Clustering coefficients as measures of the complex interactions in a directed multiplex network

Paolo Bartesaghi\textsuperscript{a}, Gian Paolo Clemente\textsuperscript{b}, Rosanna Grassi\textsuperscript{a,}\textsuperscript{*}

\textsuperscript{a}University of Milano - Bicocca, Via Bicocca degli Arcimboldi 8, 20126 Milano, Italy, email: paolo.bartesaghi@unimib.it; rosanna.grassi@unimib.it
\textsuperscript{b}Università Cattolica del Sacro Cuore di Milano, Largo Gemelli 1, 20123 Milano, Italy email: gianpaolo.clemente@unicatt.it

Abstract

In this paper, we propose new definitions of clustering coefficient for weighted and directed multiplex networks. We extend in the multiplex framework the formulas of clustering coefficients existing in the literature for weighted directed monoplex networks. We quantify how deep a node is involved in a cohesive structure focusing on a single layer, all layers or the entire system, respectively. The coefficients capture various features intrinsically related to the complex topology of the multiplex network. We test their effectiveness applying them to a particularly complex structure such as the international trade network. The trade data integrate different aspects and they can be described by a directed and weighted multiplex network, where each layer represents import and export relationships between countries for a given sector. The proposed coefficients find successful application in describing the interrelations of the trade network, allowing to disentangle the effects of countries and sectors and jointly consider the interactions between them.

Keywords: Networks, Multiplex networks, Clustering Coefficient, Data Science, World Trade

1. Introduction

The coexistence of multiple types of interactions within interconnected systems motivated the study of the multilayer nature of real-world networks. In this context, the concept of heterogeneous networks has received increasing attention in the literature. These kinds of networks are currently classified in different ways, including

*Corresponding author. email: rosanna.grassi@unimib.it
heterogeneous information networks (Gupta & Kumar (2020); Chen et al. (2021)), multilayer or multiplex networks (De Domenico et al. (2013); Kivelä et al. (2014); Boccaletti et al. (2014)) and multidimensional networks (Berlingerio et al. (2013)).

Beyond this classification, all these networks encode in the same complex structure relations of different nature and with different meaning. Indeed, it is worth pointing out that networks usually have different features in different layers that cannot be properly highlighted aggregating the layers by the overlay or the projection operation (see, e.g., Battiston et al. (2014)). Additionally, the meaning of interlinks connecting nodes in different layers goes over the simple formal representation. These aspects motivate a study of the complex object represented by multilayer networks preserving as much as possible the original structure even using sophisticated mathematical tools.

In this framework, special attention is paid to the triangles that each node can form in a multilayer network, and in particular in a multiplex network with nondiagonal couplings. Indeed, in these networks, the neighbours of a node $i$ are all nodes connected to $i$ within the same level or across different layers. As a consequence, triangles may be closed in a different layer than they originated (see Cozzo et al. (2015)). Since the incidence of triangular patterns in the network is typically measured by the well-known clustering coefficient, alternative definitions of triangles have an impact on the clustering coefficient definition in multilayer systems. The clustering coefficient is indeed an important measure of network topology (Newman (2003); Watts & Strogatz (1998)). In applied frameworks it is intended to measure the degree to which nodes in a graph tend to cluster, being an effective measure of the “cliqueness” (see Ertem et al. (2016)).

Alternative definitions of clustering coefficient in multidimensional structures have been proposed (see, in particular, De Domenico et al. (2013); Cozzo et al. (2015); Bortesaghi et al. (2021)) and all of them revealed that a univocal extension is not possible in a multiplex context. These contributions focus only on undirected networks, hence neglecting directed patterns that could be relevant empirically.

We contribute to this literature proposing new local clustering coefficients for weighted directed multiplex networks. Since in a multiplex network a node can be clustered under different perspectives, we measure how deep a node is involved in a cohesive structure focusing on the layer, all the layers or the whole system, respectively. These alternative coefficients capture various features intrinsically related to the complex topology of the multiplex network.
We extend to the multiplex structure the well-known formulas of clustering coefficients existing in the literature for weighted directed monoplex networks (see Fagiolo (2007) and Clemente & Grassi (2018)). Additionally, we generalise to the directed case the coefficient provided in De Domenico et al. (2013).

The proposed coefficients have been tested considering international trade data based on the World Input-Output Database (Timmer et al. (2015)). Traditionally, these data have been studied considering a monoplex structure (see, e.g., Fagiolo et al. (2008, 2012), Maluck & Donner (2015), Picardi & Tajoli (2012, 2018)) or a bipartite network (see Cristelli et al. (2013), Cingolani et al. (2017), Saracco et al. (2015)). However, monoplex networks, based on either countries or sectors, and bipartite networks, that connect countries to products, cannot account for the full extent of interactions that typically occur among countries in the international production network. Even when trade networks are formalized as multiplex networks without non-diagonal couplings (see Mastrandrea et al. (2014), Menichetti et al. (2014)), in which nodes are countries and layers are sectors, connections between countries are allowed within layers and any involvement of countries in relationships between sectors is therefore neglected.

Trade data integrate different aspects, such as the occurrence of various sectors, of imports and exports. They can therefore be effectively described by a directed and weighted complex structure, where each layer represents import and export relationships between countries for a given sector. International trade data have been modeled by Barigozzi et al. (2010) using a weighted directed network structured on more than one layer, where each layer refers to a commodity-specific network of import/export between countries. Classical network indicators for directed network have been studied in order to capture the statistical similarities between countries and to track how these commodity-specific networks evolve in time. The study performed by Ren et al. (2020) on the multilayer world trade network in a given time interval reveals the particular nested structure that characterises this set of networks. This property allows to quantify the product’s complexity.

