Centralities of Nodes and Influences of Layers in Large Multiplex Networks

Christoph Rahmede
Rome International Centre for Material Science,
Superstripes RICMASS 00185 Roma, Italy

Jacopo Iacovacci
School of Mathematical Sciences,
Queen Mary University of London, E1 4NS, London, UK

Alex Arenas
Departament d'Enginyeria Informática i Matemátiques,
Universitat Rovira i Virgili, 43007 Tarragona, Spain

Ginestra Bianconi
School of Mathematical Sciences,
Queen Mary University of London, E1 4NS, London, UK

September 18, 2017

We formulate and propose an algorithm (MultiRank) for the weighted ranking of nodes and layers in large multiplex networks. MultiRank takes into account the full multiplex network structure of the data and exploits the dual nature of the network in terms of nodes and layers. The proposed centrality of the layers (influences) and the centrality of the nodes are determined by a coupled set of equations. The basic idea consist on assigning more centrality to nodes that receive links from highly influential layers and from already central nodes. The layers are more influential if highly central nodes are active in them. The algorithm applies to directed/undirected as well as to weighted/unweighted multiplex networks. We discuss the application of MultiRank to three major examples of multiplex network datasets: the European Air Transportation Multiplex Network, the Pierre Auger Multiplex Collaboration Network and the FAO Multiplex Trade Network.

1   INTRODUCTION

Our world is increasingly dependent on efficient ranking algorithms [1–6]. Currently ranking algorithms based on PageRank [1] constitute by far the most diffuse way to navigate through online-stored information. As such they are perhaps one of the most significant revolutions in the field of knowledge exploration since the first publication of the Diderot Encyclopedia.

Currently the ranking of nodes in complex networks is used in a variety of different contexts [3–6], from finance to social and biological networks. In the context of economical trade networks, formed by bipartite networks of countries and products, ranking algorithms [7–9] are recognized as an important tool to evaluate the economic development of countries. Interestingly the algorithms proposed for ranking nodes and products in bipartite trade networks have been shown to be also quite useful for other applications such as ranking species in bipartite ecological networks [10].

However it has been recently recognized [11–17] that most complex systems are not simply formed by a simple or a bipartite network, but they are instead formed by multilayer networks. Multilayer networks include not just one but several layers (networks) characterizing interactions of different nature and connotation.

Multiplex networks [18–26] are a special type of multilayer networks. They are formed by a set of N nodes connected by M different types of interactions. Each set of interactions of the same type determines a distinct layer (network) of the multiplex. Examples of multiplex networks are ubiquitous. For instance trade networks [27–30] between countries are inherently multiplex structures formed by N nodes (countries) trading with each other M different kinds of products (layers). Other major examples of multiplex networks range from transportation networks [24, 31] to social [20, 22, 32–34], financial [35, 36] and biological networks [16, 30, 37, 38]. In transportation networks [24, 31] different layers can represent different means of transportation while in scientific collaboration networks the different layers can represent different topics of the collaboration [22, 34]. Finally, in biology, multiplex networks span from the multiplex connectome of C. elegans to multiplex molecular networks in the cell [16, 30, 37–39].

Given the surge of interest in multiplex networks, recently several algorithms [40–46] have been proposed to assess the centrality of nodes in these multilayer structures. In Ref. [40, 41] the proposed algorithm captures how the centrality of the nodes in a given layer (master layer) of the multiplex can affect the centrality of the nodes in other layers. This effect is modelled by considering a PageRank algorithm biased on the centrality of the nodes in the master layer. In Ref. [42] it has been proposed instead to rank simultaneously nodes and layers of the multiplex network based on a random walk process [23, 24] (defined over a 2-covariant and 2-contravariant tensor representation of the multiplex network [17]) that hops between nodes of the same layer and between nodes of different layers as well. The resulting centrality called "versatility" strongly awards nodes active (connected) in many layers. Moreover in Ref. [45] the communicability has been proposed as a centrality measure based on the flow of information in multiplex networks.

Here a different approach is considered, where we consider a random walk hopping through links of different layers with different probabilities determined by the centrality of the layers (*influences*). The role of the influences of the layers has been first introduced in Ref. [43] where different measures for the centrality of the nodes given a set of influences of the layers have been proposed. Knowing the correct values to attribute to the influences of the layers might actually be a non-trivial task as in most of the cases the influences are not known a priori. In order to address this issue in Ref. [44] we proposed the Functional Multiplex PageRank that allows to take a comprehensive approach and to establish the centrality of the nodes for a wide range of influence values. Specifically the Functional Multiplex PageRank attributes not only a single number (rank) to a node but an entire function determining the rank of the node as a function of the influences attributed to its different types of links.

Although the Functional Multiplex PageRank [44] has been show to be very informative for mul-

tiplex networks with few layers, its applications to multiplex networks with many layers is somewhat limited.

Here in order to address this question we propose a ranking algorithm, MultiRank, that is specified by a coupled set of equations that simultaneously determine the centrality of the nodes and the influences of the layers of a multiplex network. The MultiRank algorithm applies to any type of multiplex network including weighted and directed multiplex network structures.

Interestingly the MultiRank can be related to the literature of ranking algorithms for bipartite networks [7–9]. In fact, while processing all the information present in the multiplex network structure, it also exploits the bipartite nature [21] of the network between nodes and layers that can be extracted directly from the multiplex. This bipartite network indicates which node are connected in which layer and which is the strength of their incoming and outgoing links.

The MultiRank algorithm is based on a random walk among the nodes of the multiplex network. Starting from a node the random walker either hops to a neighbor node with probability depending on the *influence* of the layer or jumps to a random node of the network that is not isolated. The influences are the layer centralities that are determined by another set of equations, depending on the centrality of the nodes that are active (i.e. connected) in a given layer.

MultiRank efficiently ranks nodes and layers of large multiplex networks including many layers. Here we discuss the application of the algorithm to three major examples of multiplex networks: the European Air Transportation Multiplex Network [31], the Pierre Auger Multiplex Collaboration Network [34] and the FAO Multiplex Trade Network [30].

2   THE MULTIRANK ALGORITHM

2.1   MultiRank: the motivation

Let us consider a multiplex network of M layers $\alpha = 1, 2, \ldots, M$ and N nodes $i = 1, 2, 3, \ldots, N$. We indicate with $\mathbf{A}^{[\alpha]}$ the adjacency matrix (eventually directed and weighted) of layer α with elements $A_{ij}^{[\alpha]} > 0$ if node i points to node j with weight $A_{ij} > 0$. The aggregated network is the single network in which each pair of nodes is connected if they have at least a link in one layer of the multiplex network.

Different layers might have more or less connections and/or more or less total weight of their links. Let us characterize the potential heterogeneity of the layers by attributing to each layer α the quantity $W^{[\alpha]}$ defined as

$$W^{[\alpha]} = \sum_{i=1}^{N} \sum_{j=1}^{N} A_{ij}^{[\alpha]}. \tag{1}$$

For an undirected unweighted multiplex network $W^{[\alpha]}$ indicates the double of the total number of links in layer α. For a directed unweighted multiplex network it indicates the total number of links in layer α. For a weighted multiplex network it either indicates the total weight of the links in layer α (directed case) or the double of this number (undirected case).

Our aim is to introduce a framework to determine simultaneously the centrality X_i of the nodes $i = 1, 2, 3, \ldots, N$ and the centrality $\mathbf{z}^{[\alpha]}$ of the layers (influences) $\alpha = 1, 2, \ldots M$. On one side we will interpret the multiplex network as a colored network where different types of links are assigned a different influence. On the other side we will exploit the weighted, directed bipartite network formed by nodes and layers [11, 21] that can be constructed from the multiplex network. This bipartite network will provide us with information about the activity of the nodes [22] in each layer, i.e. if they are present (and therefore connected) in the layer. Additionally the weighted and directed links of this bipartite network will indicate which is the in-strength and out-strength of each node in any given

layer. In Figure 1 we display graphically how the bipartite network between nodes and layers can be constructed starting from any multiplex network data.

Specifically we will work with two matrices extracted from the multiplex network. The first matrix is the $N \times N$ weighted matrix $\mathbf{G}$ of the colored network, where the links of each layer $\alpha = 1, 2, \ldots, M$ are weighted with the influences $z^{[\alpha]}$ associated to it. Therefore the elements G_{ij} of the matrix $\mathbf{G}$ are given by

$$G_{ij} = \sum_{\alpha=1}^{M} A_{ij}^{[\alpha]} z^{[\alpha]}. \tag{2}$$

Secondly we extract the incidence matrices of the bipartite network constructed from the multiplex network by considering the connectivity of each node i in layer α. For directed multiplex networks we need to distinguish two $M \times N$ incidence matrices $\mathbf{B}^{in}$ and $\mathbf{B}^{out}$ of elements

$$B_{\alpha i}^{in} = \frac{\sum_j A_{ji}^{[\alpha]}}{W^{[\alpha]}},$$
$$B_{\alpha i}^{out} = \frac{\sum_j A_{ij}^{[\alpha]}}{W^{[\alpha]}}. \tag{3}$$

Therefore $B_{\alpha i}^{in}$ indicates the normalized in-strength of node i in layer α and $B_{\alpha i}^{out}$ indicates the normalized out-strength of node i in layer α. For undirected multiplex networks the matrices $\mathbf{B}^{in}$ and $\mathbf{B}^{out}$ are identical. Note that if node i is not connected in layer α we have $B_{\alpha i}^{in} = B_{\alpha i}^{out} = 0$ and we say that the node is *inactive* in layer α [21, 22]. Otherwise, if node i is connected in layer α, we will say that the node is *active* in layer α.

