PHYSICAL REVIEW E **89**, 032804 (2014)

Structural measures for multiplex networks

Federico Battiston,[1] Vincenzo Nicosia,[1,2] and Vito Latora[1,2,3]
[1]*School of Mathematical Sciences, Queen Mary University of London, London E1 4NS, United Kingdom*
[2]*Laboratorio sui Sistemi Complessi, Scuola Superiore di Catania, I-95123 Catania, Italy*
[3]*Dipartimento di Fisica ed Astronomia, Università di Catania and INFN, I-95123 Catania, Italy*
(Received 14 August 2013; published 12 March 2014)

Many real-world complex systems consist of a set of elementary units connected by relationships of different kinds. All such systems are better described in terms of multiplex networks, where the links at each layer represent a different type of interaction between the same set of nodes rather than in terms of (single-layer) networks. In this paper we present a general framework to describe and study multiplex networks, whose links are either unweighted or weighted. In particular, we propose a series of measures to characterize the multiplexicity of the systems in terms of (i) basic node and link properties such as the node degree, and the edge overlap and reinforcement, (ii) local properties such as the clustering coefficient and the transitivity, and (iii) global properties related to the navigability of the multiplex across the different layers. The measures we introduce are validated on a genuinely multiplex data set of Indonesian terrorists, where information among 78 individuals are recorded with respect to mutual trust, common operations, exchanged communications, and business relationships.

DOI: 10.1103/PhysRevE.89.032804 PACS number(s): 89.75.Hc, 87.23.Ge

I. INTRODUCTION

Much work has been done over the past few decades to investigate and characterize the structure and dynamics of complex systems. Many of these systems have been proven to be successfully described as a network whose nodes represent the different basic units of the system, and whose links represent the interactions and relationships among the units [1–4]. The standard approach to network description of complex systems consists of studying the graph resulting from the aggregation of all the links observed between a certain set of elementary units. However, such aggregation procedure might in general discard important information about the structure and function of the original system, since in many cases the basic constituents of a system might be connected through a variety of relationships which differ for relevance and meaning [5,6]. For instance, the same set of individuals in a social system can be connected through friendship, collaboration, kinship, communication, commercial, and colocation relationships, just to name some of them, while in complex multimodal transportation systems, which are typical of large metropolitan areas, a set of locations might be reached in several different ways, e.g., using bus, underground, suburban rail, or riverboat networks and the like. In these systems, each type of interaction has associated a given relevance, importance, cost, distance, or meaning, so treating all the links as being equivalent results into losing a lot of important information. A better description of such systems is in terms of *multiplex networks*, i.e., networks where each node appears in a set of different layers, and each layer describes all the edges of a given type.

Recently, a considerable amount of effort has been devoted to the characterization and modeling of multiplex networks, with the aim of creating a consistent mathematical framework to study, understand, and reproduce the structure of these systems. A number of measures have been proposed in the context of real-world multiplex networks such as air transportation systems [7] and massive multiplayer online games [8]. Some other works are pointing towards a statistical mechanics formulation of multiplex networks [9], to the extension of classical network metrics to the case of multiplexes [10,11], and to model the growth of systems of this kind [12]. Finally, another active research direction is that of characterizing the dynamics and the emergent properties of multilayer systems, especially with respect to epidemic [13] and information spreading [14,15], cooperation [16], synchronization [17], diffusion processes [18], and random walks [19]. A review of recent papers in this field can be found in Ref. [20].

In this article we introduce a set of basic measures to characterize the structural properties of multiplex networks, including their degree distributions, edge overlap, node clustering, configuration of shortest paths and spectral centrality. In particular, we focus on the quantification of the participation of single nodes to the structure of each layer and of the importance of each node for the overall efficiency of the multiplex network, in terms of node reachability and triadic closure. All the proposed measures are tested and validated on a genuinely multiplex real-world data set, the Noordin Top terrorist network, which includes detailed information about four different features, namely mutual trust, common operations, exchanged communications, and business involvement of 78 Indonesian terrorists. Thanks to its peculiar structure, this system can be naturally modeled as a four-layer multiplex. We show that, in this particular data set, one of the four layers, namely the trust layer, acts as a driver for the others, since the conditional probability for two terrorists to communicate or to participate to the same operation clearly depends on the strength of their mutual trust relationship. This result can be explained in terms of social reinforcement and reveals important details about the overall dynamics of edge formation and strengthening in the multiplex. We believe that the measures proposed hereby will have wide applicability to larger multiplexes in several different domains.

The article is organized in the following way. In Sec. II we discuss the various levels at which we can describe a multiplex network. We introduce the aggregated topological matrix, the overlapping and the weighted overlapping matrix, and the vector of adjacency matrices $\boldsymbol{A}$, which provides a complete description of the multiplex network. We also discuss basic metrics, including node degree and edge overlap. In Sec. III we introduce the multi-layer system under study, a multiplex network with four layers describing the interactions among 78 Indonesian terrorists. In Sec. IV we compare the different measures of node degree on the network under study and we introduce metrics to describe how the links of a node are distributed over the various layers. In Sec. V we quantify the edge overlap and we discuss a mechanism of social reinforcement present in the network of terrorists. In Sec. VI we generalize the concepts of clustering and transitivity to the case of multiplex networks, considering the possibility of triangles with links in different layers. In Sec. VII we investigate the number of shortest paths which make use of links in different layers. In Sec. VIII we propose a simple extension of spectral centrality to networks with multiple layers. Finally, in Sec. IX we present our conclusions.

II. GENERAL FORMALISM

Consider a complex system involving multiple kinds of relations among its basic units. When it is possible to distinguish the nature of the ties, an effective approach to describe the system consists in embedding the edges in different layers according to their type. This is the starting point of multiplex networks analysis.

In this section we propose a comprehensive approach and a coherent notation for the study of systems composed of N nodes and M layers, with ties in each layer being undirected and either unweighted or weighted [21]. Our framework does not fit, instead, the case of a multiplex of multiplexes, i.e., a system in which each layer is composed by a number of sublayers (which may in turn be composed by several subsublayers, and so on), but might be easily extended to encompass this case.

We consider, first, a system composed of N nodes and M unweighted layers, and we extend the notation to the case of weighted layers afterwards. We can associate to each layer $\alpha,\alpha = 1,\ldots,M$, an adjacency matrix $A^{[\alpha]} = \{a_{ij}^{[\alpha]}\}$, where $a_{ij}^{[\alpha]} = 1$ if node i and node j are connected through a link on layer α, so each of the M layers is an unweighted network. Such a multiplex system is completely specified by the vector of the adjacency matrices of the M layers,

$$\boldsymbol{A} = \{A^{[1]},\ldots,A^{[M]}\}. \tag{1}$$

We define the degree of a node i on a given layer as $k_i^{[\alpha]} = \sum_j a_{ij}^{[\alpha]}$, from which follows that $0 \leqslant k_i^{[\alpha]} \leqslant N-1\ \forall i,\forall\alpha$. Consequently, the degree of node i in a multiplex network is the vector

$$\boldsymbol{k}_i = \{k_i^{[1]},\ldots,k_i^{[M]}\}, \quad i = 1,\ldots,N. \tag{2}$$

We have $\sum_i k_i^{[\alpha]} = 2K^{[\alpha]}$, where $K^{[\alpha]}$ is the total number of links on layer α. As for single-layer networks, we use lowercase letters to denote node properties and capital letters for properties obtained by summing over the nodes or the edges, either at the level of single layer or at the level of the whole system.

Vectorial variables, such as $\boldsymbol{A}$ and $\boldsymbol{k}_i$, are necessary to properly store all the richness of multiplex networks. However, it is also useful to define aggregated adjacency matrices (in which we disregard the fact that the links belong to different layers) to be used as a term of comparison. As part of the goal of this paper, we will show that aggregated matrices and the corresponding aggregated measures with which one may be tempted to analyze the multilayer structure have limited potential and often fail in detecting the key structural features of a multiplex network. We define the aggregated topological adjacency matrix $\mathcal{A} = \{a_{ij}\}$ of a multiplex network, where

$$a_{ij} = \begin{cases} 1 & \text{if } \exists\,\alpha : a_{ij}^{[\alpha]} = 1 \\ 0 & \text{otherwise} \end{cases}. \tag{3}$$

This is the adjacency matrix of the unweighted network obtained from the multilayer structure joining all pairs of nodes i and j which are connected by an edge in at least one layer of the multiplex network and neglecting both the possible existence of multiple ties between a pair of nodes and the nature of each tie. For the degree of node i on the aggregated topological network, we have

$$k_i = \sum_j a_{ij}. \tag{4}$$

Summing k_i over all elements of the system, we obtain

$$\sum_i k_i = 2K, \tag{5}$$

where K is the total number of links (also called the size) of the aggregated topological network. Matrix $\mathcal{A}$ describes a single-layer binary network which can be studied using the well established set of measures defined for single-layer networks. As we will show in the following sections, this representation turns out to be very simplistic and often insufficient to unveil the key features of multilayer systems. A basic feature which is lost in the topological aggregated matrix is that in multiplex systems the same pair of nodes can be connected by ties of different kinds.

We introduce the edge overlap of edge i-j between two layers α and α' as

$$o_{ij}^{[\alpha,\alpha']} = a_{ij}^{[\alpha]} + a_{ij}^{[\alpha']} \tag{6}$$

and the edge overlap of edge i-j as

$$o_{ij} = \sum_\alpha a_{ij}^{[\alpha]}. \tag{7}$$

By definition, we have $0 \leqslant o_{ij} \leqslant M\ \ \forall i,j$. We can now define the aggregated overlapping adjacency matrix $\mathcal{O} = \{o_{ij}\}$. Matrix $\mathcal{O}$ does not differ from a standard weighted adjacency matrix of a single-layer network. Even though the overlapping matrix $\mathcal{O}$ has a richer structure compared to the purely topological matrix $\mathcal{A}$, in this paper we show that also this matrix eventually fails in featuring a number of basic structural properties of multiplex networks. In fact, although the information about the total number of connections (at different layers) between each pair of nodes is preserved, the loss of knowledge about the nature of each tie (which is instead preserved in a vectorial variable such as $\boldsymbol{A}$) will often make

$\mathcal{O}$ insufficient to catch important characteristics of multilayer systems.