To fully take into account the architecture of trade relationships, a multiplex network with non-diagonal couplings is here considered in order to disentangle the effects of countries and sectors and to jointly consider the interactions between them. Indeed, modelling a complex system as a multiplex network allows to gain information that cannot be captured taking individual layers separately. Results show that the coefficients highlight the role of countries and sectors in the whole trade taking into
account both inter-layer and intra-layer flows. In particular, we identify countries that are densely interconnected in all sectors and, at the same time, we bring out countries that are prominent in specific sectors. Additionally, the multiplex structure allows to unveil relevant sectors in terms of interconnection in the whole trade.

The paper is organized as follows. In Section 2 we introduce the mathematical notation needed for dealing with multiplex networks. In Section 3 we provide the definition of triangles in a multidimensional directed framework. In Section 4 we provide the general formulations of the local clustering coefficients by using supradjacency matrices. Also, we show how the proposed coefficients are generalisations of the classical definitions provided in the literature for monoplex networks. The different meanings of the proposed coefficients are also emphasized by means of a simple example. In Section 5 we apply our proposal to the world trade network. Conclusions follow.

2. Preliminaries

In this section we recall some notations and definitions about directed graph (Harary, 1969; Bang-Jensen & Gutin, 2008) which are essential to introduce the notation we will use in the paper.

A directed network is formally represented by a directed graph \( G = (V, E) \), where \( V \) is a set of \( N \) vertices (or nodes) and \( E \subseteq V \times V \) is the set of the ordered pairs of elements of \( V \) (arcs or directed edges). Two nodes are adjacent if there is an arc \((i, j)\) from node \(i\) to node \(j\). If both \((i, j) \in E\), and \((j, i) \in E\), we say that there is a bilateral arc between node \(i\) and node \(j\). A directed graph is weighted if a weight \( w_{ij} > 0 \) is associated with an arc \((i, j)\) in \( G \).

The binary adjacency relations between pairs of nodes can be conveniently represented by a \( N \)-square (not symmetric) matrix \( A^{[\alpha]} \), which is called binary adjacency matrix, whose entries are \( a_{ij}^{[\alpha]} = 1 \) if the \((i, j) \in E\), 0 otherwise. The entry \( a_{ij}^{[\alpha]} \in A^{[\alpha]} \) represents the arc outgoing from \(i\) and incoming to \(j\). Then, the 1’s on the row \(i\) are the arcs that leave the node \(i\). A weighted directed network is completely described by the real \( N \)-square matrix \( W^{[\alpha]} \), the weighted adjacency matrix, whose entries \( w_{ij}^{[\alpha]} \) are different from zero if there is a weighted arc \((i, j) \in E\), and \( w_{ij}^{[\alpha]} = 0 \) otherwise.

\(^1\)We use the notation \( A^{[\alpha]} \) for the adjacency matrix to be consistent with the general notation of the supradjacency matrix defined later.
In this paper we refer to multiplex networks that are node-aligned, with non-diagonal couplings (see Kivelä et al. [2014] for the taxonomy of the multidimensional networks). Specifically, a multiplex network consists of a family of networks $G_\alpha = (V_\alpha, E_\alpha)$, $\alpha = 1, ..., L$, where each network $G_\alpha = (V_\alpha, E_\alpha)$ is located in a layer $\alpha$ and a node $i \in V_\alpha$ is adjacent to $j \in V_\beta$, $\forall \alpha, \beta = 1, ..., L$ if there is an arc connecting them. Note that, in node-aligned networks, all nodes are shared between all layers, namely, $V_\alpha = V_\beta = V$, $\forall \alpha, \beta = 1, ..., L$, and that $E_\alpha$ collects all the arcs connecting nodes on layer $\alpha$ to nodes within the same layer (intra-layer connections) and to nodes on different layers (inter-layer connections).

We assume that the network is non-diagonal coupled, that is inter-layer arcs may exist not only between nodes and their counterparts, but links between a node $i$ in a given layer and a node $j \neq i$ in a different layer are allowed. From now on, we will refer to this kind of networks briefly as multiplex networks (or simply multiplex). A weight $w_{ij}^{[\alpha\beta]} > 0$ is associated with an arc $(i, j)$ in $E_{\alpha\beta}$, with $\alpha, \beta = 1, \ldots, L$. Note that when $\alpha = \beta$ we intend that there is a weighted arc $w_{ij}^{[\alpha]} > 0$ between nodes $i$ and $j$ in the network $G_\alpha$.

The adjacency relations between pair of nodes can be conveniently represented by the supradjacency matrix. It is defined as a matrix, with $L \times L$ square blocks, each one of order $N$:

$$
W = \begin{bmatrix}
W^{[1]} & W^{[12]} & \ldots & W^{[1L]} \\
W^{[21]} & W^{[2]} & \ldots & W^{[2L]} \\
\vdots & \vdots & \ddots & \vdots \\
W^{[L1]} & W^{[L2]} & \ldots & W^{[L]}
\end{bmatrix}
$$

(1)

where the diagonal blocks represent the weighted adjacency matrix of each layer $W^{[\alpha]}$, $\alpha = 1, ..., L$, whereas the out of diagonal blocks $W^{[\alpha\beta]}$ represent the weighted adjacency relations between nodes on layers $\alpha$ and nodes on layer $\beta$. We indicate its unweighted version by $A$.

We denote the generic element of $W$ as $w_{hk}$ with $h, k = 1, ..., NL$, where

$$
h = N(\alpha - 1) + i, k = N(\beta - 1) + j.
$$

(2)
Notice that the indices \( h, k \) identify the position in the supradjacency matrix \( W \) of the weight of the arc \((i, j) \in E_{\alpha, \beta}\) (i.e. \( w_{ij}^{[\alpha\beta]} = w_{hk} \)). Thus, from now on, we always assume this relation between \( h, k \) and \( \alpha, \beta, i, j \).