We assume that the centrality of a node depends on the centrality of the layers in which it is connected and additionally we assume that the centrality of the layers depends on the centrality of the nodes that are active on it.

Specifically *the centrality of a given node increases when already central nodes point to it in layers with high influence.*

Therefore, given the influences $z^{[\alpha]}$ of the layers α the centrality X_i of a node can be taken for instance to be determined by the PageRank centrality associated to the weighted adjacency matrix $\mathbf{G}$. In this case the centralities of the nodes are determined by the steady state of a random walker that, starting from a node j, with probability $\tilde{\alpha}$ hops to a neighbor node i of the aggregated network chosen with probability proportional to G_{ji} and with probability $1 - \tilde{\alpha}$ jumps to a random node which is not isolated in the weighted network described by the adjacency matrix $\mathbf{G}$.

Therefore the equations for the centrality X_i of the node i given the influences $z^{[\alpha]}$ of the layers read

$$X_i = \tilde{\alpha} \sum_{j=1}^{N} \frac{G_{ji}}{\kappa_j} X_j + \beta v_i \tag{4}$$

where $\tilde{\alpha}$ is taken to be $\tilde{\alpha} = 0.85$ (as usual in the context of the PageRank algorithms) and κ_j, v_i and
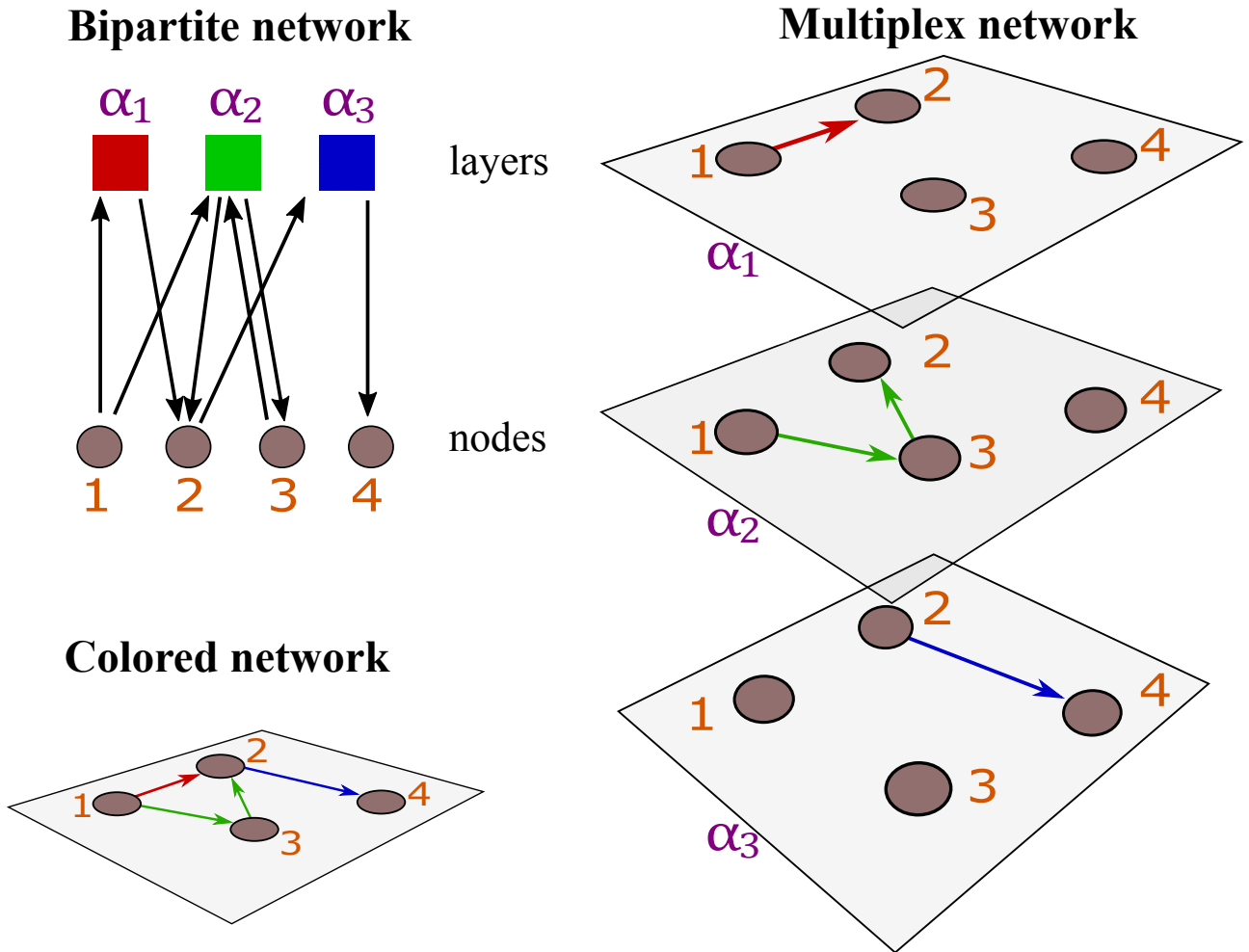
Figure 1: The schematic representation of a multiplex network formed by three different layers is here shown together with the construction of the directed bipartite network formed by nodes and layers and the multiplex representation as a colored network. The directed bipartite network indicates for each node in which layers the node is connected. Additionally, as it is explained in the text it gives direct information about in-strengths and out-strength of each node in any given layer.

β are given by

$$
\begin{aligned}
\kappa_j &= \max\left(1, \sum_{i=1}^{N} G_{ji}\right), \\
v_i &= \theta\left(\sum_{j=1}^{N}[G_{ij} + G_{ji}]\right), \\
\beta &= \frac{1}{\sum_{i=1}^{N} v_i} \sum_{j=1}^{N}\left[1 - \tilde{\alpha}\,\theta\left(\sum_{i=1}^{N} G_{ji}\right)\right] X_j.
\end{aligned}
\tag{5}
$$

This algorithm is a variation of the recently formulated Functional Multiplex PageRank [44] as long as the influences of the layers $\{z^{[\alpha]}\}$ are considered as free parameters [47]. In particular the Functional Multiplex PageRank determines the centrality of each node as a function of the set of influences associated to the layers of the multiplex network. This approach, although quite useful in the context of multiplex networks with few layers, can become prohibitive in the case of multiplex networks with many layers, i.e. for $M \gg 1$. Therefore here we propose to couple the Eqs. (4) for the centrality X_i of a generic node i of the multiplex network with another set of equations determining the best value for the influences of the layers. In order to determine these equations we will exploit the nature of the bipartite network between nodes and layers.

Our starting point will be that *layers are more influential if highly central nodes are active in them*. The simplest equation enforcing this principle is therefore

$$
z^{[\alpha]} = \frac{1}{\mathcal{N}} W^{[\alpha]} \sum_{i=1}^{N} B_{\alpha i}^{in} X_i
\tag{6}
$$

with $\mathcal{N}$ indicating a normalization constant. First of all, we note here that this equation for $z^{[\alpha]}$ rewards layers with larger total weight $W^{[\alpha]}$. For unweighted multiplex networks this implies that layers with more links are more influential while for weighted multiplex networks this implies that the layers with more total weight are more influential. Additionally this equation establishes layers having a larger number of central nodes with large in-strength as more influential. We note here that this algorithm is inspired by the algorithm proposed in Ref. [7] for ranking countries and products in bipartite trade networks but differs most notably from that model because the Eq. (4) for the centrality of the nodes X_i depends on the full underlying multiplex network structure, i.e. it depends not only on the layers where node i is active but also on the connections between different nodes existing in those layers.

The linear Eq. (6) for the centrality of the nodes X_i can be modified by including a tuning parameter $\gamma > 0$ as in the following

$$
z^{[\alpha]} = \frac{1}{\mathcal{N}} W^{[\alpha]} \sum_{i=1}^{N} B_{\alpha i}^{in}[X_i]^{\gamma},
\tag{7}
$$

where $\mathcal{N}$ indicates a normalization constant. For $\gamma = 1$ (linear case) we recover Eq. (6). For $\gamma < 1$ the nodes with low centrality contribute more than in the linear case while for $\gamma > 1$ their contribution is suppressed with respect to the linear case.