Based on the edge overlap o_{ij}, we can also define the overlapping degree of node i as follows:

$$o_i = \sum_j o_{ij} = \sum_\alpha k_i^{[\alpha]}. \tag{8}$$

with $o_i \geqslant k_i$. Slightly different measures of node overlapping were defined in Ref. [9]. Notice that the overlapping degree o_i represents the correct factor to normalize the components of the degree vector $\boldsymbol{k}_i$. In fact, we have $(1/o_i)\sum_\alpha k_i^{[\alpha]} = 1$. Summing o_i over the nodes of the system, we obtain the following:

$$\sum_i o_i = \sum_\alpha \sum_i k_i^{[\alpha]} = 2\sum_\alpha K^{[\alpha]} = 2O, \tag{9}$$

where O is the size of the overlapping network.

We now consider the case of a multiplex network composed of weighted layers. In such a case, for all the connected pairs of nodes i and j on each layer α of the multiplex, we have a positive real number $w_{ij}^{[\alpha]}$, namely the weight of the link $i - j$ at layer α. A weighted multilayer network is completely specified by the vector of its weighted adjacency matrices $\boldsymbol{W} = \{W^{[1]}, \ldots, W^{[M]}\}$, with $W^{[\alpha]} = \{w_{ij}^{[\alpha]}\}$. In analogy with the case of unweighted layers, also for weighted layers we can define the aggregated topological adjacency matrix $\mathcal{A} = \{a_{ij}\}$, where

$$a_{ij} = \begin{cases} 1 & \text{if } \exists\, \alpha : w_{ij}^{[\alpha]} > 0 \\ 0 & \text{otherwise} \end{cases}. \tag{10}$$

We can now extend all the previously introduced measures to the case of weighted multiplexes. We define the strength of node i on layer α as $s_i^{[\alpha]} = \sum_j w_{ij}^{[\alpha]}$. Similarly to the unweighted case, the strength of node i can be represented as a vector,

$$\boldsymbol{s}_i = \left\{s_i^{[1]}, \ldots, s_i^{[M]}\right\}, \quad i = 1, \ldots, N. \tag{11}$$

Summing over the nodes of the multiplex, we obtain $\sum_i s_i^{[\alpha]} = 2S^{[\alpha]}$, where $S^{[\alpha]}$ is the total strength of layer α.

We also define the weighted overlap of edge i-j as

$$o_{ij}^{\mathrm{w}} = \sum_\alpha w_{ij}^{[\alpha]}, \tag{12}$$

and, consequently, the weighted aggregated overlapping adjacency matrix $\mathcal{O}^{\mathrm{w}} = \{o_{ij}^{\mathrm{w}}\}$. We can also compute the weighted overlapping degree of node i as

$$o_i^{\mathrm{w}} = \sum_j o_{ij}^{\mathrm{w}} = \sum_\alpha s_i^{[\alpha]}. \tag{13}$$

Summing over the nodes, we obtain

$$\sum_i o_i^{\mathrm{w}} = \sum_\alpha \sum_i s_i^{[\alpha]} = 2\sum_\alpha S^{[\alpha]} = 2O^{\mathrm{w}}, \tag{14}$$

where O^{w} is the size of the weighted overlapping network, i.e., the total number of edges in the multiplex.

We would like to stress once more that the presence of different layers with links of different types confers a non-negligible added value to multilayer systems, which is lost by considering exclusively the aggregated matrices $\mathcal{A}$ and $\mathcal{O}$. As we will show in Secs. IV–VIII, a proper description of basic multiplex quantities such as degree, node clustering, and reachability, cannot disregard the explicit or implicit presence of the layer index α and of vectorial variables like $\boldsymbol{k_i}$ and $\boldsymbol{A}$.

III. THE MULTILAYER NETWORK OF INDONESIAN TERRORISTS

As a case of study, in this work we focus on the multiple relations among Indonesian terrorists belonging to the so-called Noordin Top terrorist network [22]. This data set includes information about trust (T), operational (O), communication (C) ties, and business (B) relations among a group of 78 terrorists from Indonesia active in recent years. In this data set, information for some of the layers can be split into a deeper level. This is the case of the trust and operational networks which are composed by four sublayers each, making them multiplexes inside a multiplex. Layer T is obtained as superposition of classmates, friendship, kinship, and soul-mate ties, while layer O can be split into logistics, meetings, operations and training sublayers. As a first approach, we represent this system as a multiplex network with $M = 4$ layers, namely T, O, C, and B.

We exploit the additional richness of the data set to assign a weight to the links connecting nodes in layers T and O, while we leave the analysis of multiplexes of multiplexes for future work. In particular, we associated an integer number $w_{ij}^{[\mathrm{T}]}$ with $1 \leqslant w_{ij}^{[T]} \leqslant 4$ to every edge in the trust layer, based on how many times the link i–j appears in the four corresponding sublayers. Analogously, an integer weight $w_{ij}^{[\mathrm{O}]}$ with $1 \leqslant w_{ij}^{[\mathrm{O}]} \leqslant 4$ is associated to every edge in the operational layer. For most of the following analysis we will consider also T and O as unweighted layers, while we will make explicit use of the weights of layers T and O in Sec. V.

Summing up, the multiplex network of Noordin Top terrorists has $N = 78$ nodes, $K = 623$ nonoverlapping links, $O = 911$ overlapping links, and $O^{\mathrm{w}} = 1014$. Table I reports more details about the size of each layer and sublayer, and of the corresponding aggregated adjacency matrices. We notice that some individuals are not involved in all the four layers, meaning that their activity with respect to a particular kind of social relationship has not been registered or was unknown at the time the data set was compiled. Consequently, these nodes will be isolated on one or more of the four layers. It is evident from Table I that while the trust, communication, and operational layers share approximately 90% of the nodes, the business layer has only 13 active nodes. Consequently, in the following we will consider only trust, communication, and operational relationships, with the exception of Sec. V where we will also briefly discuss the role of the business layer. For this three-layer multiplex network we have $N = 78$, $K = 620$, $O = 896$, and $O^{\mathrm{w}} = 999$. A schematic (aggregated) representation of this multiplex network is reported in Fig. 1. The color code indicates the layers in which nodes are involved, while the size of each node is proportional to its overlapping degree o_i. Notice that most of the nodes participate to all the three layers, while just a few of them are present in only one or two layers.

TABLE I. The Noordin Top terrorist network includes data about trust (T), operations (O), communication (C), and business (B) among 78 terrorists active in recent years in Indonesia. Trust and operational networks are characterized by a deeper internal structure, and they can be divided into four sublayers each. For the multiplex network (M) and each layer and sublayer we show the total number of active nodes $N_{\rm act}$, and the number of edges expressed as nonoverlapping links K, overlapping links O, and weighted overlapping links $O^{\rm w}$. For each layer α we also report the total strength $S^{[\alpha]}$.

Layer	Code	$N_{\rm act}$	K	S	O	$O^{\rm w}$
MULTIPLEX	M	78	623	/	911	1014
TRUST	T	70	259	293	/	/
Classmates	Tc	39	175	/	/	/
Friendship	Tf	61	91	/	/	/
Kinship	Tk	24	16	/	/	/
Soul-mates	Ts	9	11	/	/	/
OPERATIONAL	O	68	437	506	/	/
Logistics	Ol	16	29	/	/	/
Meetings	Om	26	63	/	/	/
Operations	Oo	39	267	/	/	/
Training	Ot	38	147	/	/	/
COMMUNICATION	C	74	200	200	/	/
BUSINESS	B	13	15	15	/	/

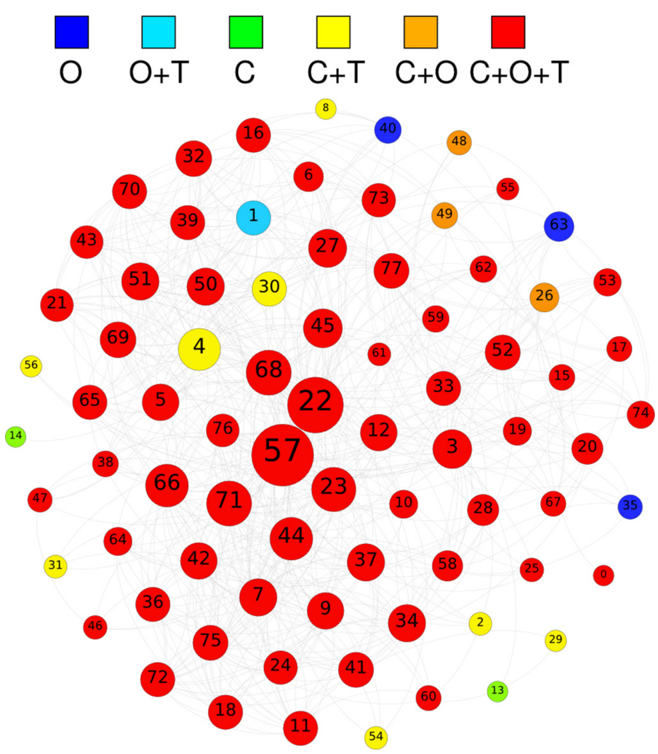

FIG. 1. (Color online) A flattened representation of the three-layer multiplex obtained by considering only trust (T), communication (C), and operations ties (O). For each node i we indicate with a color code the layers in which i is actively involved (i.e., the layers α for which $k_i^{[\alpha]} > 0$). The size of a node is proportional to its overlapping degree o_i: node 57 is the node with the largest overlapping degree.

IV. BASIC NODE PROPERTIES

One of the simplest features of a single-layer network is its degree distribution. For multiplex networks, we can study how the degree is distributed among the different nodes at each layer, but it is also important to evaluate how the degree of a node is distributed across different layers. It is in fact possible that nodes which are hubs in one layer have only few connections, or are even isolated, in another layer. Or, alternatively, nodes which are hubs in one layer are also hubs in the other layers. We have therefore computed the aggregated topological degree k_i and the degree of the nodes in each layer $k_i^{[\alpha]}$, with $\alpha \in \{\mathrm{T,O,C}\}$, ranking the nodes according to their aggregated topological degree. In Fig. 2(a) we compare with a color-code plot the values of k_i with the values $k_i^{[\alpha]}$ of the node degree at each layer α. By visual inspection, the four degree sequences appear weakly correlated, with nodes which are hubs in one level often having only few connections in another layer. In Fig. 2(b) we report the results obtained by ranking the nodes according to their overlapping degree o_i. Also in this case we observe weak correlations between the four degree sequences. To better quantify such correlations, we computed the Kendall rank correlation coefficient, τ_k, which measures the similarity of two ranked sequences of data X and Y. The correlation coefficient τ_k is a nonparametric measure of statistically dependence between two rankings, since it does not make any assumption about the distributions of X and Y and takes values in $[-1,1]$. We get $\tau_k(X,Y) = 1$ if the two rankings are identical, $\tau_k(X,Y) = -1$ if one ranking is exactly the reverse of the other and finally $\tau_k(X,Y) = 0$ if X and Y are independent. In Fig. 2(c) we report as a heat map the values of τ_k obtained for the rankings of each pair of variables. Notice that the aggregated degrees k_i and o_i are usually weakly correlated with the degree of node i on each single layer. The highest correlation is indeed found between the degree of the aggregated topological network k_i and the overlapping degree o_i.

