcal{C}$ and $\mathcal{P}$ are the set of concordant and discordant pairs $\left(a_{i}, b_{i}\right)$ and $\left(a_{j}, b_{j}\right)$, respectively. When $\mathbf{b}$ is a permutation of the components of $\mathbf{a}$, the Kendall's tau can be interpreted as a measure of the degree of shuffling of $\mathbf{b}$ with respect to $\mathbf{a}$, between minus one and one. In this sense $\tau=1$ implies $\mathbf{a}=\mathbf{b}$, while $\tau=-1$ represents the fact $\mathbf{b}$ is in reverse order with respect to $\mathbf{a}$. The closer is $\tau$ to (minus) one, therefore, the more the two rankings are (anti-) correlated, while the closer is $\tau$ to zero the more the two rankings are independent.