The Eq. (7) can be coupled to Eq. (4) to get a simultaneous ranking of nodes $\{X_i\}$ and layers $\{z^{[\alpha]}\}$. Specifically the algorithm can be used in the following cases:

i) layers with larger total weight $W^{[\alpha]}$ are assumed to be more influential;

6

ii) layers with more active nodes of high centrality and high in-strength are considered more influential.

However sometimes we might desire to assign to the layers of a multiplex network an influence that evaluates the centrality of the layers independently of their different density and/or total weight, i.e. independently of $W^{[\alpha]}$. In this case we propose the following normalized equations

$$z^{[\alpha]} = \frac{1}{\mathcal{N}} \sum_{i=1}^{N} B_{\alpha i}^{in} [X_i]^{\gamma} \tag{8}$$

with $\mathcal{N}$ indicating a normalization constant. Therefore Eq. (8) can be used when the ranking of the layers needs to satisfy the following two conditions:

i) the influence of each layer is required to be independent of $W^{[\alpha]}$;

ii) layers with more active nodes of high centrality and high in-strength are considered more influential.

A different approach consists in attributing more influence to *elite layers* with *few* highly central nodes with incoming links. In the formulation of these alternative equations for the determination of the influences $\{z^{[\alpha]}\}$ of the layers, we have been inspired by the recently proposed algorithm [8, 9] for ranking nodes in the bipartite network of countries and products. Specifically it is possible to propose the following set of equations for determining the values of the layers influences,

$$z^{[\alpha]} = \frac{1}{\mathcal{N}} W^{[\alpha]} \left[\sum_{i=1}^{N} B_{\alpha i}^{in} (X_i)^{-1} \right]^{-1}, \tag{9}$$

where $\mathcal{N}$ is a normalization constant. Here layers with more overall weights $W^{[\alpha]}$ are more influential, and whereas the overall weight of the layers is the same, layers with fewer highly central nodes receiving in-links are more influential. For instance in the Trade Multiplex Network a product is more important if the overall world-wide trade is very conspicuous for that product and only few central nodes are exporting it.

Similarly to the previously discussed case, we can also introduce a parameter $\gamma > 0$ and consider the following equations for the influences $z^{[\alpha]}$ of the layers,

$$z^{[\alpha]} = \frac{1}{\mathcal{N}} W^{[\alpha]} \left[\sum_{i=1}^{N} B_{\alpha i}^{in} (X_i)^{-\gamma} \right]^{-1} \tag{10}$$

where $\mathcal{N}$ is a normalization constant. For $\gamma = 1$ (linear case) we recover Eq. (9). For $\gamma \neq 1$ the contribution of low centrality nodes in determining the centrality of the layers in which they are active can be either enhanced ($\gamma > 1$) or suppressed ($\gamma < 1$).

The Eqs. (10) can be used if the ranking of the layer must satisfy the following set of conditions:

i) layers with larger total weight $W^{[\alpha]}$ are associated to a larger influence;

ii) layers with fewer active nodes of high centrality are considered more influential.

Alternatively we might consider equations in which we normalize by the total weights of the layers $W^{[\alpha]}$, getting

$$z^{[\alpha]} = \frac{1}{\mathcal{N}} \left[\sum_{i=1}^{N} B_{\alpha i}^{in} (X_i)^{-\gamma} \right]^{-1}. \tag{11}$$

This last set of equations can be used when the ranking of the layers must satisfy the following conditions:

i) the influence of each layer is required to be independent of $W^{[\alpha]}$.

ii) layers with more active nodes of high centrality and high in-strength are considered more influential.

2.2 MultiRank: The definition

Summarizing the discussion of the previous section here we introduce the MultiRank algorithm depending on three parameters: $\gamma > 0$, s taking values $s = 1, -1$ and a taking values $a = 1, 0$.

The MultiRank algorithm assigns a centrality X_i to each node i and an influence $z^{[\alpha]}$ to each layer α which can be found by solving the following set of coupled equations

$$X_i \;=\; \tilde{\alpha} \sum_{j=1}^{N} \frac{G_{ji}}{\kappa_j} X_j + \beta v_i \tag{12}$$

$$z^{[\alpha]} \;=\; \frac{1}{\mathcal{N}} \left[W^{[\alpha]} \right]^a \left[\sum_{i=1}^{N} B_{\alpha i}^{in} \left(X_i \right)^{s\gamma} \right]^s \tag{13}$$

where $\tilde{\alpha}$ is taken to be $\tilde{\alpha} = 0.85$ and κ_j, v_i and β are given by Eq. (5) and $\mathcal{N}$ is a normalization constant.

- For $a = 1$ the influence of a layer is proportional to $W^{[\alpha]}$.

- For $a = 0$, the influence of a layer is normalized with respect to $W^{[\alpha]}$.

- For $s = 1$ layers have larger influence if they include more central nodes. In this case the parameter γ can be tuned to either suppress ($\gamma > 1$) or enhance ($\gamma < 1$) the contribution of low centrality nodes in determining the centrality of the layers in which they are active.

- For $s = -1$ however, layers have larger influence if they include *fewer* highly influential nodes. In other words this algorithm awards *elite layers*. In this case the parameter γ can be tuned to either enhance ($\gamma > 1$) or suppress ($\gamma < 1$) the contribution of low centrality nodes in determining the centrality of the layers in which they are active.

- For $a = 0$ and $\gamma = 0$ the influences of the layers are all equal and the MultiRank reduces to the Pagerank of the aggregated network. For all the other possible values of the parameters the MultiRank is not reducible to the PageRank of the aggregated network.

We implemented the MultiRank algorithm in a MATLAB code that is publicly available [48]. The code has a computational complexity $\mathcal{O}(N^2 M)$ as tested on an ensemble of multiplex networks formed by uncorrelated Poisson layers.

2.3 Possible variations of the MultiRank algorithm

The MultiRank algorithm, is fully defined by the Eqs. (12) and (13). However this algorithm can be modified by changing the first equation (i.e. Eq. (12)) determining the centrality of the nodes. In fact instead of adopting a PageRank algorithm for the centrality of the nodes, it is possible to consider either the eigenvector centrality or the Katz centrality [2]. In the first case Eq. (12) could be substituted by

$$X_i \;=\; \frac{\sum_{j=1}^{N} G_{ji} X_j}{\sum_{r=1}^{N} \sum_{j=1}^{N} G_{jr} X_j}. \tag{14}$$
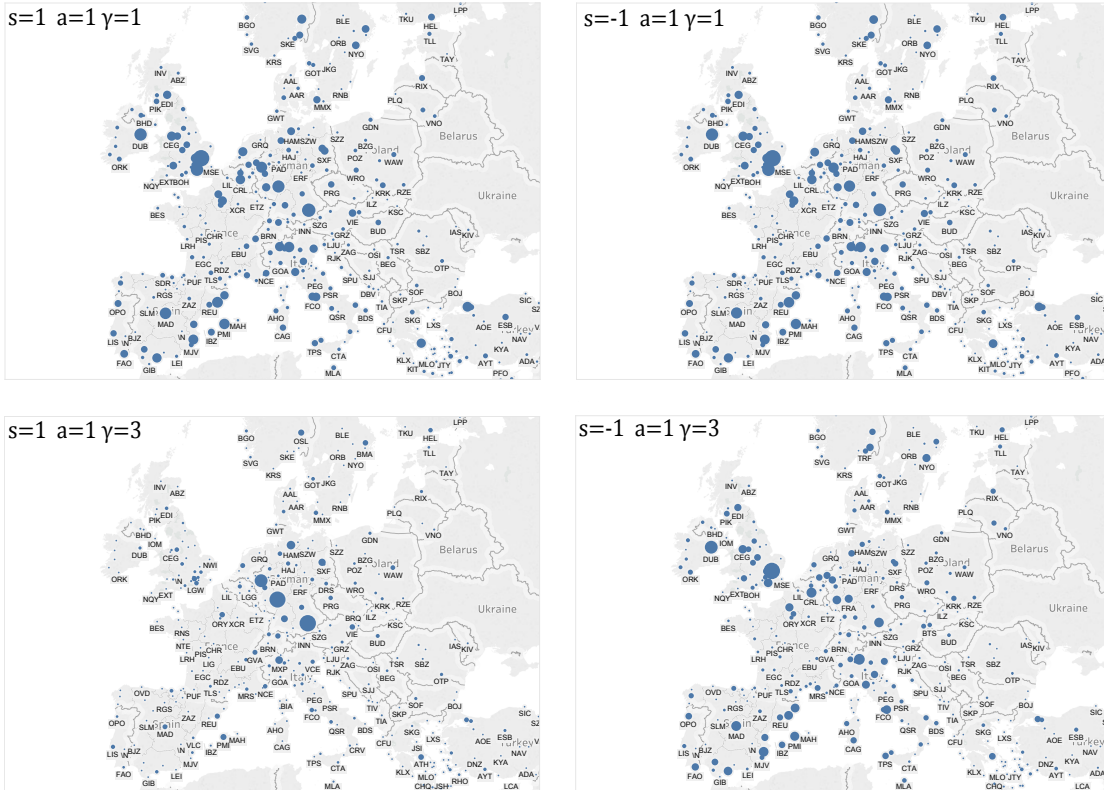
Figure 2: The maps representing the centrality X_i of the European airports in the European Air Transportation Multiplex Network according to the MultiRank algorithm are here shown for different values of the parameters $s = -1, 1$; $a = 1$ and $\gamma = 1, 3$. In the case $s = 1, a = 1$ we observe that the centrality of the nodes are more heterogeneously distributed for $\gamma = 3$ than for $\gamma = 1$ with few airports acquiring a centrality much higher than the others.