The in-degree \( d_{i,\text{in}}^{[\alpha]} \) of a node \( i \) on layer \( \alpha \) is the number of arcs pointing towards \( i \) from any layer. This degree can be expressed as follows:

\[
d_{i,\text{in}}^{[\alpha]} = (A^T \mathbf{1})_h
\]

where \( \mathbf{1} \) is the \( NL \)-vector of ones. Definition (3) represents the \( h \)-th component of the in-degree vector, with \( h \) as in (2). By the definition of the matrix \( A \), \( d_{i,\text{in}}^{[\alpha]} \) is obtained as the sum of the in-degree related to arcs in the same layer \( \alpha \) and the in-degrees related to incoming arcs from other layers \( \beta \).

Similarly, the out-degree is:

\[
d_{i,\text{out}}^{[\alpha]} = (A \mathbf{1})_h. \tag{4}
\]

The degree \( d_{i}^{[\alpha]} \) of a vertex \( i \) in the layer \( \alpha \) is then:

\[
d_{i}^{[\alpha]} = d_{i,\text{in}}^{[\alpha]} + d_{i,\text{out}}^{[\alpha]} = ([A^T + A] \mathbf{1})_h. \tag{5}
\]

Bilateral arcs between the node \( i \) and its adjacent nodes, if any, are represented as:

\[
d_{i,\leftrightarrow}^{[\alpha]} = (A^2)_{hh}. \tag{6}
\]

The in-degree of a node \( i \) with respect to all layers is defined as:

\[
d_{i,\text{in}} = \sum_{\alpha=1}^{L} d_{i,\text{in}}^{[\alpha]}. \tag{7}
\]

Out-degree \( d_{i,\text{out}} \) with respect to all layers is defined similarly. The total degree on the multilayer is then \( d_i = d_{i,\text{in}} + d_{i,\text{out}} \).

Moving to the weighted case, the previous definitions can be replaced by the strength of a node \( i \):

\[
s_{i,\text{in}}^{[\alpha]} = (W^T \mathbf{1})_h \tag{8}
\]

\[
s_{i,\text{out}}^{[\alpha]} = (W \mathbf{1})_h \tag{9}
\]
The strength $s_i^{[\alpha]}$ of a vertex $i$ in the layer $\alpha$ is then:

$$s_i^{[\alpha]} = s_i^{[\alpha],in} + s_i^{[\alpha],out} = [(W^T + W)1]_i.$$

(10)

The in-strength of a node $i$ with respect to all layers is defined as:

$$s_{i,in} = \sum_{\alpha=1}^{L} s_i^{[\alpha],in}.$$  

(11)

Out-strength $s_{i,out}$ with respect to all layers is defined similarly. The total strength
of $i$ is then $s_i = s_{i,in} + s_{i,out}$.

We define the strength related to bilateral arcs between the node $i$ and its adjacent nodes as:

$$s_i^{[\alpha],\leftrightarrow} = (WA + AW)_{hh} \frac{1}{2}.$$  

(12)

Formula (12) extends formula (6) to the weighted case, multiplying each bilateral link by the arithmetic mean of its weights. With this choice, we are assuming that the bilateral strength sums the average weight of each bilateral arc.

3. Triangles in directed multilayer networks

To formally provide a definition of clustering coefficients in multiplex context, it is worth at first to introduce the definition of triangle.

A triangle in a multiplex is a closed triplet $i, j, k$ such that the nodes can belong to up to three different levels so that they are connected by inter or intra-layer arcs, independently of their orientation. By this definition, we include all possible closed triplets, moving in all directions, along inter or intra-layer arcs. This definition extends to the directed case the one adopted in Bartesaghi et al. (2021) for multiplex weighted networks.

The respective number of triangles can be calculated by counting the accordingly oriented 3-walks around the node of interest. Similarly, in a weighted directed

---

2 Other choices are possible but could be not equally effectiveness in computation.

3 Unless the undirected case, in direct multiplex networks we can identify four types of triangles around a given node $i$ on layer $\alpha$ according to the orientation of the arcs (named out, in, cycle and middleman)
multiplex network a weight can be given to a triangle for instance by simply multiplying the weights of its three arcs. This is not a univocal choice, and other ways to attribute weights have been proposed in the literature, giving rise to different clustering coefficients in this context.

The total number $T_{i}^{[\alpha]}$ of real triangles around the node $i$ on layer $\alpha$ is provided by the following expression in terms of supradjacency matrix $A$:

$$T_{i}^{[\alpha]} = \frac{1}{2}((A + A^{T})^{3})_{hh}$$

(13)

Formula (13) allows to extend the definitions on each layer and on the whole network. Indeed, the number of triangles around the node $i$ on any layer is then $T_{i} = \sum_{\alpha=1}^{L} T_{i}^{[\alpha]}$, and the total number of triangles in a single layer $\alpha$ is $T^{[\alpha]} = \sum_{i=1}^{N} T_{i}^{[\alpha]}$, finally the number of total triangles in the network is $T = \sum_{i=1}^{N} \sum_{\alpha=1}^{L} T_{i}^{[\alpha]}$.

Figure 1 illustrates what has been introduced so far on a simple example. Four nodes on two distinct levels are connected by oriented arcs. For the sake of simplicity, there are no bilateral arcs and all the weights are set equal to 1.

![Figure 1: Example of a binary directed multiplex network with 4 nodes and 2 layers](image)

We can observe that, for instance, on layer 1 node 3 belongs to 3 directed triangles, whereas on layer 2 node 3 belongs to only 2 ones. Similarly, we can compute the number of actual and potential triangles for each node on each level.
4. Clustering coefficients in weighted directed multiplex networks

In this section we provide the mathematical definitions of clustering coefficients for multiplex directed networks. To do this, we have at first to decide under which perspective we want to consider triangles around \( i \). This is the reason why, in this framework, we can define four different types of clustering coefficients, namely three coefficients, in dependence on which node and/or layer is taken into account, and a global one, for the whole network. Moreover, in a weighted and directed multiplex network, triangles can be weighted in more than one way, giving rise to more than one possible definition of weighted clustering coefficient.

