Silesian J. Pure Appl. Math.
vol. 10 (2020), 45–102

Piotr WILCZEK

Computer Laboratory, Poznań, Poland

A COMPARATIVE PERFORMANCE ANALYSIS OF THE EXPONENTIAL-BASED AND RESOLVENT-BASED CENTRALITY MEASURES

Abstract. The present study aims to quantitatively assess the effect of the attenuation factor on the resolution, performance, rank correlations, robustness and assortativity of the Katz centrality measure. We found that the granularity of the exponential-based and resolvent-based ranking algorithms is strongly correlated with the number of automorphically equivalent nodes within the network. Thus, this result can be viewed as a bridge between *algebraic* and *quantitative graph theory*. Moreover, we substituted the dichotomous adjacency matrix in the definitions of the exponential-based and resolvent-based centrality indices by its weighted (normalized) version and, therefore, we obtained two novel ranking algorithms. The deliberate attack simulation experiments carried out on four empirical and on two model networks showcased that the newly suggested ranking methods considerably outperform their unweighted counterparts as well as the classical degree centrality measure. In the last part of the paper, we introduced the concept of the *centrality assortativity profile* of a complex network. The extensive numerical results demonstrated that this novel theoretical notion is useful in complex network mining.

Mathematics Subject Classification: 05C82, 05C90.
Keywords: complex networks, centrality measures, attack simulations.
Corresponding author: P.Wilczek (piotr.wilczek.net@onet.pl).
Received: 25.08.2020.

1. Introduction

A long-standing open problem in *network science* is to develop algorithms capable to rank nodes in complex networks with respect to their relative significance. Such algorithms, in the form of centrality measures, should take into account the heterogeneity and function-oriented specificity of nodes in the network. Designing novel and more accurate node ranking methods affects our abilities to understand and control, for instance, pandemics, cascading failure processes, dissemination and diffusion phenomena as well as the robustness of critical infrastructures.

In the *network science* literature, there exist many different proposals how to identify influential nodes in complex networks [16, 17, 24, 25, 38]. For instance, a popular way to define the relative importance of a given node in a complex network relies on quantifying its abilities to initiate walks around the network. This concept leads to the use of different matrix functions of the general form $f(M(G))$ where $M(G)$ is some graph-theoretical matrix. If f is the resolvent function and $M(G)$ is the adjacency matrix, then we obtain the resolvent-based subgraph centrality measure [11] or the Katz centrality measure [5, 17, 29]. In turn, if f is the exponential function and $M(G)$ is the adjacency matrix, then we obtain the subgraph centrality measure [19] or the total communicability centrality measure [10].

Note that the oldest of the centrality metrics defined *via* the matrix function, i.e., the Katz centrality index, denoted by KC, is a parameter-dependent node significance ranking algorithm. This means that the values of this measure hinge on the choice of the attenuation factor (hereinafter referred to as the Katz parameter). However, the choices of this factor have not received much attention in the *network science* literature. Some insightful discussions are provided by M. Benzi and C. Klimko who scrutinized the behavior of the KC index as the attenuation factor decreases to zero or increases to $\frac{1}{\rho(A(G))}$ (where $\rho(A(G))$ is the spectral radius of the adjacency matrix of the network G) [11] and by M. Aprahamian et. al who analysed how the attenuation factor affects the linear and rank correlations between the KC index and the total communicability centrality measure, denoted by TC [5]. Nevertheless, as far as we know, the KC and TC indices were not tested with respect to their performance in dismantling the networks or with respect to their robustness and assortativeness. Consequently, in the light of the above, it seems desirable to carry out the computer simulation experiments in order to quantitatively assess the efficiency of the KC and TC indices in attacking the empirical networks as well as to estimate the reliability of both ranking methods in the face of measurement errors. Furthermore, it would be valuable to determine the assortativeness of the above centrality metrics. In all simulation experiments, the KC index will be calculated with four different values of the Katz parameter considered in the reference [5] and the resulting ranking lists will be comparatively evaluated. Thus, the present work will complement the analyses contained in [5].

As was stated before, the Katz and total communicability centrality measures are defined by the functions applied to the adjacency matrix of the network under consideration. As is well know, the entries of the adjacency matrix of any network indicate if pairs of nodes are connected or not in this network without providing any information regarding the strength of such connections. Therefore, we will propose to introduce the weighted (normalized) variants of the KC and TC indices. These modified vertex importance ranking algorithms are defined by applying the resolvent and exponential functions to the edge-weighted adjacency matrix developed in the field of *chemical graph theory* by S.B. Bozkurt and coworkers [12]. In the simulation experiments, we will juxtapose the original KC and TC measures with their newly suggested weighted counterparts.