In the second case Eq. (12) could be substituted by

$$X_i = \hat{\alpha} \sum_{j=1}^{N} G_{ji} X_j + v_i,$$

(15)

where $\hat{\alpha} > 0$ is suitably chosen to ensure the convergence of the algorithm and v_i is defined in Eq. (5).

3   APPLICATION TO REAL DATASETS

In this section we will consider the application of the MultiRank algorithm defined in Sec. 2.2 to three main examples of multiplex datasets: the European Air Transportation Multiplex Network [31], the Pierre Auger Multiplex Collaboration Network [34] and the FAO Multiplex Trade Network [30]. All datasets are publicly available at the website [49].

Table 1 summarizes the main properties of the three datasets and the number of iterations necessary for convergence of the MultiRank algorithm.

3.1   European Air Transportation Multiplex Network

The European Air Transportation Multiplex Network [31] is a multiplex network dataset comprising $M = 37$ layers (European airline companies) and $N = 450$ nodes (European airports). Each layer
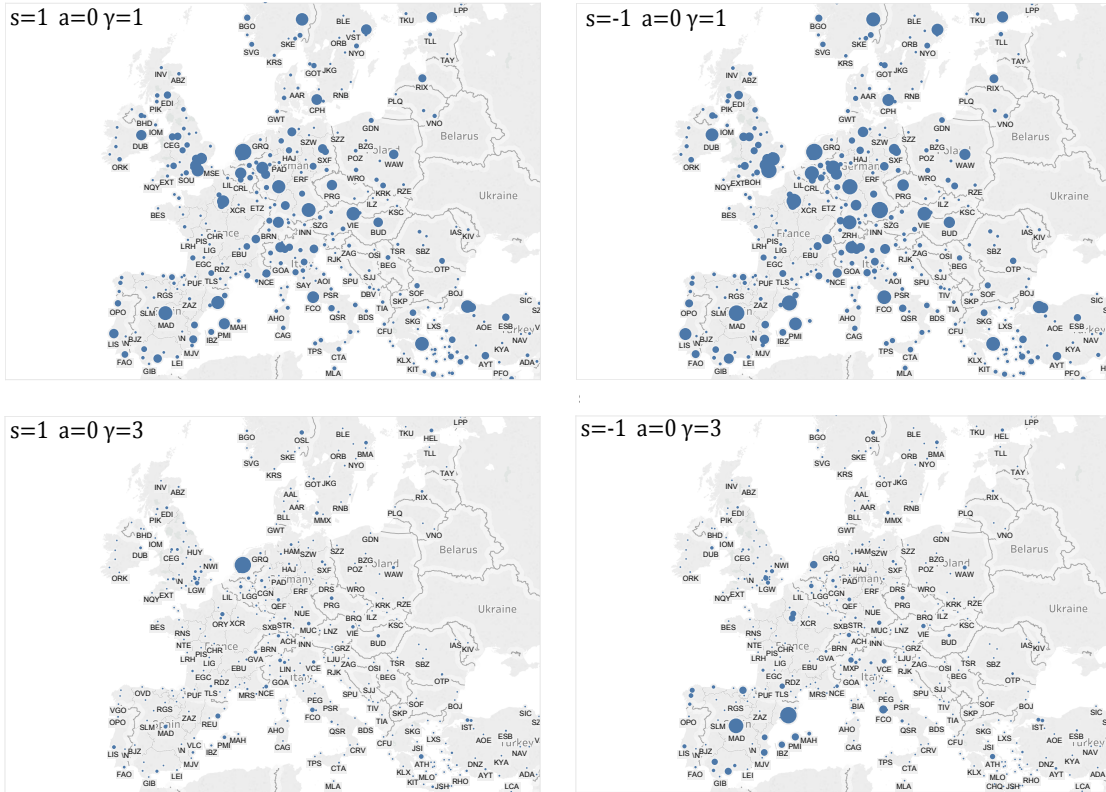
9

Figure 3: The maps representing the centrality X_i of the European airports in the European Air Transportation Multiplex Network according to the MultiRank algorithm are here shown for different values of the parameters $s = -1, 1$; $a = 0$ and $\gamma = 1, 3$. As for $s = 1$ and $a = 1$ (maps reported in Figure 2) we observe that for $\gamma = 3$ the centralities of the nodes are more heterogeneously distributed than for $\gamma = 1$.

is unweighted and undirected and is formed by the flight connections operated by a given European airline company. The dataset is widely used in the context of multiplex network research as one of the most interesting, clean, well defined and freely available datasets where to test new measures and algorithms for the extraction of relevant information. Here we discuss the application of the MultiRank algorithm to this dataset. Therefore our goal will be to rank simultaneously European airports and airline companies.

In Figure 2 and Figure 3 we show a map of Europe where different airports are assigned the MultiRank centrality for different values of the parameters s, a and γ. The distribution of the node centralities tends to become significantly more skewed for larger values of γ. In order to capture this phenomena in Figure 4 we report the distribution $P(X)$ of the airports MultiRank centralities X for the same parameters used in Figure 2 and in Figure 3.

In Figure 5 and in Figure 6 we plot respectively the rank of the top 15 European airline companies and of the rank of the top 15 European airports according to the MultiRank evaluated for $s = 1, -1$ and $a = 0, 1$. For the relevant case $s = -1, a = 1$ we observe a stable ranking of the top 4 ranked airline companies as a function of γ given by Ryanair, Easyjet, Lufthansa, Air Berlin and rather stable ranking of airports with Stansted airport (STN) remaining the rank one airport for all the considered range of values of γ. When we calculate the MultiRank algorithm for different values of the parameter s and a, with $(s, a) \neq (-1, 1)$ we note that the algorithm experience an instability for $\gamma = \gamma^\star > 1$.

The instability observed for $s = 1$ at large values of γ can be explained as the result of the suppressed role that low-rank airports have in determining the centrality of the layers in which they are active.

Dataset	N	M	I	Directed	Weighted
EA	450	37	20-60	No	No
PA	514	16	42-44	No	No
FAO	214	364	16-18	Yes	Yes

Table 1: Main properties of the three analyzed datasets: the European Air Transportation Multiplex Network (EA) [31], the Pierre Auger Multiplex Collaboration Network (PA) [34] and the FAO Multiplex Trade Network (FAO) [30]. The quantities specified are N, the total number of nodes, M the total number of layers and I, the number of iterations necessary for a convergence of the MultiRank algorithm within a tolerance of 10^{-4} on the square error of the vector of node centralities X_i. Since for every value of the parameters the number of iterations I can vary we have indicated the range of variability of I for each one of the three datatsets. Additionally we have indicated if the network is directed and/or weighted.
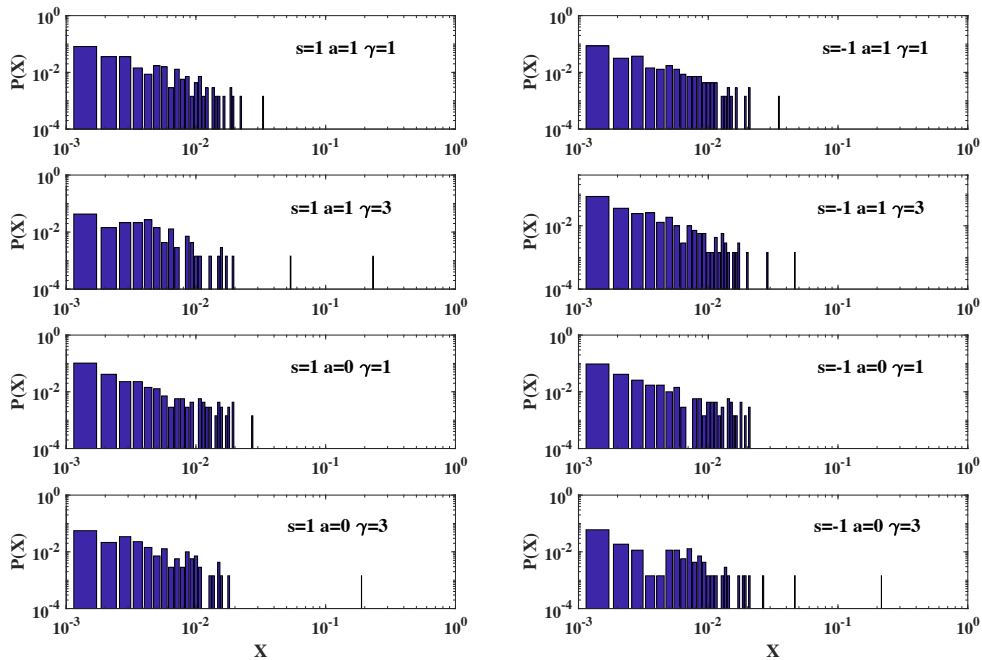

Figure 4: Distribution $P(X)$ of the MultiRank centralities X of the airports in the European Air Transportation Multiplex Network for the same parameters $s = 1, -1$, $a = 0, 1$ and $\gamma = 1, 3$ considered in Figure 2 and Figure 3.