Due to the heterogeneity in the degree distribution of each layer and to the weak correlation observed between the degrees of the same node at different layers, it is necessary to introduce a measure to quantify the richness of the connectivity patterns across layers. For instance, consider two nodes i and j having exactly the same value of overlapping degree $o_i = o_j$, and imagine that i is a massive hub on a layer α and an isolated node on the other layers, so $o_i = k_i^{[\alpha]}$, while j has the same number of edges on each layer, so $o_j = M k_j^{[\alpha]}, \forall \alpha$. From a multiplex perspective i and j have radically different roles, but this fact is not detectable by comparing their overlapping degrees, which have the same value. Conversely, even if o_i and o_j differ substantially, i and j can look very similar if one considers the contribution of each layer to the total overlapping degree of the two nodes.

A suitable quantity to describe the distribution of the degree of node i among the various layers is the entropy of the multiplex degree,

$$H_i = -\sum_{\alpha=1}^{M} \frac{k_i^{[\alpha]}}{o_i} \ln\left(\frac{k_i^{[\alpha]}}{o_i}\right). \qquad (15)$$

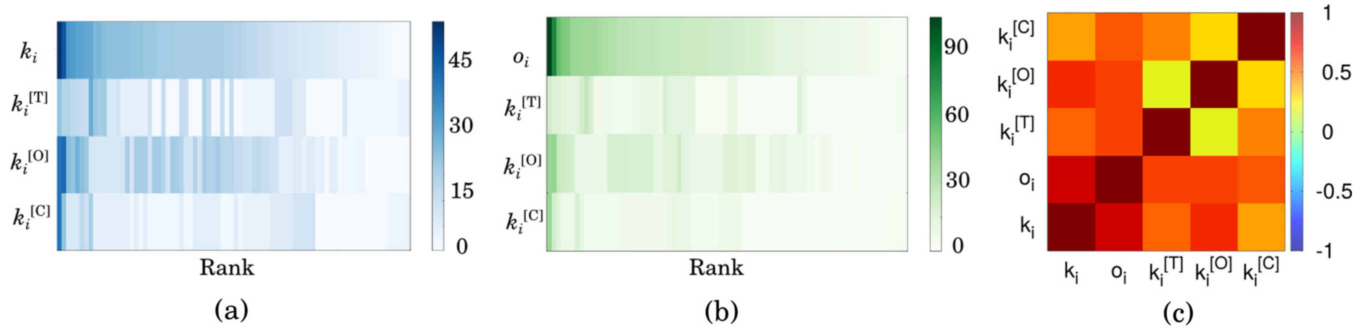

FIG. 2. (Color online) (a) The top row of the panel shows with a color code the degree k_i of each node, from the largest (darkest, leftmost) to the smallest (brightest, rightmost). Keeping fixed the ranking induced by k_i, in the other three rows we report, respectively, the degree in the trust layer $k_i^{[T]}$, operational layer $k_i^{[O]}$, and communication layer $k_i^{[C]}$. (b) Same as panel (a) but in the first row nodes are ranked according to their overlapping degree o_i. (c) The heat map represents the Kendall τ correlation coefficient among k_i, o_i, $k_i^{[T]}$, $k_i^{[O]}$, and $k_i^{[C]}$. Notice that the degree of a node in the operational layer O is poorly correlated with its degree in the communication and trust layers (bright yellow regions in the heat map).

This entropy is equal to zero if all the links of node i are in a single layer, while it takes its maximum value when the links are uniformly distributed over the different layers. In general, the higher the value of H_i, the more uniformly the links of node i are distributed across the layers. A similar quantity is the multiplex participation coefficient P_i of node i,

$$P_i = \frac{M}{M-1}\left[1-\sum_{\alpha=1}^{M}\left(\frac{k_i^{[\alpha]}}{o_i}\right)^2\right]. \tag{16}$$

The definition of the multiplex participation coefficient is in the same spirit of that of participation coefficient introduced in Refs. [23,24] to quantify the participation of a node to the different communities of a network. In this adaptation to multilayer networks, P_i takes values in the interval [0,1] and measures whether the links of node i are uniformly distributed among the M layers, or are instead primarily concentrated in just one or a few layers. Namely, the coefficient P_i is equal to 0 when all the edges of i lie in one layer, while $P_i = 1$ only when node i has exactly the same number of edges on each of the M layers. In general, the larger the value of P_i, the more equally distributed is the participation of node i to the M layers of the multiplex. The participation coefficient P of the whole multiplex is defined as the average of P_i over all nodes, i.e., $P = N^{-1}\sum_i P_i$. The two quantities P_i and H_i give very similar information, so in the following we will discuss the results for P_i only.

In Fig. 3(a) we plot the distribution of P_i for the multilayer network of Indonesian terrorists under study. Although the average participation coefficient of the multiplex is equal to $P = 0.72$, we observe a quite broad distribution of P_i in the range [0,1]. This variance suggests the existence in the network of various levels of node participation to each of the three layers. Since the overlapping degree of a node represents its overall importance in terms of number of incident edges, while the multiplex participation coefficient gives information about the distribution of incident edges across the layers, we propose to classify the nodes of a multiplex by looking, at the same time, at their multiplex participation coefficient and at their overlapping degree. With respect to the multiplex participation coefficient, we identify three classes. We call *focused* those nodes for which $0 \leqslant P_i \leqslant 1/3$, *mixed* the nodes having $1/3 < P_i \leqslant 2/3$, and *truly multiplex* (or even simply *multiplex*) the nodes for which $P_i > 2/3$. Instead of the overlapping degree we consider the associated Z score, which allows us to compare multiplex networks of different size,

$$z(o_i) = \frac{o_i - \langle o\rangle}{\sigma_o}, \tag{17}$$

where $\langle o\rangle$ is equal to the average overlapping degree of the nodes of the system and σ_o is the corresponding standard deviation. With respect to the Z score of their overlapping degree, we distinguish *hubs*, for which $z(o_i) \geqslant 2$, from regular nodes, for which $z(o_i) < 2$. Consequently, by considering the multiplex participation coefficient P_i of a node and its total overlapping degree o_i we can define six classes of nodes, as depicted in Fig. 3(b), where we represent each node as a point in the $(P_i, z(o_i))$ plane. Notice that the distribution of $z(o_i)$ is asymmetric and unbalanced towards positive values, and this is a sign of the heterogeneity of the total overlapping degree. Moreover, there is a quite large heterogeneity in the values of P_i for a fixed value of $z(o_i)$. Let us focus, for instance, on two specific nodes, namely node 16 and 34. These two nodes have the same overlapping degree, namely $o_{16} = o_{34} = 25$, corresponding to $z(o_{16}) = z(o_{34}) = 0.12$, but very different participation coefficient across layers T, O, and C, respectively, $P_{16} = 0.915$ and $P_{34} = 0.23$. Consequently, even if the overall number of edges of node 16 and node 34 is the same (which would make these two nodes indistinguishable in the aggregated overlapping network), they play radically different roles, as becomes evident by looking at their ego networks, reported in Fig. 3(c). In fact, while node 34 is highly focused on the operational layer (blue edges), with only one edge in the trust layer (green edge) and one edge in the communication layer (red edge), node 16 is instead involved in all the three layers, with a comparable number of edges in each of them. This implies that the removal of node 34 would primarily affect just the operational layer, while the absence of node 16 could cause major disruptions in the trust, operational, and communication networks. Similar results are obtained by considering the Z score of the degree k_i of node i in the aggregated topological network.

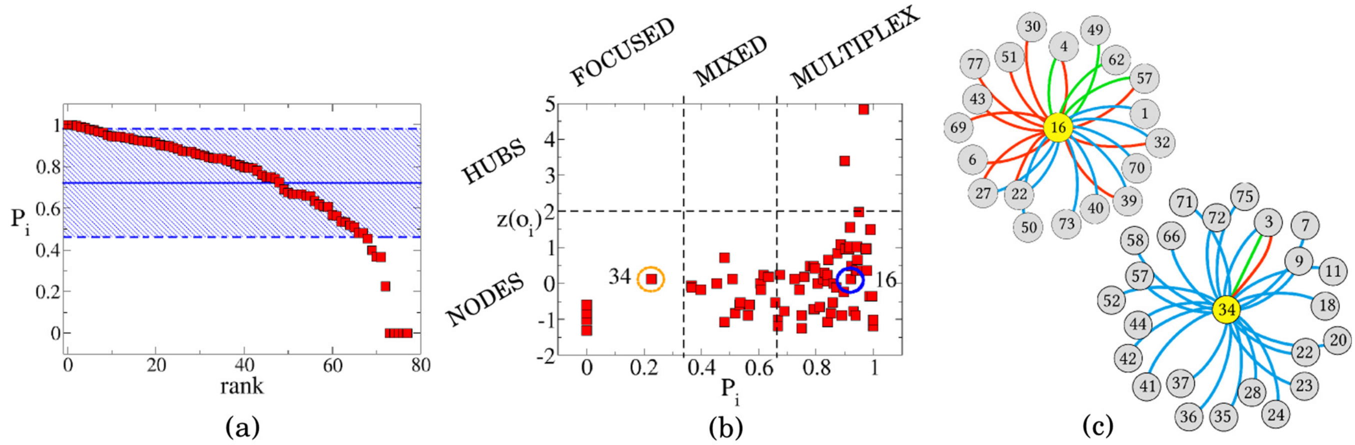

FIG. 3. (Color online) (a) Rank distribution of the participation coefficient P_i for the multilayer network of Noordin Top Indonesian terrorists. $M = 3$ layers were considered, namely trust, operational, and communication. The average value $P = 0.72$ is shown as a horizontal blue line, while the shaded band indicates the standard deviation. (b) A cartography of the roles of the nodes in a multilayer network can be obtained by plotting, for each node i, the multiplex participation coefficient P_i versus the Z score of the total overlapping degree $z(o_i)$. Even if two nodes have exactly the same value of $z(o_i)$ (like node 16 and node 34, indicated by the orange and blue circles, respectively), they can have pretty different roles, according to the value of the multiplex participation coefficient. (c) The ego networks of node 16 and 34, in which edges are colored according to the layer to which they belong, respectively green (trust), blue (operational), and red (communication). It is evident that the connectivity pattern of node 16, whose links are homogeneously distributed across the three layers, is "more *multiplex*" than that of node 34, which is instead *focused* on the operational layer.