The rest of the paper is structured in the following manner: Section 2 contains the necessary background on *graph theory* and on the centrality measures defined *via* the expression $f(M(G))$. Section 3 proposes the weighted (normalized) versions of the indices presented in Section 2. Section 4 describes the datasets and computational methodology used in Section 5. The numerical results and discussion are included in Section 5. Finally, Section 6 concludes the article.

2. Background

In the current work, all considered networks are represented by undirected graphs without loops or multiple edges of the general form $G = (V(G), E(G))$ where $V(G) = \{v_1, v_2, \ldots, v_n\}$ is the *vertex* (*node*) *set* and $E(G) = \{e_1, \ldots, e_m\}$ is the *edge* (*link*) *set*. For two vertices $v_i, v_j \in V(G)$, $v_i v_j$ means that v_i and v_j are *adjacent*, i.e., $v_i v_j \in E(G)$. A complex network can be represented by the *adjacency matrix* $A(G) = [a]_{ij}$ whose elements are given by the term $[a]_{ij} = 1$ if $v_i v_j \in E(G)$ and $[a]_{ij} = 0$ otherwise. The symbol k_i denotes the *degree* of the vertex $v_i \in V(G)$. The degree of v_i is the number of edges that are incident to the vertex v_i. The *degree centrality* corresponding to the node $v_i \in V(G)$, denoted by $DC(v_i)$, is given by $DC(v_i) = \frac{k_i}{n-1}$ or simply by $DC(v_i) = k_i$ [17, 38]. Several topological metrics have been introduced in order to study the way in which nodes with a given degree are connected within a complex network. A paradigmatic example of such metrics is the *degree assortativity coefficient* suggested by M. Newman in [37]. Namely, for a complex network $G = (V(G), E(G))$, the degree assortativity coefficient (also known as the *degree-degree correlation coefficient*), denoted by $A_k(G)$, is expressed by the following condition [47]:

$$A_k(G) := \frac{4\langle k_i k_j \rangle - \langle k_i + k_j \rangle^2}{2\left\langle k_i^2 + k_j^2 \right\rangle - \langle k_i + k_j \rangle^2}, \tag{1}$$

where $v_i v_j \in E(G)$ and $\langle\rangle$ denotes the arithmetic mean over all edges in G. Note that the equation (1) is simply the Pearson correlation coefficient of the degrees at both ends of a link. The network G in which high-degree nodes tend to be connected to each other possesses a positive value of $A_k(G)$ and is said to be *degree-assortative*. In turn, the network G in which high-degree nodes tend to constitute edges with low-degree nodes has a negative value of $A_k(G)$ and is said to be *degree-disassortative*. When $A_k(G) \approx 0$, the network is said to be *degree-neutral* [17, 38].

The degree assortativity coefficient can be easily generalized to any centrality measure C [35]. Accordingly, for a complex network $G = (V(G), E(G))$, the *centrality assortativity coefficient* (also known as the *centrality-centrality correlation coefficient*), denoted by $A_C(G)$, is identified with the subsequent formula

$$A_C(G) := \frac{4\langle c_i c_j \rangle - \langle c_i + c_j \rangle^2}{2\left\langle c_i^2 + c_j^2 \right\rangle - \langle c_i + c_j \rangle^2} \tag{2}$$

where $v_i v_j \in E(G)$, c_i is the value of any centrality measure C at the node $v_i \in V(G)$ and $\langle\rangle$ denotes the arithmetic mean over all links in G.

For a complex network G, we adopt the ranges of the indices defined by the equations (1) and (2) as well as the corresponding *degree* and *centrality assortativity levels* of G from the reference [35]. They are included in Table 1.

The degree assortativity of complex networks is studied in many articles (cf. the review paper [40] and the references cited therein), whereas the assortativity determined by other centrality measures has received much less attention. As far as we know, only several references consider the assortativity of complex networks induced by other than the degree node importance ranking algorithms (cf. [3, 6, 20, 32, 34, 35]).