For instance this observation well describes the fact that for $s = 1, a = 1$ a low cost airline (such as Ryanair) that typically serve either its own hub airports or much smaller airports, decrease its rank for larger value of γ to the advantage of a major carrier such as Lufthansa.

On the contrary for $s = -1$ at large value of γ low rank airports acquire a major role in determining the centrality of the layers in which they are active. This effect is even more pronounced for $a = 0$ when the influence of the airline companies (layers) is not proportional to the number of flight connections. This explains for instance that for $s = -1, a = 0$ Lufthansa decreases its rank as γ increases.

The maps reported in Figure 3 showing the MultiRank centrality of the airports for different MultiRank parameters describe visually the effect of the instability observed for $(s, a) \neq (-1, 1)$. For $\gamma = 3$, after these instabilities have taken place, we observe that few airports acquire a centrality much higher than all the other airports. For instance in the case $s = 1, a = 1$ well before the instability, for $\gamma = 1$, the airline Ryanair and its major airport Stansted (STN) are on the top of the rank, for $\gamma = 3$, above the instability, we observe the two major Lufthansa airports, i.e. Munich Airport (MUC) and
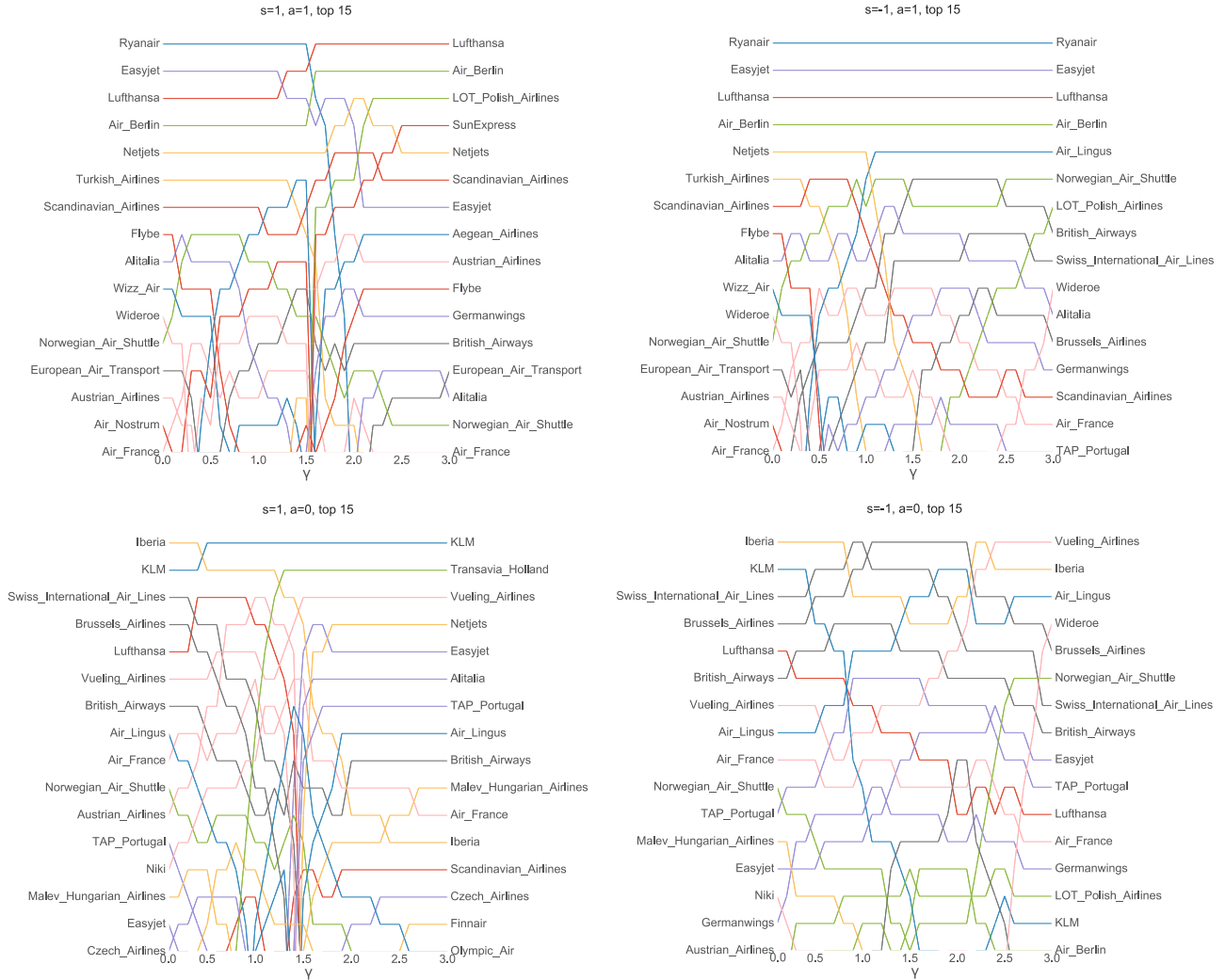
Figure 5: Ranking of the top 15 European airline companies in the European Air Transportation Multiplex Network (listed from top to bottom in order of decreasing centrality) according to the MultiRank algorithm are here shown for different values of the parameters $s = -1, 1$ and $a = 0, 1$ as a function of $\gamma \in (0, 3)$.

Frankfurt Airport (FRA), at the top of the airport rankings acquiring a centrality much higher than all the other European airports.

3.2 Pierre Auger Multiplex Collaboration Network

The Pierre Auger Multiplex Collaboration Network [34] describes the scientific collaboration networks established in the framework of the Pierre Auger experiment. The Pierre Auger experiment [50] aims at constructing a large array of detectors for studying cosmic rays.

The Pierre Auger Multiplex Collaboration Network is formed by $M = 16$ layers (scientific topics) and $N = 514$ nodes (scientists). Each layer is a network characterizing the collaborations in a given scientific topic. Each layer of this multiplex network is undirected and unweighted. The names of the nodes (scientists) have been anonymized while the names of the layers are available. These layers are: Neutrinos, Detector, Enhancements, Anisotropy, Point-source, Mass-composition, Horizontal, Hybrid-reconstruction, Spectrum, Photons, Atmospheric, SD-reconstruction, Hadronic-interactions, Exotics, Magnetic, Astrophysical-scenarios.