V. EDGE OVERLAP AND SOCIAL REINFORCEMENT

After having proposed some measures of the role of individual nodes in the multiplex, we now aim at quantifying the importance of each layer as a whole. For instance, we can detect the existence of correlations across the layers of a multiplex by computing the edge overlap o_{ij} of Eq. (7) for each edge i-j, and by looking at how this quantity is distributed. We now consider the multiplex formed by all the four layers of the Noordin Top Indonesian terrorist network, i.e., the trust, operational, communication, and business layers, so $1 \leqslant o_{ij} \leqslant 4$ for all possible pairs of nodes connected by at least one edge. If we look at the distribution of o_{ij}, we see that 46% of the edges exist in just one of the four layers, 27% are present in two layers, 23% exist in three layers, and only 4% are present in all four layers.

Besides the distribution of o_{ij} gives some information about the existence of interlayer correlations, it is not able to disentangle the relevance of single layers. A slightly more sophisticated quantity we can look at is the conditional probability of finding a link at layer α' given the presence of an edge between the same nodes at layer α,

$$P\left(a_{ij}^{[\alpha']}|a_{ij}^{[\alpha]}\right) = \frac{\sum_{ij} a_{ij}^{[\alpha']} a_{ij}^{[\alpha]}}{\sum_{ij} a_{ij}^{[\alpha]}}. \qquad (18)$$

The denominator of Eq. (18) is equal to the number $K^{[\alpha]}$ of edges at layer α, while the numerator is equal to the number of such edges which are also present at the layer α'. The conditional probability $P(a_{ij}^{[\alpha']}|a_{ij}^{[\alpha]})$ is shown as a heat map in Fig. 4(a) for the four layers. For instance, the first column shows with a color code the probability to find a link on layer T given its existence on layer B, C, O, or T (obviously, we have $P(a_{ij}^{[T]}|a_{ij}^{[T]}) = 1$), while the last row represents the fraction of edges in layer T which also exist in layer T, O, C, and B. Since layers T and O have a composite internal structure of four levels each, which allows us to assign a weight $w_{ij}^{[\mathrm{T}]}$ and $w_{ij}^{[\mathrm{O}]}$ to each pair of connected nodes i and j, it is interesting to study the probability $P^{\mathrm{w}}(a_{ij}^{[\alpha']}|w_{ij}^{[\alpha]})$ of having a link on layer α' given its weight on the leading layer α, with α corresponding to layers O and T. In Fig. 4(b) we plot the probability of finding a link at layer O, C, and B, given the weight $w_{ij}^{[\mathrm{T}]}$ of the link at layer T. Even though in principle $w_{ij}^{[\mathrm{T}]} = 4$ is possible, none of the edges appears in all the sublayers of the trust layer. In all the three cases, P^{w} is an increasing function of $w_{ij}^{[\mathrm{T}]}$. Figure 4(b) suggests that the stronger the trust connection between two terrorists the higher the probability for them to operate together, communicate, or have a common business. In particular, for layers O and C, which are the ones that have a number of nodes comparable to the one of layer T, already a value of $w_{ij}^{[\mathrm{T}]} = 2$ implies that the two people have common operations and communications in 80% of the cases. If $w_{ij}^{[\mathrm{T}]} = 3$, then the probability that the edge i-j exists in all the three remaining layers is equal to 1.

This phenomenon can be explained in terms of social reinforcement, meaning that the existence of strong connections in the Trust layer, which represents the strongest relationships between two people, actually fosters the creation of links in other layers and produces a measurable effect on the probability to operate, communicate, and do business together. Despite we do not have longitudinal information to test the hypothesis that original trust connections actually caused the creation of links in other layers by means of social reinforcement, in this particular case we have to stress that the strength of the trust relationship between two individuals is higher if they had been kin, classmates, soul mates, and/or friends, respectively [22]. This means that, with high probability, the establishment of any of the four Trust relationships between i and j preceded

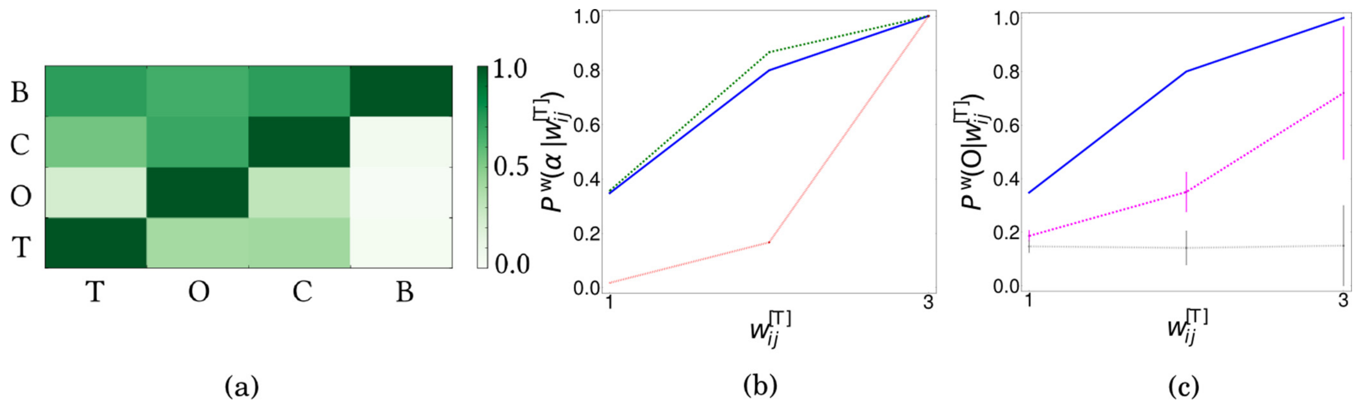

FIG. 4. (Color online) (a) For each layer α, we show in the color map the fraction of edges which is also present in each other layer α'. (b) Probability P^{w} of finding a certain link at layer O (solid blue line), C (dotted green line), and B (dashed red line), conditional to the weight $w_{ij}^{[T]}$ of the same link at layer T. (c) The values of P^{w} computed on real data for layer O (solid blue line) are compared to those obtained by randomizing the edges keeping fixed the total number of links $K^{[O]}$ (dashed gray line) or the degree distribution $P(k^{[O]})$ (dotted magenta line) for the operational layer.

by several years the establishment of any communication, operational, or business relationship registered during the collection of the data set. Consequently, it is not too pretentious to suggest that a social reinforcement mechanism took place in this small social system and that trust relationships have actually *caused* subsequent communication and collaboration among the terrorists.

In order to statistically validate these results, in Fig. 4(c) we report the expected values of P^{w} obtained by randomizing the nonleading layers while keeping fixed either the total number of links $k^{[\alpha]}$ or the degree distribution $P(k^{[\alpha]})$. In the first case, each nonleading layer is an Erdös-Renyi random graph and P^{w} is not even correlated with the weights on layer T, as expected. In the second case, which is an extension to multiplexes of the configuration model, for each weight $w_{ij}^{[T]}$ the conditional probability to find an edge on the operational layer is systematically lower in the randomized networks than in the original one. Hence, we can conclude that interlayer correlations among the heterogeneous degree distributions of the various levels do not provide an ultimate explanation for the behavior of P^{w} that the Trust layer is genuinely *driving* the observed connection pattern. Results analogous to that of layer O were also found for layers C and B.

Similar results are obtained considering the operational network (instead of the trust network) as a leading layer, but in this case the conditional probability of finding an edge in T, C, and B given its weight in O was substantially smaller than those reported in Figs. 4(b) and 4(c). This is not surprising at all, since while it is clear that a stronger level of trust between two individuals can boost their communications and their common operations, we expect a weaker causality between the strength of different operations two individuals have shared and their trust and communications. The existence of a weaker interaction between the operational layer and the other three layers increases the validity of our hypothesis that the trust layer is indeed controlling the overall structure of the multiplex network through a social reinforcement mechanism and that the relative importance of the trust layer for the formation of edges on other layers is not a mere consequence of existence of sublayers.

VI. TRANSITIVITY AND CLUSTERING

One of the most remarkable characteristic of complex real-world single-layer networks, especially acquaintance and collaboration networks, is the tendency of nodes to form triangles, i.e., simple cycles involving three nodes. This widely observed tendency is concisely expressed by the popular saying "the friend of your friend is my friend" and is usually quantified through the so-called node clustering coefficient [25]. The clustering coefficient of node i is defined as

$$C_i = \frac{\sum_{j\neq i,m\neq i} a_{ij}a_{jm}a_{mi}}{\sum_{j\neq i,m\neq i} a_{ij}a_{mi}} = \frac{\sum_{j\neq i,m\neq i} a_{ij}a_{jm}a_{mi}}{k_i(k_i-1)} \quad (19)$$

and quantifies how likely it is that two neighbors of node i are connected to each other. In fact, Eq. (19) measures the fraction of triads centered in i that close into triangles. By definition C_i takes values in the interval $[0,1]$. Averaging this quantity over all the nodes in a network, one gets the network clustering coefficient:

$$C = \frac{1}{N}\sum_i C_i. \quad (20)$$

A similar—although not identical—measure of local cohesion [26], which is commonly used in the social sciences, is the network transitivity [27],

$$T = \frac{3 \times \text{No. of triangles in the graph}}{\text{No. of triads in the graph}}. \quad (21)$$

This is defined as the proportion of triads, i.e., connected triples of nodes, which close into triangles.

Since each layer of a multiplex can be seen as a single-layer network, the definitions of network clustering coefficient and network transitivity can be used to characterize the abundance of triangles on each layer. In general, different layers may show similar or dissimilar patterns of clustering. In Table II we report the average clustering coefficient and the transitivity

TABLE II. The average clustering coefficient C and the transitivity T for layers T, O, and C and for the aggregated topological network $\mathcal{A}$.

Layer	C	T
T	0.38	0.53
O	0.67	0.62
C	0.45	0.27
$\mathcal{A}$	0.66	0.56

for each layer of the terrorist network and for its topological aggregate.

Notice that each layer has quite peculiar values of clustering and transitivity, which in turn differ from those measured on the aggregated topological network. In particular, the highest values of clustering and transitivity are observed in the Operations layer, probably due to the fact that terrorist missions usually involve more than two people at the same time. In Fig. 5(a) we focus on the node clustering coefficient, we rank the nodes of the multiplex according to the value of C_i for the aggregated topological network, and we compare this value with the clustering coefficient calculated on each layer $C_i^{[\alpha]}$. As shown, many nodes display quite different values of clustering coefficient across the layers. We have computed the Kendall correlation coefficient τ_k between each pair of layers and between each layer and the topological aggregate. The results are shown in Fig. 5(b) as a heat map. Notice that at the best the sequences of clustering coefficient are weakly correlated, when not uncorrelated or even anticorrelated. In particular, the ranking of clustering coefficient for the Operations layer is anticorrelated with that of the other three layers and of the topological aggregated network.