A first possibility is to take into account all directed triangles that a node \( i \) actually forms with its neighbours, weighted with the average weight of the links connecting a node \( i \) to its adjacent nodes \( j \) and \( k \). This is in line with the coefficient proposed in Clemente & Grassi (2018) for monoplex networks. Hence, we define the coefficient \( C(i, \alpha) \) for a node \( i \) and level \( \alpha \) as:

\[
C(i, \alpha) = \frac{\left[ (W + W^T)(A + A^T)^2 \right]_{hh}^{\alpha}}{2 \left[ d^{\alpha}_i (d^{\alpha}_i - 1) - 2s^{\alpha}_i s^{\alpha}_i,\leftrightarrow \right]^{\alpha}}.
\]

(14)

where \( W \) is defined as in (1) and \( A \) accordingly, \( d^{\alpha}_i, s^{\alpha}_i \) are defined as in formulas (5), (10) and (12) and \( h \) is the position index defined in (2).

In particular, the numerator extends to the weighted case formula (13) that counts the number of directed triangles \( T^{\alpha}_i \). The denominator provides all possible (weighted) directed triangles that \( i \) could form in or outside the layer \( \alpha \).

An alternative way to weight the triangles is to use the geometric mean of the weights of the arcs. This choice is in line with the coefficient proposed by Fagiolo (2007) for monoplex networks. Hence, we define the coefficient, denoted by \( \tilde{C}(i, \alpha) \), as:

\[
\tilde{C}(i, \alpha) = \frac{\left( \hat{W} + \hat{W}^T \right)^3_{hh}}{2 \left[ d^{\alpha}_i (d^{\alpha}_i - 1) - 2d^{\alpha}_i,\leftrightarrow \right]^{\alpha}}.
\]

(15)

where \( \hat{W} \) is the matrix whose entries are \( \hat{w}^{\alpha,\beta}_{ij} = \left( \frac{w^{[\alpha,\beta]}_{ij}}{\max(w_{ij}^{[\alpha,\beta]})} \right)^{\frac{1}{3}} \) \( \forall i, j, \alpha, \beta \), being \( w_{ij}^{[\alpha,\beta]} \) defined as in Section 2. Elements \( d^{\alpha}_i, d^{\alpha}_i,\leftrightarrow \) are defined as in formulas (5) and (6) and \( h \) is the position index defined as in (2).
It is noticeable that, in (14) and (15), bilateral arcs $2s_{i,\alpha}^{[\alpha]}$ and $2d_{i,\alpha}^{[\alpha]}$ have to be removed as they represent “false” triangles, being formed by $i$ and by a pair of directed arcs pointing to the same node, e.g., $i \to j$ and $j \to i$. In directed framework, a node $i$ can form up to two triangles with each pair, also including two “false” potential triangles for each bilateral link.

A third possibility to assign a weight to actual and potential triangles can be found in the recent literature (see De Domenico et al. (2013) and Jia et al. (2021)). It consists in taking the product of the weights of the three symmetrized arcs, assuming that the adjacency matrix is normalized as for the coefficient in formula (15). According to this choice, the clustering coefficient can be defined as

$$\tilde{C}(i,\alpha) = \frac{\left(\tilde{W} + \tilde{W}^T\right)_{hh}^3}{s_{i}^{[\alpha]} - \sum_{k \neq h} (w_{hk} + w_{kh})^2} \quad (16)$$

where $\tilde{W}$ is the normalized matrix whose entries are $\tilde{w}_{ij}^{[\alpha\beta]} = \frac{w_{ij}^{[\alpha\beta]}}{\max(w_{ij}^{[\alpha\beta]})} \forall i, j, \alpha, \beta$, and $h, k$ are the position indices defined as in (2). The denominator in formula (16) accounts for the weights of all the connected triads centred in the node $i$ on level $\alpha$ and it can be considered the weight of all the potential triangles. Let us notice that to the (missing or actual) arc opposite to the node of interest is assigned a weight equal to 1. Of course, $s_{i}^{[\alpha]}$ is again the total strength of node $i$ on level $\alpha$ and it is defined as in formula (10).

Formulas (14), (15) and (16) provide local coefficients, since they refer to a single node in a single layer. It turns out to be useful to define coefficients that account for an intermediate clustering structure of the network, both considering homologous nodes on different layers or all the nodes within a given layer. It is also possible to introduce an average coefficient on the whole network that traces the concept of transitivity, well-known for monolayer networks.

In order to introduce these further generalizations, we refer to formula (14) but quite naturally similar extensions of coefficients (15) and (16) can be provided.

First we can consider a single node on all the layers on which it lies. This involves calculating the total weight of triangles to which these homologous nodes belong and dividing by the potential ones:
\[
C_N(i) = \frac{\sum_{\alpha=1}^{L} \left[ \left( W + W^T \right) (A + A^T)^2 \right]_{hh}}{2 \sum_{\alpha=1}^{L} \left[ s^{[\alpha]}_i (d^{[\alpha]}_i - 1) - 2s^{[\alpha]}_{i,\leftrightarrow} \right]},
\]

where \( h = N(\alpha - 1) + i \) (by formula (2)) and we sum triangles over \( \alpha = 1, \ldots, L \).