Table 1
The ranges of the degree-degree or centrality-centrality correlation coefficient and the corresponding degree or centrality assortativity levels

Range	Assortativity level	Range	Assortativity level
0.0 to 0.19	neutral	-0.19 to -0.01	neutral
0.2 to 0.59	weakly assortative	-0.59 to -0.20	weakly disassortative
0.6 to 1	strongly assortative	-1 to -0.60	strongly disassortative

In Subsection 5.5, we will estimate the centrality assortativity coefficients (according to the equation (2)) induced by the exponential-based and resolvet-based centrality measures of the empirical and generated networks and we will compare the obtained results with the degree assortativity coefficient of these networks. We will try to corroborate the hypothesis that the centrality assortativity coefficients conditioned by the centrality metrics defined *via* the matrix functions under consideration provide valuable information concerning the structural characteristics of the networks.

The *spectral radius* of the adjacency matrix, denoted by $\rho(A(G))$, is identified with the largest absolute value of its eigenvalues [17]. A *walk* of length l in a network $G = (V(G), E(G))$ is a list of (not necessary different) nodes $v_i, v_2, \ldots, v_l, v_{l+1}$ such that for each $i = 1, 2, \ldots, l$ there is a link from v_i to v_{i+1}. A *closed walk* of length l is a walk $v_i, v_2, \ldots, v_l, v_{l+1}$ in which $v_{l+1} = v_l$ [17]. For a complex network $G = (V(G), E(G))$ represented by the adjacency matrix $A(G)$ and for any positive l, the entry $\left[a^l\right]_{ij}$ of the matrix $A^l(G)$ is equal to the number of walks of length l that start at the vertex v_i and end at the vertex v_j [17]. For a network $G = (V(G), E(G))$ with the vertex set $V(G)$ where $|V(G)| = n$, its *Randić matrix* (also known as the *product connectivity matrix*), denoted by $R(G) = [r]_{ij}$, is the square real matrix of order n whose entries are given by the term $[r]_{ij} = (k_i k_j)^{-0.5}$ if $v_i v_j \in E(G)$ and $[r]_{ij} = 0$ otherwise [12]. The quantity $(k_i k_j)^{-0.5}$ assigned to the edge $v_i v_j \in E(G)$ is known as its *Randić weight.* The *Randić degree* of the node $v_i \in V(G)$, denoted by rk_i, is given by the formula: $rk_i = \sum\limits_{j=1}^{n} [r]_{ij}$. The *Randić eigenvalues* of the network G are the eigenvalues of its Randić matrix $R(G)$. The *Randić spectral radius* of G, denoted by $\rho(R(G))$, is the largest absolute value of its Randić eigenvalues. In [31], it is proved that for any network G, its Randić spectral radius is always equal to one.

As previously mentioned, a popular way of defining a centrality measure for a complex network G is to use the functions of some graph-theoretical matrices

$M(G)$. The present paper focuses on the Katz centrality measure KC (defined *via* the matrix resolvent function) and on the total communicability centrality measure TC (defined *via* the matrix exponential function). Recall that for a complex network $G = (V(G), E(G))$ with the vertex set $V(G)$ where $|V(G)| = n$, the *Katz centrality* of the vertex $v_i \in V(G)$ is given by the following expression

$$KC_i(\alpha) := \sum_{j=1}^{n} \left[(I - \alpha\, A(G))^{-1} \right]_{ij}, \tag{3}$$

where the term $(I - \alpha A(G))^{-1}$ is the matrix resolvent (I refers to the *identity matrix* of order n) and α is the attenuation parameter (known as the *Katz parameter*) which satisfies the condition $0 < \alpha < \frac{1}{\rho(A(G))}$. The resolvent in the equation (3) has the following power series expansion

$$\begin{aligned}(I - \alpha\, A(G))^{-1} = I + \alpha\, A(G) + \alpha^2 A^2(G) + \ldots + \alpha^l A^l(G) + \ldots = \\ = \sum_{l=0}^{\infty} \alpha^l A^l(G).\end{aligned} \tag{4}$$

The restrictions imposed on α ensure that the matrix $I - \alpha\, A(G)$ is invertible and that the power series expansion given by the equation (4) converges to its inverse. These restrictions also force that the resolvent in the equation (3) is nonnegative as $I - \alpha\, A(G)$ is a nonsingular M-matrix, i.e., a matrix of the form $W = qI - B$, where $B = [b]_{ij}$ with $b_{ij} \geq 0$ for all $1 \leq i, j \leq n$ and q exceeds the dominant eigenvalue of B (for the general properties of matrix functions and M-matrices, we refer to [22, 30, 43]). Consequently, the row sums of $(I - \alpha\, A(G))^{-1}$ are nonnegative and can be used to rank vertices in the networks with respect to their relative significance. From the equations (3) and (4), it follows that the Katz centrality score of the vertex $v_i \in V(G)$ counts all walks that originate at v_i and simultaneously penalizes the contribution of walks of length l by the factor α^l. The Katz centrality measure is a parameter-dependent node importance ranking algorithm. Accordingly, this node significance ranking algorithm has infinitely many numerical instances. In the *network science* literature, several different choices of the Katz parameter have been suggested and, consequently, several numerical instances of the KC index have been used. The most frequently encountered choices of α were discussed and tested in [5]. These choices are expressed by the subsequent formulae