However, comparing the sequences of C_i for each layer tells us very little about the interplay between the several levels of the system in terms of clustering. In particular, it is interesting to study to which extent the multiplexicity affects the formation of triangles, i.e., how the presence of different layers can give rise to triangles which were impossible to close at the level of single layers. For this reason we need to extend the notion of triangle to take into account the richness added by the presence of more than one layer. We define a two-triangle a triangle which is formed by an edge belonging to one layer and two edges belonging to a second layer. Similarly, we call a

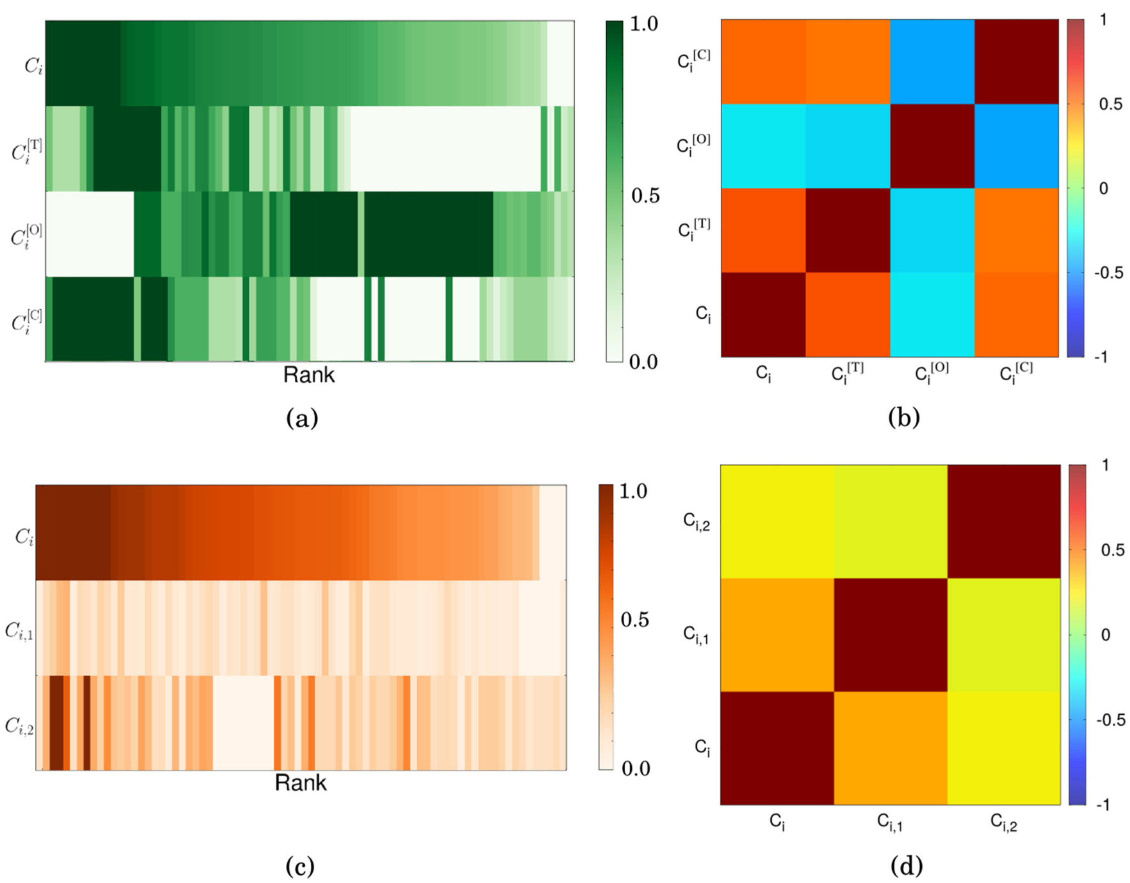

FIG. 5. (Color online) (a) The node clustering coefficient C_i of the aggregated topological network and of the three layers T, O, and C, respectively denoted as $C_i^{[T]}$, $C_i^{[O]}$, and $C_i^{[C]}$. The nodes are ranked according to their value of C_i on the aggregated topological network. (b) The heat map represents the correlation between the rankings of nodes according to their clustering coefficients on the three layers and on the topological aggregated network. Notice that $C_i^{[\alpha]}$ is weakly correlated with C_i for $\alpha \in \{T,O,C\}$, and that such a correlation might also be negative, as in the case of $C_i^{[O]}$. (c) Comparison among the clustering coefficient C_i of the aggregated topological network and the multilayer clustering coefficients $C_{i,1}$ and $C_{i,2}$. The nodes are ranked according to their value of C_i. (d) The heat map represents the correlation between the rankings of nodes according to C_i, $C_{i,1}$, and $C_{i,2}$.

three-triangle a triangle which is composed by three edges all lying in different layers. In order to quantify the added value provided by the multiplex structure in terms of clustering, we define two parameters of clustering interdependence I_1 and I_2. I_1 is the ratio between the number of triangles in the multiplex which can be obtained only as two-triangles, and the number of triangles in the aggregated system. I_2 is the ratio between the number of triangles in the multiplex which can be obtained only as three-triangles and the number of triangles in the aggregated system. Then $I = I_1 + I_2$ is the total fraction of triangles of the aggregated topological network which cannot be found entirely in one of the layers. For the multilayer network of terrorists we obtain $I_1 = 0.31$ and I_2 of the order of 10^{-3}, which indicates that almost no triangle is formed exclusively by the interplay of three different layers. This result suggests the presence of nontrivial patterns in clustering and triadic closure in multilayer systems.

In this work we also aim at generalizing the notion of clustering coefficient to multilayer networks. Recalling the definition of two-triangles and three-triangles, we define a one-triad centered at node i, for instance, j-i-k, a triad in which both edge j-i and edge i-k are on the same layer. We also define a two-triad as a triad whose two links belong to two different layers of the systems. We are now ready to give two definitions of clustering coefficient for multiplex networks. Similar definitions have been recently—and independently—proposed in Ref. [28]. The first coefficient $C_{i,1}$ is defined, for each node i, as the ratio between the number of two-triangles with a vertex in i and the number of one-triads centered in i. We can express this clustering coefficient in terms of the multilayer adjacency matrix as follows:

$$C_{i,1} = \frac{\sum_\alpha \sum_{\alpha' \neq \alpha} \sum_{j\neq i, m\neq i} \left(a_{ij}^{[\alpha]} a_{jm}^{[\alpha']} a_{mi}^{[\alpha]}\right)}{(M-1)\sum_\alpha \sum_{j\neq i, m\neq i} \left(a_{ij}^{[\alpha]} a_{mi}^{[\alpha]}\right)} = \frac{\sum_\alpha \sum_{\alpha' \neq \alpha} \sum_{j\neq i, m\neq i} \left(a_{ij}^{[\alpha]} a_{jm}^{[\alpha']} a_{mi}^{[\alpha]}\right)}{(M-1)\sum_\alpha k_i^{[\alpha]}\left(k_i^{[\alpha]}-1\right)}. \qquad (22)$$

Since each one-triad can theoretically be closed as a two-triangle on each of the M layers of the multiplex excluding the layer to which its edges belong, in order to have a normalized coefficient we have to divide the term by $M-1$. In addition to this, we define a second clustering coefficient for multiplex networks as the ratio between the number of three-triangles with node i as a vertex, and the number of two-triads centered in i. In terms of adjacency matrices, we have

$$C_{i,2} = \frac{\sum_\alpha \sum_{\alpha' \neq \alpha} \sum_{\alpha'' \neq \alpha,\alpha'} \sum_{j\neq i, m\neq i} \left(a_{ij}^{[\alpha]} a_{jm}^{[\alpha'']} a_{mi}^{[\alpha']}\right)}{(M-2)\sum_\alpha \sum_{\alpha' \neq \alpha} \sum_{j\neq i, m\neq i} \left(a_{ij}^{[\alpha]} a_{mi}^{[\alpha']}\right)}. \qquad (23)$$

where a normalization coefficient $M-2$ has been added. While $C_{i,1}$ is a suitable definition for multiplexes with $M \geqslant 2$, $C_{i,2}$ can only be defined for systems composed of at least three layers. Averaging over all the nodes of the system, we obtain the network clustering coefficients C_1 and C_2.

In Fig. 5(c) we rank the nodes of the terrorist network according to their value of C_i for the aggregated system and compare this sequence of values with the ones obtained with the two measures of multiplex clustering, $C_{i,1}$ and $C_{i,2}$. As shown in the figure, $C_{i,1}$ and $C_{i,2}$ capture different effects

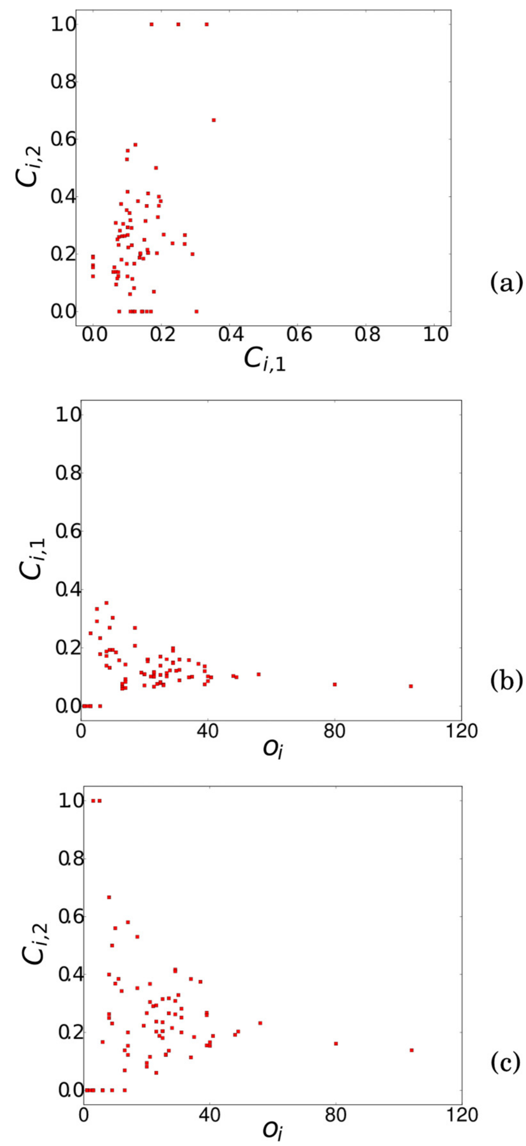

FIG. 6. (Color online) Scatter plots of (a) $C_{i,1}$ versus $C_{i,2}$ (b) $C_{i,1}$ versus o_i and (c) $C_{i,2}$ versus o_i. The values of the Kendall's τ and of the Pearson's linear correlation coefficient r for any pair of measures are, respectively, $\tau(C_{i,1},C_{i,2}) = 0.61$, $r(C_{i,1},C_{i,2}) = 0.76$, $\tau(C_{i,1},o_i) = -0.11$, $r(C_{i,1},o_i) = -0.13$, $\tau(C_{i,2},o_i) = 0.01$, $r(C_{i,2},o_i) = 0.04$. It is worth noticing that both $C_{i,1}$ and $C_{i,2}$ are almost uncorrelated with the overlapping degree o_i, a fact that confirms their truly multiplex nature.

of multilayer clustering. This fact is confirmed by the heat map reported in Fig. 5(d), which shows with a color code the nonparametric correlations among $C_{i,1}$, $C_{i,2}$, and C_i. Notice that, in general, the correlation between C_i and both $C_{i,1}$ and $C_{i,2}$ is pretty small.