It is noteworthy that \( C_N(i) \) can be obtained as a weighted average of the local coefficients \( C(i, \alpha) \) defined in formula (14). Each local coefficient has a weight in the weighted average equal to

\[
\frac{s^{[\alpha]}_i (d^{[\alpha]}_i - 1) - 2s^{[\alpha]}_{i,\leftrightarrow}}{\sum_{\alpha=1}^{L} \left[ s^{[\alpha]}_i (d^{[\alpha]}_i - 1) - 2s^{[\alpha]}_{i,\leftrightarrow} \right]}.
\]

In other words, the more the strength (and the degree) of the node \( i \) in the layer has a high incidence on the total strength (and degree) of the node in all the layers, the more \( C(i, \alpha) \) affects \( C_N(i) \).

Alternatively, we can focus on a single layer and define a coefficient regarding all the nodes in that layer. In particular, these nodes contribute to the coefficient on the basis of all their arcs, within the layer or outside:

\[
C_L(\alpha) = \frac{\sum_{i=1}^{N} \left[ \left( W + W^T \right) (A + A^T)^2 \right]_{hh}}{2 \sum_{i=1}^{N} \left[ s^{[\alpha]}_i (d^{[\alpha]}_i - 1) - 2s^{[\alpha]}_{i,\leftrightarrow} \right]},
\]

where, again, \( h = N(\alpha - 1) + i \) and the sums are over \( i = 1, \ldots, N \).

Also in this case, the clustering coefficient for a layer can be written as a weighted average of the local coefficients \( C(i, \alpha) \). The incidence of \( C(i, \alpha) \) on \( C_L(\alpha) \) depends on the relation between the strength and the degree of the node \( i \) in the layer \( \alpha \) and the total strength and degree of all nodes in the layer.

Finally, we can take into account all nodes on all levels, with all their actual or potential triangles:

\[
C = \frac{\sum_{h=1}^{N_L} \left[ \left( W + W^T \right) (A + A^T)^2 \right]_{hh}}{2 \sum_{i=1}^{N} \sum_{\alpha=1}^{L} \left[ s^{[\alpha]}_i (d^{[\alpha]}_i - 1) - 2s^{[\alpha]}_{i,\leftrightarrow} \right]}.
\]

Note that formula (19) extends the definition of transitivity for the entire network to the directed multiplex case.
5. Computational experiment

5.1. Dataset, multiplex network and preliminary analyses

To test the effectiveness of the proposed coefficients, we run a computational experiment applying them to real-world data. We refer to the international trade data collected in the World Input-Output Database (WIOD) 2016 Release (Timmer et al. (2015)). WIOD consists of a series of databases and covers 28 EU countries and 15 other major countries. Including the biggest countries in the world, this set covers more than 85 per cent of world GDP (see, e.g., Timmer et al. (2015)). However, to complete the data and make them suitable for various modeling purposes, it is also considered a region called the Rest of the World (RoW) that proxies for all other countries in the world. In Table A.1 we display the list of countries that are included in our analysis. The WIOD contains annual time-series of world input–output tables covering the period from 2000 to 2014. Data consider bilateral trades of exports, expressed in millions of U.S. dollars, for each couple of origin and destination countries and distinct between 56 sectors (see Table A.2 for a list of sectors). Sectors are classified according to the International Standard Industrial Classification (revision 4) (see Department of Economic and Social Affairs (2008)). In this analysis, we focus on the data of last available year, namely, 2014.

By these data we construct a multiplex weighted directed network, where each layer is represented by a sector. In each layer, a node is a country and a directed weighted link represents the total amount of incoming or out-coming flows between a couple of countries in the same sector. In the same layer, we exclude self-loops, i.e. trans-action of a country in the same sector. Inter-layer arcs are instead weighted with the amount of imports or exports between (the same or different) countries in different sectors. We obtain a multiplex network with 44 nodes and 56 sectors, represented by an asymmetric and weighted supradjacency matrix with 2464 rows and columns.

To give a first idea of the structure of the network, we display in Figure 2 the density of each layer and the average inter-layer density. The average inter-layer density of a sector has been computed as the average value of the densities of the graphs characterised by the connections between the sector and each other layer. It is noticeable how several sectors show a high internal density in the layer, while a lower inter-layer density is observed on average. Sector “T” is instead very sparse and tends to be more connected with other layers. Sector “U” is completely isolated
having neither internal connections nor connections with other layers. Therefore, we decided to remove this layer from the network and to deal with 55 sectors.

![Intra-layer and inter-layer density of the multilayer network for each sector](image)

**Figure 2:** Intra-layer and inter-layer density of the multilayer network for each sector

We report in Figure 3 the total strength for each combination of sector and country. The total strength has been obtained as the sum of in and out strength of the country in the layer. It is interesting to note how manufacturing sectors (from “C10” to “C33” in the plot) are characterised by the highest volumes. In terms of country, China, USA and ROW show the highest strengths. On the one hand, China tends to be the dominant country in several sectors but shows also a strength equal to zero in specific sectors (as “C33”, “G45”, “J58”, etc.). On the other hand, USA and ROW have average strength a bit lower than China but these countries show connections in all sectors.

A specific focus on countries’ strength is made in Figure 4. In particular, we distinguish between in and out strength and between strength related to inter-layer and intra-layer connections. A high volume of trade due to inter-layer connections is observed. In particular, USA, AUS and CHN show the highest exposition towards inter-layer arcs. We have values lower than 3% for the ratio between the strength related to arcs in the same layer and the strength due to arcs that connect different layers. Smaller countries (as MLT, LUX, TWN, SVK, HUN) are instead more concentrated on intra-layer flows.
According to imports and exports, NOR, RUS, NLD are characterised by a positive trade balance with a ratio between in and out strength around 80%. Vice versa MLT, HUN and MEX have the opposite behaviour with the highest ratios (more than 110%).