We have run the MultiRank algorithm on this dataset ranking simultaneously nodes and layers.
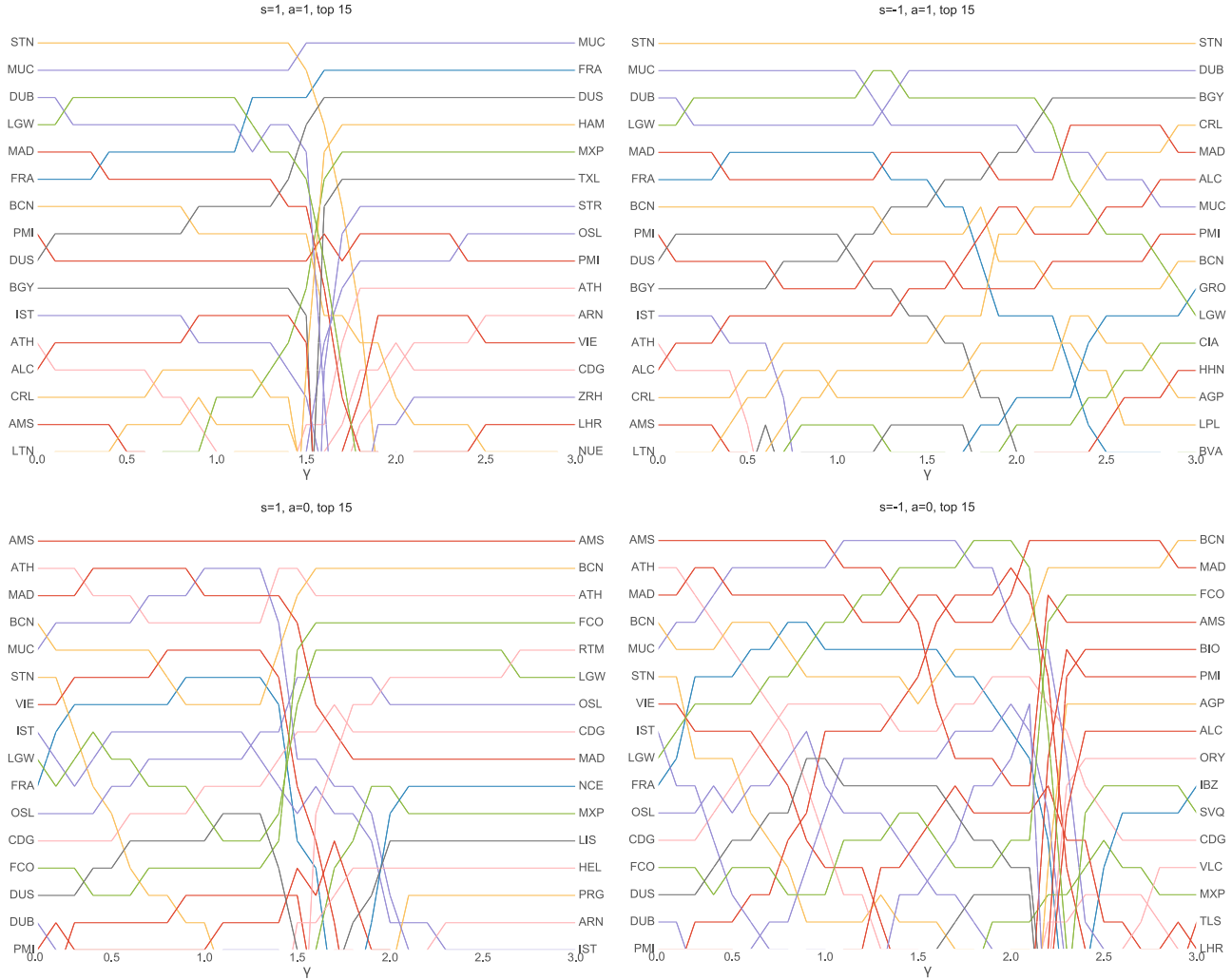
Figure 6: Ranking of the top 15 European airports in the European Air Transportation Multiplex Network (listed from top to bottom in order of decreasing centrality) according to the MultiRank algorithm are here shown for different values of the parameters $s = -1, 1$ and $a = 0, 1$ as a function of $\gamma \in (0, 3)$.

However in Figure 7 we display only the results relative to the layers due to the fact that the node labels and identities are not disclosed.

The MultiRank results clearly establish the importance of experimental advances for the Pierre Auger Multiplex Collaboration Network. We observe that for each of the cases $s = 1, -1$ and $a = 1, 0$ the topic Enhancements is at the top of the ranking. Its position is actually stable in every case with the exception of the case $s = 1, a = 0$. For cases $a = 1$ we observe in second position (rank 2) the topic Detector both in case $s = 1$ and in case $s = -1$.

Overall, the ranking of the layers seems to be more stable for the cases $a = 1$ than for the cases $a = 0$. Notably, while running the MultiRank on this dataset we have not detected any instability in this case.

3.3 FAO Multiplex Trade Network

As a third example of a multiplex network, we have considered here the FAO Multiplex Trade Network [30] collecting data coming from the Food and Agriculture Organization of the United Nations (FAO).

This is a major example of a Trade Network between countries. Other similar multiplex trade

Figure 7: Ranking of the top 15 topics in the Pierre Auger Multiplex Collaboration Network (listed from top to bottom in order of decreasing centrality) according to the MultiRank agorithm are here shown for different values of the parameters $s = -1, 1$ and $a = 0, 1$ as a function of $\gamma \in (0, 3)$.
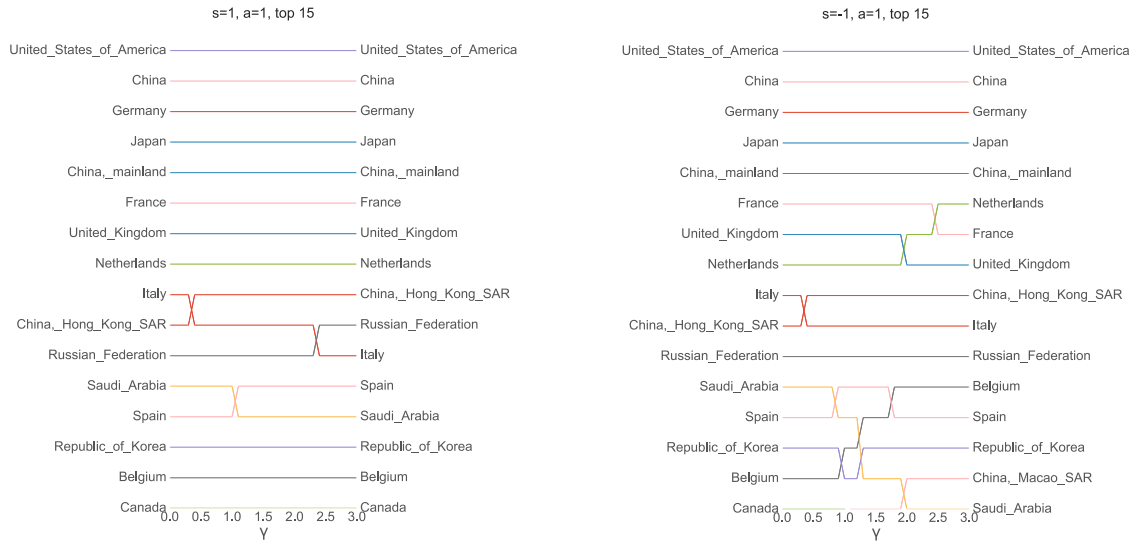
Figure 8: Ranking of the top 15 countries in the FAO Multiplex Trade Network (listed from top to bottom in order of decreasing centrality) according to the MultiRank algorithm are here shown for different values of the parameters $s = -1, 1$ and $a = 1$ as a function of $\gamma \in (0, 3)$.
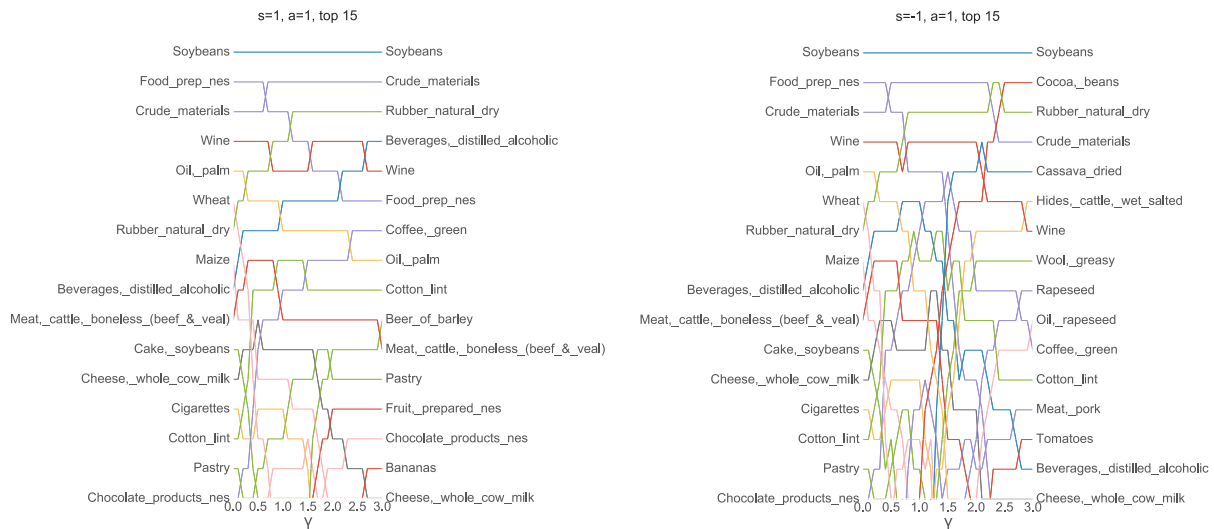

Figure 9: Ranking of the top 15 products in the FAO Multiplex Trade Network (listed from top to bottom in order of decreasing centrality) according to the MultiRank algorithm are here shown for different values of the parameters $s = -1, 1$ and $a = 1$ as a function of $\gamma \in (0, 3)$.

networks including also other types of products have been extensively studied in the literature [27–29, 51].