These results indicate that multiplex clustering provides information which substantially differ from those obtained by looking at the clustering of the aggregated network. In addition to this, $C_{i,1}$ and $C_{i,2}$ are poorly correlated, as is also evident from Fig. 6(a). In practice, for a given value of $C_{i,1}$, we have nodes with a wide range of values of $C_{i,2}$ and vice versa. Consequently, it is necessary to use both clustering coefficients in order to properly quantify the abundance of triangles in multilayer networks. In Figs. 6(b) and 6(c) we report the scatter plots of $C_{i,1}$ and $C_{i,2}$ versus o_i. Multiplex clustering coefficients are genuine multiplex variable and appear to be not correlated with the degree of the nodes of the system. We also found that the clustering coefficient is not correlated with other measures of aggregated degree, such as k_i and $o_i^{\rm w}$.

We can also generalize the definition of transitivity T to the case of multilayer networks. Similarly to the case of the clustering coefficient, we propose two measures of transitivity. We define T_1 as the ratio between the number of two-triangles and $M-1$ times the number of one-triads in the multilayer network. Moreover, we introduce T_2 as the ratio between the number of three-triangles and $M-2$ times the number of two-triads in the system.

Notice that clustering interdependences I_1 and I_2, average multiplex clustering coefficients C_1 and C_2, and multiplex transitivities T_1 and T_2 are all global graph variables which give a different perspective on the multilayer patterns of clustering and triadic closure with respect to the clustering coefficient and the transitivity computed for each layer of the network. We have computed all such quantities for the multilayer network of the Indonesian terrorists and, as a term of comparison, we have constructed a configuration model for multiplex networks, which will be useful to investigate the nontrivial organization of the network under study.

In analogy with the case of a single-layer network, for a multiplex with M layers, where each node is characterized by a degree vector $\boldsymbol{k}_i$, we call configuration model the set of multiplexes obtained from the original system by randomizing edges and keeping fixed the sequence of degree vectors $\{\boldsymbol{k}_1,\boldsymbol{k}_2,\ldots,\boldsymbol{k}_N\}$, i.e., keeping fixed the degree sequence at each layer α. We can now compare the values of C and T, C_1 and C_2, T_1 and T_2, and I_1 and I_2 obtained on real data with the average values found for the multilayer configuration model. The comparison is shown in Table III. As expected C and T computed on the aggregated topological network for real data are systematically higher than the ones obtained on randomized data, where edge correlations are washed out by the randomization. For the same reason, C_1, C_2, T_1, and T_2 are higher on real data. Conversely, we obtained higher values on randomized data for I_1 and I_2. This is not surprising, since the measures of clustering interdependence tell us about the fraction of triangles which can be exclusively found as multitriangles in the system. Since the configuration model washes out interlayer correlations, it is generally easier to find multitriangles on a randomized multiplex network rather than on a real one where edges have a higher overlap.

All these results demonstrate that, as previously shown for the overlap, also the clustering coefficient appears to be affected by the presence of nontrivial structural properties across the different layers of the multiplex network under study.

TABLE III. Values of clustering C and transitivity T computed on the aggregated topological network, and values of the introduced measures for clustering in multilayer networks, namely the multiplex clustering coefficients C_1 and C_2, the multiplex transitivities T_1 and T_2, and the clustering interdependences I_1 and I_2. For comparison we report also the results for a randomized system obtained through a multilayer configuration model.

Variable	Real data	Randomized data
C	0.66	0.46
T	0.56	0.41
C_1	0.13	0.08
C_2	0.26	0.18
T_1	0.10	0.07
T_2	0.21	0.16
I_1	0.31	0.60
I_2	0.005	0.047

VII. REACHABILITY, SHORTEST PATHS, AND INTERDEPENDENCE

Reachability is an important feature in networked systems. In single-layer networks it has to do with the existence and length of shortest paths connecting pairs of nodes. In multilevel systems, shortest paths may significantly differ between different layers, and between layer and the aggregated topological networks as well. To capture the multiplex contribution to the reachability of each unit of the network, the so-called node interdependence has been recently introduced in Refs. [12,29].

The interdependence λ_i of node i is defined as

$$\lambda_i = \sum_{j\neq i} \frac{\psi_{ij}}{\sigma_{ij}}, \tag{24}$$

where σ_{ij} is the total number of shortest paths between node i and node j on the multiplex network and ψ_{ij} is the number of shortest paths between node i and node j which make use of links in two or more than two layers. Hence, the node interdependence is equal to 1 when all shortest paths make use of edges laying at least on two layers and equal to 0 when each of the shortest paths makes use of only one of the M layers of the system. Averaging λ_i over all nodes, we obtain the network interdependence $\lambda = (1/N)\sum_i \lambda_i$. In Fig. 7 we display the rank distribution of λ_i. The network has a large variety of node interdependencies: although most of the nodes have a value of λ_i in the range [0.27,0.56] around the average value $\lambda = 0.41$, there are also nodes with values as small as $\lambda_i = 0.1$, and two nodes with values larger than 0.8.

The interdependence is a genuine multiplex measure and, as shown in Fig. 7(b), provides information in terms of reachability which is slightly anticorrelated to measures of degree such as o_i. In fact, a node with a high overlapping degree quite likely will have a number of different possibilities to choose the first edge to go towards the other nodes, and in this way it will have a low value of λ_i. Conversely, a node with low degree will more likely have a high value of λ_i, being its shortest paths constrained to a limited selection of edges and layers from the first step. Moreover, λ_i appears to be slightly anticorrelated with P_i, as confirmed by the values

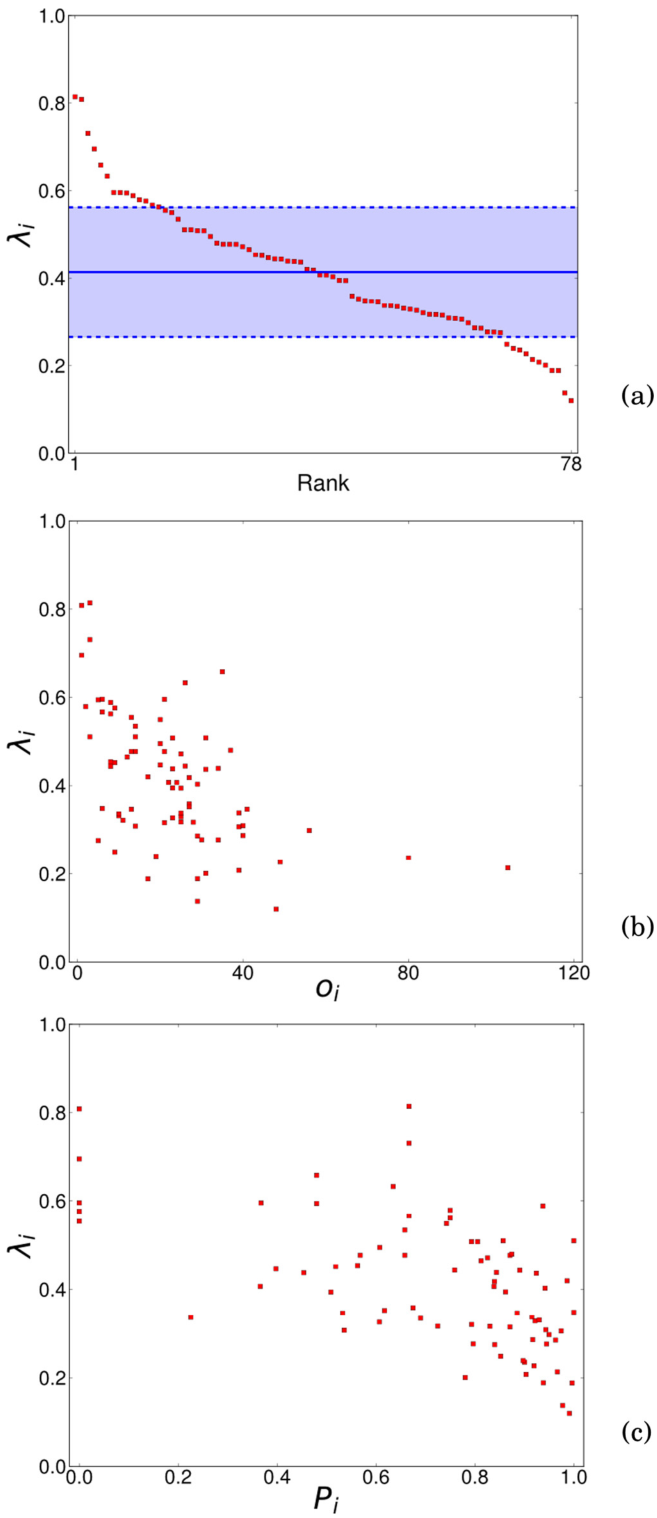

FIG. 7. (Color online) (a) Rank distribution of the node interdependence λ_i in the Indonesian terrorist multiplex network. (b) Scatter plot of the interdependence λ_i versus o_i and (c) versus P_i. The corresponding value of Kendall's τ and Pearson's r correlation coefficients are, respectively, $\tau(\lambda_i,o_i) = -0.41$, $r(\lambda_i,o_i) = -0.56$, $\tau(\lambda_i,P_i) = -0.41$, and $r(\lambda_i,P_i) = -0.57$.

of Kendall's and Pearson's correlation coefficients (see the caption to Fig. 7).