In Figure 3, we focus on sectors’ strength. It is noticeable a very high in-strength for the sector “F” (Constructions), due in particular to directed connections with some manufacturing sectors. Important out-flows are instead observed for the sector “B” (Mining and quarrying), probably due to the use of these materials in other sectors. Finally, we observe how the sector “C26” (Manufacture of computer, electronic and optical products) is characterised by the highest intra-layer transactions.
Figure 4: “In” and “out” strength for each country. Strength is also differentiated between intra-layer and inter-layer connections because of important trades between countries.

Figure 5: Intra-layer strength for each sector (i.e. transactions in the same sectors). Inter-layer “in” and “out” strength for each sector; in this case, we consider directed flows between different sectors.
5.2. Main Results

Alternative coefficients described in Section 4 have been tested on the whole multiplex network. We start by comparing the local coefficients provided by formulas (14), (15) and (16). To this end, we display in Figure 6 for each sector the ranking based on the local clustering coefficients. It allows to emphasize which countries appear prominent in terms of interconnections in each sector and at the same time, we can appreciate different patterns between coefficients. We notice that coefficients $\hat{C}(i,a)$ and $\tilde{C}(i,a)$ tend to provide similar rankings, while coefficients $C(i,a)$ behave in a different way. However, it has to be stressed that values of $C(i,a)$ are very close to one and hence the ranking is in this case not so interesting because very limited differences are observed between countries. Main justification is related to the fact that this coefficient is more affected by the number of triangles than by the weights. Therefore, given the high density of the network, very few differences can be noticed between countries. Focusing on rankings given by $\hat{C}(i,a)$ and $\tilde{C}(i,a)$, it is noticeable how CHN, USA and ROW appear highly interconnected. In particular, CHN have the highest ranking in the largest number of sectors (27 and 38 sectors according to $\hat{C}(i,a)$ and $\tilde{C}(i,a)$, respectively), but USA and ROW are well connected in all sectors. We have indeed that these two countries belong to the top quartile of the clustering distribution in all sectors. We have instead that CHN shows a very low ranking in specific sectors where it is not well represented. In terms of average ranking, these countries are then followed by FRA, GBR, JAP, DEU. In particular, FRA and GBR tend to be well clustered in several sectors (especially considering the coefficient $\hat{C}(i,a)$), while JAP and DEU are instead not covered in some specific sectors. In particular, FRA has the highest ranking in sector “C33” (Air transport). France is indeed one of the largest and most symbolically important aviation markets in Europe, not only in terms of traffic due to its size and geographic location, but also because it is home to some of the industry’s flagship names. At a lower average ranking we find ITA, RUS, KOR, NLD, ESP, countries that belong to the top quartile in specific sectors but also show a lower clustering in other sectors. It is also noteworthy how Denmark appears as a top country in sector “T”, Activities of households as employers; undifferentiated goods producing activities of households for own use, and Taiwan is highly interconnected in sector “C26” Manufacture of computer, electronic and optical products. Electronic component manufacturing is indeed a pillar of Taiwan’s economy, and its role is increasing over time also thanks to the development of technology.
Figure 6: Ranking of clustering coefficients $\hat{C}(i,a)$, $C(i,a)$ and $\tilde{C}(i,a)$ for each sector. Darker red means higher ranking.
To better emphasize the differences between the alternative clustering coefficients, we display in Figure 7 the countries’ ranking computed on the three multiplex coefficients. To obtain these rankings, we considered the coefficients of the single node $i$ over all the layers (defined by formula (17)). The comparison is also extended to the application of the local clustering coefficients for monoplex networks. In particular, we compute on each layer a coefficient $\hat{C}_i$, based on the formula provided by Fagiolo (2007) and then we average the results between the layers. Similarly, the same procedure has been applied in order to obtain the coefficient $C_i$ based on the formula provided in Clemente & Grassi (2018). Therefore, inter-layer connections are not considered in the computation of $\hat{C}_i$ and $C_i$.

It is interesting to note that the inclusion of inter-layer effects provides significant differences in the ranking. This pattern is also justified by the fact that the inter-layer connections are relevant, as shown by the behaviour of density and strength in the previous section. The coefficients based on monoplex networks show indeed a positive rank correlation but far from one, with the results based on multiplex formulas. Finally, we notice a higher level of similarity between $\hat{C}_N(i)$ and $\hat{C}_i$ than between $C_N(i)$ and $C_i$. 
Figure 7: Comparison of rankings of clustering coefficients $\hat{C}_N(i)$, $C_N(i)$ and $\tilde{C}_N(i)$ computed for each country $i$. The comparison is also extended to the average clustering based on monoplex coefficients. In particular $\tilde{C}_i$ and $C_i$ are the average values of the clustering coefficients of each country based on formulas provided in [Fagiolo (2007)] and [Clemente & Grassi (2018)] for weighted and directed networks. Darker red means higher ranking.

In Figure 8, attention has been also paid to the level of interconnection of sectors. Also in this case, the average multiplex coefficients has been compared with the monoplex versions. We notice again a positive correlation between them but with some significant differences. Focusing on specific sectors, we notice a higher level of interconnections in manufacturing sectors. In particular, sectors related to chemical, mineral, metal products and machinery and equipments (as “C20”, “C23”, “C25”, “C28”) show the highest coefficients. Other relevant sectors are wholetrade (“G46”) and constructions (“F”).
6. Conclusions

Clustering coefficient has gained increasing attention in network theory. Although several proposals have been provided for the case of monoplex networks, the assessment of local and global clustering coefficients for multiplex networks deserves a specific focus.

In this paper, we focus on clustering coefficients for a weighted and directed multiplex node-aligned networks. The proposed coefficients generalize the alternative coefficients already provided in the literature for directed monoplex networks and extend existing coefficients for undirected multilayer networks.

The approach has been tested using the data taken from the World Input-Output Database. Results show how the coefficients are able to grasp the degree of interconnection at different levels, emphasizing the role of countries and sectors in the whole trade taking into account both inter-layer and intra-layer flows.
References

Bang-Jensen, J., & Gutin, G. Z. (2008). *Digraphs: theory, algorithms and applications*. Springer Science & Business Media.