This dataset includes $N = 214$ nodes (countries) and $M = 364$ layers (food and agriculture products). Each layer is weighted and directed characterizing which country is importing from which country a given product, and indicating the amount of the imported values in US dollars. Therefore the network is in this case both directed and weighted.

We characterized the MultiRank of this dataset by ranking simultaneously countries and products. Note that given the definition of MultiRank we are taking into account not only the bipartite nature of the graph determining which country exports which product as in Refs. [7–9] but instead we consider the overall multiplex network structure of the FAO Multiplex Trade Network. In fact we process the information of which country is trading a given product with which other country.

In this case it is natural to focus our attention on the MultiRank algorithm with $a = 1$, since the importance of a product should be clearly related to the total amount in US dollars of its international trade.

In Figures 8 and 9 we report respectively the rank of the top 15 countries and 15 products as a function of $\gamma \in (0, 3)$ for different values of the parameter s, i.e. for $s = 1$ and $s = -1$, and for $a = 1$.

We observe a very stable list of top ranked countries: United States, China, Germany, Japan, China Mainland, France both for $s = 1$ and for $s = -1$. For the products we observe in top rank positions a composition of prime necessity products such as Soybeans and Crude-materials and of more expensive products which are in high demand such as Wine, and Cocoa-beans.

4 Comparison with other rankings

The MultiRank algorithm is a novel centrality measure that unlike previously existing algorithms does not only rank nodes but also ranks the layers of a multiplex network. Also in the case in which we just focus on the ranking of the nodes, as long as $\gamma > 0$ the MultiRank is not reducible to previously considered rankings. This is a desirable property because it implies that the MultiRank is an independent centrality measure. However it is desirable to test whether the MultiRank it a meaninful centrality measure and whether it is in line with other observations and indicators of the centrality of the nodes. A way to check the consistency of the results obtained with the MultiRank algorithm is to measure the correlations between the MultiRank of the nodes and the ranking of the nodes provided by some other independent measure.

For the European Air Transportation Multiplex Network we have considered the correlations between the MultiRank of the airports and their ranking according to their total number of passengers. The data on the total number of passengers for each European airport has been taken from the Eurostat webpage [52] for the same year (2011) in which the European Air Multiplex Transportation Network has been collected. The ranking of the top 100 airports with largest number of passengers is compared with their MultiRank. The Spearman rank correlations between the rankings is shown in Figure 10A and is very significant as the P-value ranges between 1×10^{-17} and 5×10^{-17}.

For the FAO Multiplex Trade Network it is possible to compare the MultiRank of the countries with their ranking according to the GDP. The data of the GDP of the countries has been collected from the World Bank database [53] for the same year (2010) in which the FAO Multiplex Trade Network has been collected. The ranking of the top 100 countries with largest GDP is compared with their MultiRank. The Spearman rank correlations between the rankings shown in Figure 10B are very significant and correspond to a P-value ranging between 2×10^{-27} and 2×10^{-14} depending on the MultiRank parameters.

Therefore in both cases we observe that the ranking provided by the MultiRank is in line with other independent observations. Additionally we observe that the Spearman rank correlations are in some cases higher than the Spearman correlations calculated using instead of the MultiRank the PageRank

of the aggregated network. In both cases we observe that the MultiRank with $s = -1$ and $a = 0$ achieves higher Spearman correlations.

However in none of the considered parameter values the ranking provided by MultiRank is reducible to the rank obtained by the passenger number or the GDP showing that the MultiRank is an independent measure of centrality.
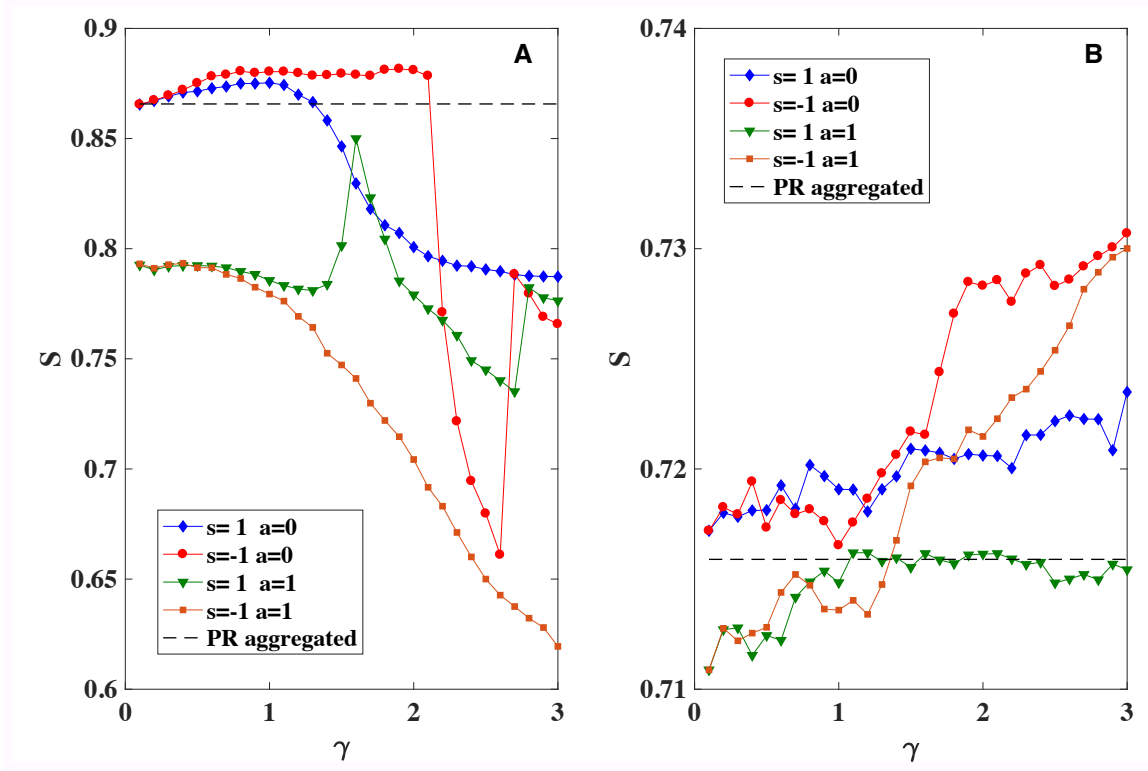

Figure 10: In panel A the Spearman rank correlation S between the ranking of the top 100 airports with largest number of passengers and their MultiRank is shown as a function of γ for the parameter values $s = 1, -1$ and $a = 0, 1$. The solid dashed line in panel A indicates the Spearman rank correlation between the ranking of the airports determined by the number of passengers and the ranking provided by the PageRank of the aggregated network. In panel B the Spearman rank correlation S between the ranking of the top 100 countries with largest GDP and their MultiRank is shown as a function of γ for the parameter values $s = 1, -1$ and $a = 0, 1$. The solid dashed line in panel B indicates the Spearman rank correlation between the ranking determined by the GDP and the ranking provided by the PageRank of the aggregated network.

5 CONCLUSIONS

In this paper we have presented the MultiRank algorithm for ranking nodes and layers in large multiplex networks. The ranking of nodes and layers is determined by a coupled set of equations. The centrality of the nodes is evaluated according to a PageRank algorithm based on a weighted random walk where the probability to hop to a neighbor node of a given layer α depends on the *influence* (centrality) of the layer. Vice-versa, the influences of the layers are determined by the centrality of the nodes that are active on it. The equations determining the influence of the layers exploit the bipartite structure of the network between nodes and layers that can be constructed starting from the complete information about the multiplex structure.

The MultiRank depends on two main parameters s and a. The parameter a allows to attribute to

layers an influence proportional to their total weight or alternatively to attribute an influence that does not take into account the total weight of the layers. The parameter s instead, allows us to attribute more influence to more popular layers or to elite layers.

The proposed algorithm is very flexible and is able to efficiently rank multiplex networks with many layers. It can be applied to undirected and directed multiplex networks with weighted or unweighted links.

We have shown that the MultiRank is able to extract relevant information from the datasets of three major examples of multiplex networks (including the European Air Transportation Multiplex Network, the Pierre Auger Multiplex Collaboration Network and the weighted and directed FAO Trade Multiplex Network).

In conclusion this algorithm provides a framework to evaluate the centrality of nodes and layers in multiplex networks. It combines recent developments obtained for evaluating the centrality of bipartite networks with state of the art research on multilayer networks.

Finally we believe that the proposed algorithm could be useful for a variety of different applications, including not only transportation, social and economic multiplex networks but also molecular and brain networks. We hope that this work will stimulate further research in this direction.

Funding

Acknowledgments

Author Contributions Statement

C. R., A. A. and G.B. designed the research, C. R. and G. B conducted the research, C. R. and J. I. prepared the figures and all the authors wrote the main manuscript text.