VIII. CENTRALITY

The concept of node centrality is an important and well-studied issue in network theory. Various measures of centrality, such as the node degree, the closeness, and the betweenness, have been proposed and used over the years to quantify the importance of a node in a single-layer network [27]. The extension of these concepts to multiplex networks is still an open research question. In Sec. II we have proposed various ways to extend the definition of node degree to the case of a multilayer system. Here we will focus our attention on the eigenvector centrality, which is a generalization of the concept of degree centrality. In a single-layer network the eigenvector centrality of a node i is defined as the i-th component of the eigenvector associated to the leading eigenvalue of the adjacency matrix of the network [30]. For a multiplex network, we can calculate the eigenvector centrality at each layer. If we denote as $E_i^{[\alpha]}$ the eigenvector centrality of node i at layer α, then the eigenvector centrality of node i in the multiplex network is a vector,

$$\boldsymbol{E}_i = \{E_i^{[1]}, \ldots, E_i^{[M]}\}. \tag{25}$$

We can also compute the eigenvector centrality on the aggregated topological and on the aggregated overlapping network. We indicate the results respectively as $E_i(\mathcal{A})$ and $E_i(\mathcal{O})$. In Figs. 8(a), 8(b), and 8(c) we compare the eigenvector centrality computed on each layer with that evaluated on the aggregated topological and overlapping networks. We notice only very weak correlations between the different centrality sequences. Such results are very similar to those obtained in Sec. IV for the case of node degree, as a consequence of the fact that, at order zero, the eigenvector centrality reduces to the node degree. The Kendall correlation coefficients obtained for pairs of centralities are reported in Fig. 8(d) as a heat map. For a large fraction of nodes, the rankings induced by the eigenvector centrality at different layers differ significantly. A slightly higher value of correlation is found between centrality at different layers and the centrality of the aggregated network, while the maximum correlation is observed between the values of eigenvector centrality computed on the aggregated topological network and on the aggregated overlapping network.

It is interesting to notice, as shown in Fig. 9, that the centrality computed on the aggregated networks (e.g., on the overlapping network) is not correlated with the multiplex participation coefficient of the nodes. In fact, if we fix the value of $E_i(\mathcal{O})$, we observe a large heterogeneity in the values of P_i and vice versa.

Until now we have computed and compared the eigenvector centralities at each layer of the network. As already done for other metrics, we will now propose a proper multiplex definition of the eigenvector centrality which takes into account the presence of all layers at the same time. We follow a similar but relatively simpler approach than the one recently proposed in Ref. [31]. Given a two-layer multiplex network (a duplex) and the corresponding adjacency matrices $A^{[1]}$ and $A^{[2]}$, we can construct the following adjacency matrix:

$$\mathcal{M}(b) = bA^{[1]} + (1-b)A^{[2]}, \tag{26}$$

which is a convex combination of $A^{[1]}$ and $A^{[2]}$, where b is a parameter taking values in the interval [0,1]. We

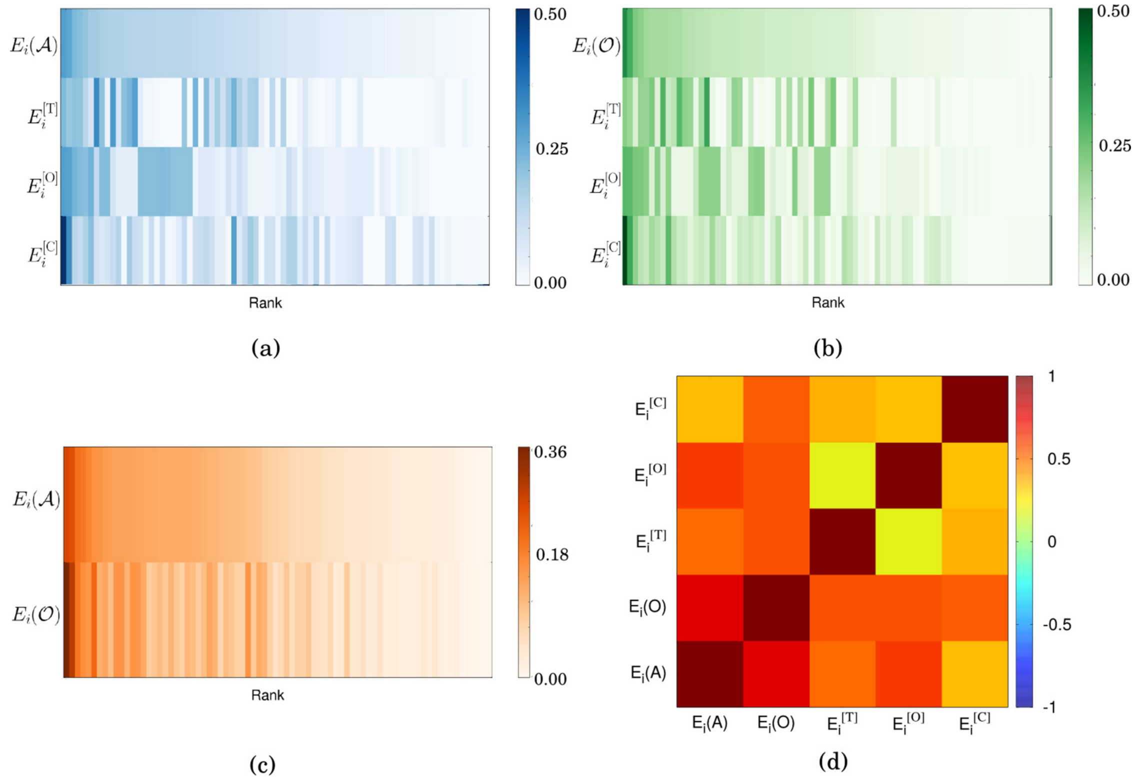

FIG. 8. (Color online) (a) Eigenvector centrality of the aggregated topological network $E_i(\mathcal{A})$ and of the trust $E_i^{[T]}$, operational $E_i^{[O]}$, and communication layer $E_i^{[C]}$. The nodes are ranked according to their value of $E_i(\mathcal{A})$ on the aggregated topological network. (b) Similar to panel (a) but here the nodes are ranked according to their eigenvector centrality computed on the aggregated overlapping network $E_i(\mathcal{O})$. (c) Comparison of the rankings of eigenvector centrality computed on the aggregated topological network and on the aggregated overlapping network, respectively, $E_i(\mathcal{A})$ and $E_i(\mathcal{O})$. (d) The heat map shows the nonparametric correlation between the rankings induces by the different centralities.

call such matrix the multiadjacency matrix. Notice that the parameter b sets the relative contribution of each layer to the multiplex structure. In fact, if $b = 0$ (respectively, $b = 1$) the multiadjacency matrix of the duplex reduces to $A^{[2]}$ (respectively $A^{[1]}$). We can consider $b = 0.5$ as the benchmark case, where the two layers are given the same weight. Notably, we have $\mathcal{M}(b = 0.5) = \mathcal{O}/2$, i.e., for $b = 0.5$ the multiadjacency matrix is proportional to the aggregated overlapping network.

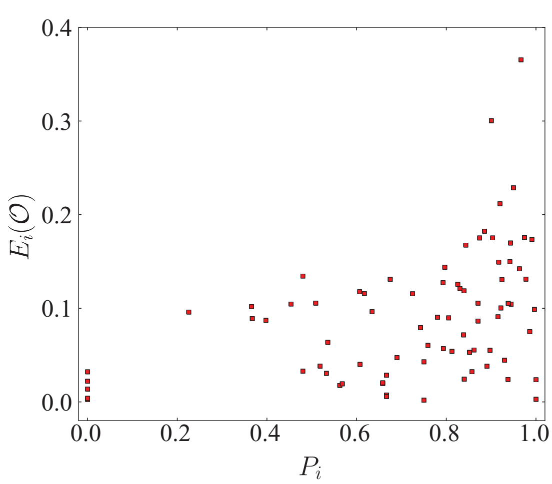

FIG. 9. (Color online) Scatter plot of the eigenvector centrality on the aggregated overlapping network $E_i(\mathcal{O})$ and the participation coefficient P_i. Notice that there is indeed a slightly positive correlation between these metrics ($\tau(E_i(\mathcal{O}), P_i) = 0.31$, $r(E_i(\mathcal{O}), P_i) = 0.43$).

For each value of b, $\mathcal{M}$ is a square matrix with non-negative entries. Thus, with all the hypotheses of the Perron-Frobenius theorem being satisfied, we can calculate the eigenvector centrality of $\mathcal{M}$ as a function of b. In order to assess the role of each layer in determining the multiplex centrality, we follow this approach: We compute the eigenvector centrality of the benchmark case $\bar{b} = 0.5$ (corresponding to matrix $\mathcal{O}$); we then compute the eigenvector centrality of $\mathcal{M}$ for a generic value of b, and we evaluate the Kendall correlation coefficient τ_k between the centrality ranking obtained for a generic value of b and the benchmark case $\bar{b} = 0.5$. Since the multiplex network of the Indonesian terrorists has three layers, we can construct three different duplex networks. The results are shown in Fig. 10, where we plot the Kendall coefficient τ_k as a function of b.

As expected, the three duplex have a peak $\tau_k = 1$ for $b = 0.5$. By comparing the three curves we can deduce that T and O have a similar role in determining the centrality of the multilayer system, in both cases stronger than layer C. In fact, the slopes of the curves, as well as their symmetry (asymmetry), and the symmetry (asymmetry) of the extreme

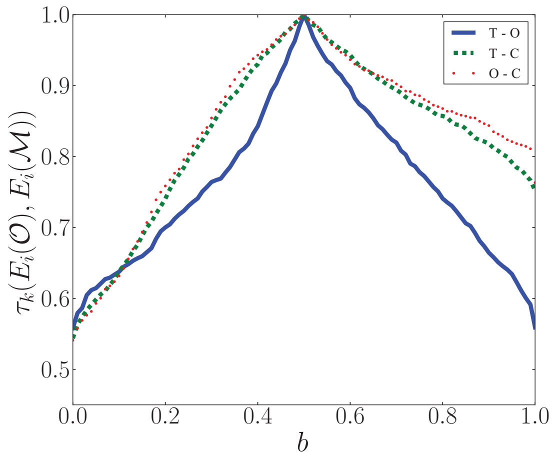

FIG. 10. (Color online) For each couple of layers (duplex) of the multiplex network of Indonesian terrorists we plot the Kendall correlation coefficient τ_k between the eigenvector centrality of the benchmark case ($b = 0.5$, i.e., equal weights on both layers) and the generic case of Eq. (26).

cases $b = 0$ and $b = 1$, tell us about the interplay between the two layers in determining the centrality of the multilayer system. The curve corresponding to the duplex T-O is quite symmetrical, indicating that the effect of T and O on the centrality is very similar. Conversely, the curves corresponding to T-C and O-C are asymmetrical. This means that both layers T and O dominate layer C in determining the centrality of the nodes. If we focus on the case $b = 0$, we obtain three similar values of τ_k. Instead, the three curves display different behavior in the range $0 \leqslant b \leqslant 0.5$. In particular, the solid blue curve shows the steepest decrease from the peak (this is also true for $b \geqslant 0.5$), indicating that layers T and O differ more than layers T and C or layers O and C. For this reason, a small perturbation of the coefficients of $\mathcal{M}$ from the benchmark case affects the centrality of the multilayer system more for the duplex T-O than for the duplexes T-C and O-C. The largest dissimilarity of the pair T-O is also confirmed by the smallest value of τ_k found for the couple $E_i^{[T]}$ and $E_i^{[O]}$, as shown in Fig. 8(d).