Barigozzi, M., Fagiolo, G., & Garlaschelli, D. (2010). Multinetwork of international trade: A commodity-specific analysis. *Physical Review E, 81*, 046104.

Bartesaghi, P., Clemente, G. P., & Grassi, R. (2021). Clustering Coefficients in Weighted Undirected Multilayer Networks. *arXiv preprint arXiv:2105.14325*, .

Battiston, F., Nicosia, V., & Latora, V. (2014). Structural measures for multiplex networks. *Physical Review E, 89*, 032804.

Berlingerio, M., Coscia, M., Giannotti, F., Monreale, A., & Pedreschi, D. (2013). Multidimensional networks: foundations of structural analysis. *World Wide Web, 16*, 567–593.

Boccaletti, S., Bianconi, G., Criado, R., Del Genio, C. I., Gómez-Gardenes, J., Romance, M., Sendina-Nadal, I., Wang, Z., & Zanin, M. (2014). The structure and dynamics of multilayer networks. *Physica Reports, 544*, 1–122.

Chen, Z.-Y., Fan, Z.-P., & Sun, M. (2021). Tensorial graph learning for link prediction in generalized heterogeneous networks. *European Journal of Operational Research, 290*, 219–234.

Cingolani, I., Panzarasa, P., & Tajoli, L. (2017). Countries’ positions in the international global value networks: Centrality and economic performance. *Applied Network Science, 2*.

Clemente, G. P., & Grassi, R. (2018). Directed clustering in weighted networks: a new perspective. *Chaos, Solitons & Fractals, 107*.

Cozzo, E., Kivelä, M., De Domenico, M., Solé-Ribalta, A., Arenas, A., Gómez, S., Porter, M. A., & Moreno, Y. (2015). Structure of triadic relations in multiplex networks. *New Journal of Physics, 17*, 073029.

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*, 1–20. URL: https://doi.org/10.1371/journal.pone.0070726 doi: 10.1371/journal.pone.0070726
De Domenico, M., Solé-Ribalta, A., Cozzo, E., Kivelä, M., Moreno, Y., Porter, M. A., Gómez, S., & Arenas, A. (2013). Mathematical formulation of multilayer networks. *Physical Review X, 3*, 041022.

Department of Economic and Social Affairs (2008). *International Standard Industrial Classification of All Economic Activities*. Technical Report United Nations.

Ertem, Z., Veremyev, A., & Butenko, S. (2016). Detecting large cohesive subgroups with high clustering coefficients in social networks. *Social Networks, 46*, 1–10.

Fagiolo, G. (2007). Clustering in complex directed networks. *Physical Review E, 76*. doi:10.1103/physreve.76.026107.

Fagiolo, G., Reyes, J., & Schiavo, S. (2008). On the topological properties of the world trade web: A weighted network analysis. *Physica A: Statistical Mechanics and its Applications, 387*, 3868–3873.

Fagiolo, G., Squartini, T., & Garlaschelli, D. (2012). Null models of economic networks: the case of the world trade web. *Journal of Economic Interaction and Coordination, 8*, 75–107.

Gupta, M., & Kumar, P. (2020). Recommendation generation using personalized weight of meta-paths in heterogeneous information networks. *European Journal of Operational Research, 284*, 660–674.

Harary, F. (1969). *Graph theory*. Addison-Wesley, Reading, MA.

Jia, M., Gabrys, B., & Musial, K. (2021). Directed closure coefficient and its patterns. *PLOS ONE, 16*, 1–23. doi:10.1371/journal.pone.0253822.

Kivelä, M., Arenas, A., Barthelemy, M., Gleeson, J. P., Moreno, Y., & Porter, M. A. (2014). Multilayer networks. *Journal of Complex Networks, 2*, 203–271.

Maluck, J., & Donner, R. V. (2015). A network of networks perspective on global trade. *PLOS ONE, 10*, 1–24.

Mastrandrea, R., Squartini, T., Fagiolo, G., & Garlaschelli, D. (2014). Reconstructing the world trade multiplex: The role of intensive and extensive biases. *Phys. Rev. E, 90*, 062804.
Menichetti, G., Remondini, D., Panzarasa, P., Mondragón, R. J., & Bianconi, G. (2014). Weighted multiplex networks. *PLOS ONE*, 9, 1–8.

Newman, M. E. J. (2003). The structure and function of complex networks. *SIAM Review*, 6, 28384. doi:10.1038/srep28384

Piccardi, C., & Tajoli, L. (2012). Existence and significance of communities in the world trade web. *Physical Review E*, 85, 066119.

Piccardi, C., & Tajoli, L. (2018). Complexity, centralization, and fragility in economic networks. *PLOS ONE*, 13, 1–13.

Ren, Z.-M., Zeng, A., & Zhang, Y.-C. (2020). Bridging nestedness and economic complexity in multilayer world trade networks. *Humanities and Social Sciences Communications*, 7, 1–8.

Saracco, F., Di Clemente, R., & Gabrielli, A. (2015). Randomizing bipartite networks: the case of the world trade web. *Scientific Report*, 5.