Additional Information

The authors declare that they have no competing financial interests.
Correspondence to christoph.rahmede@kit.edu, j.iacovacci@qmul.ac.uk, alexandre.arenas@urv.cat, ginestra.bianconi@gmail.com

References

[1] Brin, S., & Page, L. (1998). The anatomy of a large-scale hypertextual web search engine. *Computer networks and ISDN systems*, **30**, 107.

[2] Newman, M. E. J. (2010) *Networks: An introduction*, Oxford University Press, Oxford.

[3] Radicchi, F., Fortunato, S., Markines, B., & Vespignani, A. (2009). Diffusion of scientific credits and the ranking of scientists. *Phys. Rev. E* **80**, 056103.

[4] Meiss, M. R., Menczer, F., Fortunato, S., Flammini, A.,& Vespignani, A. (2008, February). Ranking web sites with real user traffic. *In Proceedings of the 2008 International Conference on Web Search and Data Mining* (pp. 65-76). ACM.

[5] Ghoshal, G., & Barabási, A. L. (2011) Ranking stability and super-stable nodes in complex networks. *Nature Comm.* **2**, 394.

[6] Blumm N, et al. (2012) Dynamics of ranking processes in complex systems. *Phys. Rev. Lett.* **109**, 128701.

[7] Hidalgo, C. A., Klinger, B., Barabási, A. L., & Hausmann, R. (2007). The product space conditions the development of nations. *Science* **317**, 482.

[8] Tacchella, A., Cristelli, M., Caldarelli, G., Gabrielli, A., & Pietronero, L. (2012). A new metrics for countries' fitness and products' complexity. *Sci. Rep.* **2**, 723.

[9] Cristelli, M., Gabrielli, A., Tacchella, A., Caldarelli, G., & Pietronero, L. (2013). Measuring the intangibles: A metrics for the economic complexity of countries and products. *PloS one* **8**, e70726.

[10] Domínguez-García, V., & Muñoz, M. A. (2015). Ranking species in mutualistic networks. *Sci. Rep.* **5**, 8182.

[11] Boccaletti, S., et al. (2014). The structure and dynamics of multilayer networks. *Phys. Rep.* **544**, 1.

[12] Kivelä, M. et al. (2014). Multilayer networks. *J. Complex Net.* **2**, 203.

[13] Bianconi, G. (2015). Interdisciplinary and physics challenges of network theory. *EPL* **111**, 56001.

[14] De Domenico, M., Granell, C., Porter, M. A., & Arenas, A. (2016). The physics of spreading processes in multilayer networks. *Nature Physics* **12**, 901.

[15] Buldyrev, S. V., Parshani, R., Paul, G., Stanley, H. E., & Havlin, S. (2010). Catastrophic cascade of failures in interdependent networks. *Nature* **464**, 1025 (2010).

[16] De Domenico, M., Porter, M. A., & Arenas, A. (2015). MuxViz: a tool for multilayer analysis and visualization of networks. *Jour. Comp. Net.* **3**, 159.

[17] De Domenico, M., et al. (2013). Mathematical formulation of multilayer networks. *Phys. Rev. X* **3**, 041022.

[18] Lee, K. M., Kim, J. Y., Lee, S., & Goh, K. I. (2014). Multiplex networks. In *Networks of networks: The last frontier of complexity* (pp. 53-72). Springer International Publishing.

[19] Bianconi, G. (2013). Statistical mechanics of multiplex networks: Entropy and overlap. *Phys. Rev. E* **87**, 062806.

[20] Menichetti, G., Remondini, D., Panzarasa, P., Mondragón, R. J., & Bianconi, G. (2014). Weighted multiplex networks. *PloS One* **9**, e97857.

[21] Cellai, D., & Bianconi, G. (2016). Multiplex networks with heterogeneous activities of the nodes. *Phys. Rev. E* **93**, 032302.

[22] Nicosia, V., & Latora, V. (2015). Measuring and modeling correlations in multiplex networks. *Phys.Rev. E* **92**, 032805.

[23] Gomez, S., et al. (2013). Diffusion dynamics on multiplex networks. *Phys. Rev. Lett.* **110**, 028701.

[24] De Domenico, M., Solé-Ribalta, A., Gómez, S., & Arenas, A. (2014). Navigability of interconnected networks under random failures. *Proc. Nat. Aca. Sci.* **111**, 8351.

[25] Baxter, G. J., Dorogovtsev, S. N., Goltsev, A. V.,& Mendes, J. F. F. (2012). Avalanche collapse of interdependent networks. *Phys. Rev. Lett.* **109**, 248701.

[26] Radicchi, F., & Bianconi, G. (2017). Redundant interdependencies boost the robustness of multi-layer networks. *Phys. Rev. X* **7**, 011013.

[27] Barigozzi, M., Fagiolo, G., & Garlaschelli, D. (2010). Multinetwork of international trade: A commodity-specific analysis. *Phys. Rev. E* **81**, 046104.

[28] Barigozzi, M., Fagiolo, G., & Mangioni, G. (2011). Identifying the community structure of the international-trade multi-network. *Physica A* **390**, 2051.

[29] Ducruet, C. (2013). Network diversity and maritime flows. *Jour. Trans. Geo.* **30**, 77.

[30] De Domenico, M., Nicosia, V., Arenas, A., & Latora, V. (2015). Structural reducibility of multi-layer networks. *Nature Comm.* **6**, 6864.

[31] Cardillo, A., et al. (2013). Emergence of network features from multiplexity. *Sci. Rep.* **3**, 1344.

[32] Mucha, P. J., Richardson, T., Macon, K., Porter, M. A., & Onnela, J. P. (2010). Community structure in time-dependent, multiscale, and multiplex networks. *Science* **328**, 876.

[33] Szell, M., Lambiotte, R., & Thurner, S. (2010). Multirelational organization of large-scale social networks in an online world. *Proc. Nat. Aca. Sci.* **107**, 13636.

[34] De Domenico, M., Lancichinetti, A., Arenas, A., & Rosvall, M. (2015). Identifying modular flows on multilayer networks reveals highly overlapping organization in interconnected systems. *Phys. Rev. X* **5**, 011027.

[35] Battiston, S., Caldarelli, G., May, R. M., Roukny, T., & Stiglitz, J. E. (2016). The price of complexity in financial networks. *Proc. Nat. Aca. Sci.* **113** 10031.

[36] Musmeci, N., Nicosia, V., Aste, T., Di Matteo, T., & Latora, V. (2016). The multiplex dependency structure of financial markets. arXiv preprint arXiv:1606.04872.

[37] Bullmore, E., & Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Rev. Neuro.* **10**, 186.

[38] Bentley, B., et al. (2016). The multilayer connectome of Caenorhabditis elegans. *PLOS Comp. Bio.* **12**, e1005283.

[39] Cantini, L., Medico, E., Fortunato, S., & Caselle, M. (2014). Detection of gene communities in multi-networks reveals cancer drivers. *Sci. Rep.* **5**, 17386.

[40] Halu, A., Mondragón, R. J., Panzarasa, P., & Bianconi, G. (2013). Multiplex pagerank. *PloS one* **8**, e78293.

[41] Iacovacci, J., & Bianconi, G. (2016). Extracting information from multiplex networks. *Chaos* **26**, 065306.

[42] De Domenico, M., Solé-Ribalta, A., Omodei, E., Gómez, S., & Arenas, A. (2015). Ranking in interconnected multilayer networks reveals versatile nodes. *Nature Comm.* **6**, 6868.

[43] Solá, L., Romance, M., Criado, R., Flores, J., García del Amo, A., & Boccaletti, S. (2013). Eigenvector centrality of nodes in multiplex networks. *Chaos* **23**, 033131.

[44] Iacovacci, J., Rahmede, C., Arenas, A., & Bianconi, G. (2016). Functional Multiplex PageRank. *EPL* **116**, 28004.

[45] Estrada, E., & Gómez-Gardeñes, J. (2014). Communicability reveals a transition to coordinated behavior in multiplex networks. *Phys. Rev. E*, **89**, 042819.

[46] Reiffers-Masson, A. & Labatut, V. (2017). Opinion-Based Centrality in Multiplex Networks: A Convex Optimization Approach. arXiv preprint arXiv: 1703.03741.

[47] Note that also in [44] the influences are not simply attributed to layers but they are assignment to each different pattern of connection that can exist between two nodes of the multiplex, i.e. every type of possible multilink.

[48] https://github.com/ginestrab/MultiRank

[49] http://deim.urv.cat/ manlio.dedomenico/data.php

[50] https://www.auger.org/

[51] Gemmetto, V., & Garlaschelli, D. (2015). Multiplexity versus correlation: the role of local constraints in real multiplexes. *Sci. Rep.* **5**, 9120.

[52] http://ec.europa.eu/eurostat/web/transport/data/database

[53] http://databank.worldbank.org/data