A slightly different approach provides useful insights about the distribution of centrality in the system under study. Given the three duplex networks, for each one of them we can compute the Kendall coefficient τ_k between the values of centrality obtained for $\mathcal{M}$ and different values of b and those obtained for each single layer. Results are shown in Fig. 11. We note that the value of $\tau_k(E_i^{[\alpha]}, E_i(\mathcal{M}(b = 0.5)))$ in each panel of Fig. 11 is equal, respectively, to the value of $\tau_k(E_i(\mathcal{O}), E_i(\mathcal{M}(b = 1)))$ for $\alpha = 1$ and to $\tau_k(E_i(\mathcal{O}), E_i(\mathcal{M}(b = 0)))$ for $\alpha = 2$ on the corresponding curve in Fig. 10. In Fig. 11(a) the two curves are quite symmetrical and intersect around $b = 0.5$, indicating that the contributions of layers T and O to centrality is similar. Conversely, for both T-C and O-C [respectively, Figs. 11(b) and 11(c)] the two curves are asymmetrical and intersect at $0.35 < b < 0.40$, indicating that both layer T and O have stronger impact on centrality than layer C.

These results indicate that multilayer systems are characterized by nontrivial organization also with respect to centrality. We conclude this section by noticing that the definition of

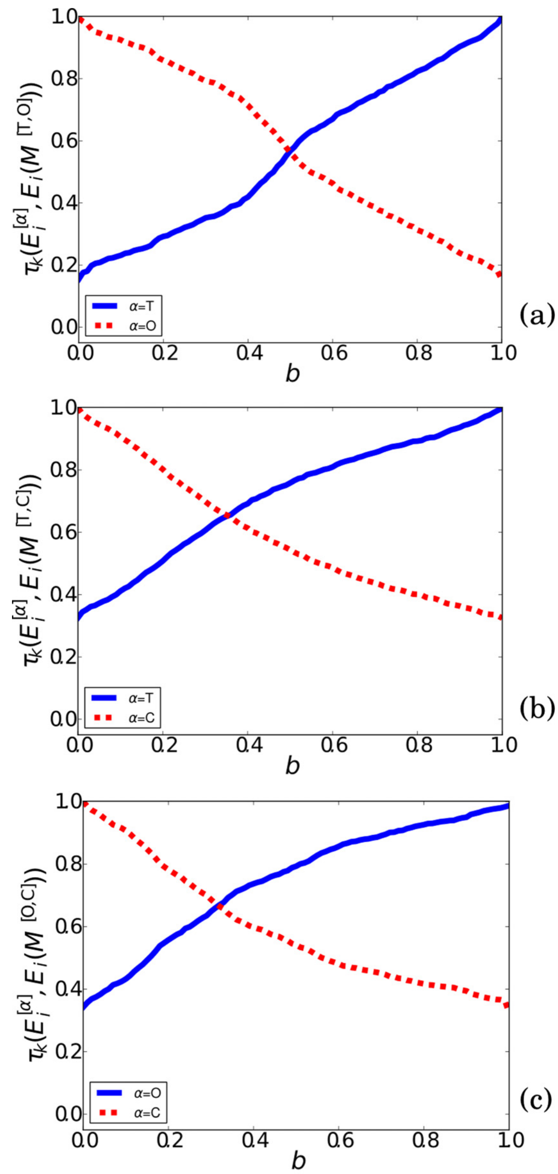

FIG. 11. (Color online) As subsets of the original overlapping network, we consider the three duplex $M^{[T,O]} = bA^{T} + (1-b)A^{[O]}$, $M^{[T,C]} = bA^{T} + (1-b)A^{[C]}$ and $M^{[O,C]} = bA^{O} + (1-b)A^{[C]}$. For each possible duplex, we report the Kendall coefficient τ_k between the centrality of each single layer and the corresponding $\mathcal{M}$ as a function of b.

multiplex centrality can be easily generalized to a system of M levels by constructing the following adjacency matrix:

$$\mathcal{M} = b_1 A^{[1]} + b_2 A^{[2]} + \cdots + b_M A^{[M]}, \qquad (27)$$

with the condition that $\sum_{i=1}^{M} b_i = 1$. Once again the benchmark case obtained by fixing $b_1 = \cdots = b_M = \frac{1}{M}$ coincides with the aggregated overlapping network.

IX. CONCLUSIONS

The basic units of many real-world systems are connected through a large variety of different relations. One of the new challenges in network theory is therefore to treat together ties of different kind preserving existing differences. The multiplex metaphor, which allows us to distinguish the different kinds of relationships among a set of nodes, constitutes a promising framework to study and model multilayer systems. In this paper we proposed a comprehensive formalism to deal with systems composed of several layers, both with binary or weighted links. In particular, we provided a clear distinction about the different levels of description of a multiplex network: the aggregated topological, the overlapping, and the weighted overlapping network, which are simpler but less rich structures than the vector of adjacency matrix $\boldsymbol{A}$. We also proposed a number of metrics to characterize multiplex systems with respect to node degree, edge overlap, node participation to different layers, clustering coefficient, reachability, and eigenvector centrality. All these measures were tested on the multiplex network of Indonesian terrorists, a system with 78 nodes and four layers. Admittedly, the notation proposed in this work, based on the explicit vectorial representation of node and edge properties, is just one of the possible ways of dealing with multiplex networks, and indeed there have been other recent attempts to define a consistent framework for the analysis and characterization of multilayer systems. In particular, the tensorial formalism proposed in Refs. [11,28] seems a promising approach, since it allows us to express some multiplex metrics in a synthetic and compact way. However, we believe that the notation we proposed here, which makes explicit the role of single layers, is somehow more immediately understandable and easier to use for the study of real-world multiplex networks. We really hope that the set of tools and metrics presented in this paper will trigger further research on the characterization of the structural properties of multilayer complex systems.

ACKNOWLEDGMENTS

This work was supported by the Project LASAGNE, Contract No. 318132 (STREP), funded by the European Commission.

[1] S. H. Strogatz, Nature **410**, 268 (2001).
[2] R. Albert and A.-L. Barabasi, Rev. Mod. Phys. **74**, 47 (2002).
[3] M. E. J. Newman, SIAM Rev. **45**, 167 (2003).
[4] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D. U. Hwang, Phys. Rep. **424**, 175 (2006).
[5] M. Kurant and P. Thiran, Phys. Rev. Lett. **96**, 138701 (2006).
[6] S. V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley, and S. Havlin, Nature **464**, 1025 (2010).
[7] A. Cardillo, J. Gomez-Gardenes, M. Zanin, M. Romance, D. Papo, F. del Pozo, and S. Boccaletti, Sci. Rep. **3**, 1344 (2013).
[8] M. Szell, R. Lambiotte, and S. Thurner, Proc. Natl. Acad. Sci. USA **107**, 13636 (2010).
[9] G. Bianconi, Phys. Rev. E **87**, 062806 (2013).
[10] A. Solé-Ribalta, M. De Domenico, N. E. Kouvaris, A. Díaz-Guilera, S. Gómez, and A. Arenas, Phys. Rev. E **88**, 032807 (2013).
[11] M. De Domenico, A. Solé-Ribalta, E. Cozzo, M. Kivelä, Y. Moreno, M. A. Porter, S. Gómez, and A. Arenas, Phys. Rev. X **3**, 041022 (2013).
[12] V. Nicosia, G. Bianconi, V. Latora, and M. Barthelemy, Phys. Rev. Lett. **111**, 058701 (2013).
[13] A. Saumell-Mendiola, M. A. Serrano, and M. Boguna, Phys. Rev. E **86**, 026106 (2012).
[14] E. Cozzo, R. A. Baños, S. Meloni, and Y. Moreno, Phys. Rev. E **88**, 050801(R) (2013).
[15] B. Min and K.-I. Goh, arXiv:1307.2967.
[16] J. Gomez-Gardenes, I. Reinares, A. Arenas, and L. M. Floria, Sci. Rep. **2**, 620 (2012).
[17] V. Nicosia, M. Valencia, M. Chavez, A. Diaz-Guilera, and V. Latora, Phys. Rev. Lett. **110**, 174102 (2013).
[18] S. Gomez, A. Diaz-Guilera, J. Gomez-Gardenes, C. J. Perez-Vicente, Y. Moreno, and A. Arenas, Phys. Rev. Lett. **110**, 028701 (2013).
[19] M. De Domenico, A. Sole, S. Gomez, and A. Arenas, arXiv:1306.0519.
[20] M. Kivelä, A. Arenas, M. Barthelemy, J. P. Gleeson, Y. Moreno, and M. A. Porter, arXiv:1309.7233.
[21] Throughout the paper we make the assumption that the graph at each layer is simple, i.e., without self-loops.
[22] N. Roberts and S. F. Everton, *Roberts and Everton Terrorist Data: Noordin Top Terrorist Network (Subset)* (2011).
[23] R. Guimerà and L. A. Nunes Amaral, J. Stat. Mech. (2005) P02001.
[24] R. Guimera, L. A. N. Amaral, Nature **433**, 895 (2005).
[25] D. Watts and S. Strogatz, Nature **393**, 440 (1998).
[26] V. Latora and M. Marchiori, Eur. Phys. J. B **32**, 249 (2003).
[27] S. Wasserman and K. Faust, *Social Networks Analysis* (Cambridge University Press, Cambridge, 1994).
[28] E. Cozzo *et al.*, arXiv:1307.6780.
[29] R. G. Morris and M. Barthelemy, Phys. Rev. Lett. **109**, 128703 (2012).
[30] P. Bonacich, Am. J. Sociol. **92**, 1170 (1987).
[31] L. Solá, M. Romance, R. Criado, J. Flores, A. García del Amo, and S. Boccaletti, Chaos **23**, 033131 (2013).