Timmer, M., Dietzenbacher, E., Los, B., Stehrer, R., & de Vries, G. J. (2015). An illustrated user guide to the world input–output database: the case of global automotive production. *Review of International Economics*, 23, 575–605.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of small-world networks. *Nature*, 393, 440–442. doi:10.1038/30918
## Appendix A. Lists of sectors and countries

| Number | Name             | Code | Number | Name            | Code |
|--------|------------------|------|--------|-----------------|------|
| 1      | Australia        | AUS  | 2      | Austria         | AUT  |
| 3      | Belgium          | BEL  | 4      | Bulgaria        | BGR  |
| 5      | Brazil           | BRA  | 6      | Canada          | CAN  |
| 7      | Switzerland      | CHE  | 8      | China           | CHN  |
| 9      | Cyprus           | CYP  | 10     | Czech Republic  | CZE  |
| 11     | Germany          | DEU  | 12     | Denmark         | DNK  |
| 13     | Spain            | ESP  | 14     | Estonia         | EST  |
| 15     | Finland          | FIN  | 16     | France          | FRA  |
| 17     | United Kingdom   | GBR  | 18     | Greece          | GRC  |
| 19     | Croatia          | HRV  | 20     | Hungary         | HUN  |
| 21     | Indonesia        | IDN  | 22     | India           | IND  |
| 23     | Ireland          | IRL  | 24     | Italy           | ITA  |
| 25     | Japan            | JPN  | 26     | Korea           | KOR  |
| 27     | Lithuania        | LTU  | 28     | Luxembourg      | LUX  |
| 29     | Latvia           | LVA  | 30     | Mexico          | MEX  |
| 31     | Malta            | MLT  | 32     | Netherlands     | NLD  |
| 33     | Norway           | NOR  | 34     | Poland          | POL  |
| 35     | Portugal         | PRT  | 36     | Romania         | ROU  |
| 37     | Russian Federation| RUS | 38     | Slovak Republic | SVK  |
| 39     | Slovenia         | SVN  | 40     | Sweden          | SWE  |
| 41     | Turkey           | TUR  | 42     | Taiwan          | TWN  |
| 43     | United States    | USA  | 44     | Rest of the World | ROW |

Table A.1: Lists of 44 countries and areas in the WIOD table.
| Sector | Description                                                                 | Code  |
|--------|-------------------------------------------------------------------------------|-------|
| 1      | Crop and animal production, hunting and related service activities            | A01   |
| 2      | Forestry and logging                                                          | A02   |
| 3      | Fishing and aquaculture                                                       | A05   |
| 4      | Mining and quarrying                                                          | B     |
| 5      | Manufacture of food products, beverages and tobacco products                  | C10-C12 |
| 6      | Manufacture of textiles, wearing apparel and leather products                 | C13-C15 |
| 7      | Manufacture of wood and of products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials | C16   |
| 8      | Manufacture of paper and paper products                                       | C17   |
| 9      | Printing and reproduction of recorded media                                   | C18   |
| 10     | Manufacture of coke and refined petroleum products                            | C19   |
| 11     | Manufacture of chemicals and chemical products                                | C20   |
| 12     | Manufacture of basic pharmaceutical products and pharmaceutical preparations  | C21   |
| 13     | Manufacture of rubber and plastic products                                    | C22   |
| 14     | Manufacture of other non-metallic mineral products                            | C23   |
| 15     | Manufacture of basic metals                                                   | C24   |
| 16     | Manufacture of fabricated metal products, except machinery and equipment      | C25   |
| 17     | Manufacture of computer, electronic and optical products                      | C26   |
| 18     | Manufacture of electrical equipment                                           | C27   |
| 19     | Manufacture of machinery and equipment n.e.c.                                 | C28   |
| 20     | Manufacture of motor vehicles, trailers and semi-trailers                     | C29   |
| 21     | Manufacture of other transport equipment                                      | C30   |
| 22     | Manufacture of furniture; other manufacturing                                | C31-C32 |
| 23     | Repair and installation of machinery and equipment                            | C33   |
| 24     | Electricity, gas, steam and air conditioning supply                           | D35   |
| 25     | Water collection, treatment and supply                                        | E36   |
| 26     | Sewage, waste collection, treatment and disposal activities; materials recovery; remediation activities and other waste management services | E37-E39 |
| 27     | Construction                                                                  | F     |
| 28     | Wholesale and retail trade and repair of motor vehicles and motorcycles        | G45-G46 |
| 29     | Wholesale trade, except of motor vehicles and motorcycles                     | G46   |
| 30     | Retail trade, except of motor vehicles and motorcycles                        | G47   |
| 31     | Land transport and transport via pipelines                                    | H49   |
| 32     | Water transport                                                               | H50   |
| 33     | Air transport                                                                 | H51   |
| 34     | Warehousing and support activities for transportation                         | H52   |
| 35     | Postal and courier activities                                                 | H53   |
| 36     | Accommodation and food service activities                                     | I     |
| 37     | Publishing activities                                                         | J58   |
| 38     | Motion picture, video and television programme production, sound recording and music publishing activities; programming and broadcasting activities | J59-J60 |
| 39     | Telecommunications                                                            | J61   |
| 40     | Computer programming, consultancy and related activities; information service activities | J62-J63 |
| 41     | Financial service activities, except insurance and pension funding            | K64   |
| 42     | Insurance, reinsurance and pension funding, except compulsory social security  | K65   |
| 43     | Activities auxiliary to financial services and insurance activities           | K66   |
| 44     | Real estate activities                                                        | L68   |
| 45     | Legal and accounting activities; activities of head offices; management consultancy activities | M69-M70 |
| 46     | Architectural and engineering activities; technical testing, and analysis     | M71   |
| 47     | Scientific research and development                                           | M72   |
| 48     | Advertising and market research                                                | M73   |
| 49     | Other professional, scientific and technical activities; veterinary activities | M74-M75 |
| 50     | Administrative and support service activities                                 | N     |
| 51     | Public administration and defence; compulsory social security                 | O84   |
| 52     | Education                                                                     | P95   |
| 53     | Human health and social work activities                                       | Q     |
| 54     | Other service activities                                                      | R8    |
| 55     | Activities of households as employers; undifferentiated goods- and services-producing activities of households for own use | T     |
| 56     | Activities of extraterritorial organizations and bodies                       | U     |

Table A.2: Lists of 56 sectors in the WIOD table