$$\alpha_{0.5} = \frac{1}{2\rho\left(A\left(G\right)\right)}, \tag{5}$$

$$\alpha_{0.85} = \frac{0.85}{\rho\left(A\left(G\right)\right)}, \tag{6}$$

$$\alpha_{k} = \frac{1}{\left\|A\left(G\right)\right\|_{\infty}+1}, \tag{7}$$

and

$$\alpha_{\min} = \frac{1-\exp^{-\rho(A(G))}}{\rho\left(A\left(G\right)\right)}. \tag{8}$$

Recall that in undirected and unweighted networks, the ∞-norm of $A\left(G\right)$ in the equation (7) is equal to the largest node degree. The justification for the choice of $\alpha_{\min}$ discussed in depth in [5] is as follows: Suppose that $KC\left(\alpha\right)\left(G\right)$ is the vector of the Katz centrality scores calculated on the network $G=\left(V\left(G\right),E\left(G\right)\right)$ with the vertex set $V\left(G\right)$ where $\left|V\left(G\right)\right|=n$ under the parameter α, i.e.,

$$KC\left(\alpha\right)\left(G\right) := \left(I-\alpha A\left(G\right)\right)^{-1}\mathbf{1}, \tag{9}$$

where $\mathbf{1}=\left[1,1,\ldots,1\right]^{T}$ and $TC\left(G\right)$ is the vector of the total communicability centrality scores calculated on the same network G, i.e.,

$$TC\left(G\right) := \exp^{A(G)}\mathbf{1}, \tag{10}$$

where the term $\exp^{A(G)}$ is the matrix exponential discussed below. The authors of the reference [5] proposed to determine the Katz parameter α that aims to minimize the 2-norm of the difference between the vectors given by the equations (9) and (10). This is equivalent to find α that solves

$$\min_{\alpha} err\left(\alpha\right) = \min_{\alpha}\left\|TC\left(G\right)-KC\left(\alpha\right)\left(G\right)\right\|_{2}, \tag{11}$$

where the 2-norm is given by the condition $\left\|x\right\|_{2}=\left(x^{T}x\right)^{0.5}$. Accordingly, the approach from [5] naturally forces that the centrality scores obtained from the Katz centrality vector $\left(I-\alpha A\left(G\right)\right)^{-1}\mathbf{1}$ (the equation (9)) and the action of the exponential of the adjacency matrix $\exp^{A(G)}\mathbf{1}$ (the equation (10)) are similar.

By contrast, applying the matrix exponential to the adjacency matrix of a complex network $G=\left(V\left(G\right),E\left(G\right)\right)$ with the vertex set $V\left(G\right)$ where $\left|V\left(G\right)\right|=n$

gives rise to the subsequent power series expansion of the term $\exp^{A(G)}$ [10]:

$$\exp^{A(G)} = I + A(G) + \frac{A^2(G)}{2!} + \frac{A^3(G)}{3!} + \ldots + \frac{A^l(G)}{l!} + \ldots = \\ = \sum_{l=0}^{\infty} \frac{A^l(G)}{l!}. \quad (12)$$

Then, the so-called *total (subgraph) communicability centrality measure* TC of the vertex $v_i \in V(G)$ is identified with the condition [10]:

$$TC_i := \sum_{j=1}^{n} \left[\exp^{A(G)} \right]_{ij}. \quad (13)$$

The power series expansion given by the equation (12) and the definition of the TC index given by the equation (13) imply that this node significance ranking algorithm counts all walks between a focal node v_i and all other nodes in the network (with the node v_i included) and weights walks of length l by a penalty factor equal to $\frac{1}{l!}$. Accordingly, the TC index evaluates the relative importance of nodes in a network based on their communicability with other nodes in this network [10].

3. The weighted (normalized) exponential-based and resolvent-based ranking algorithms

For a complex network $G = (V(G), E(G))$ with the vertex set $V(G)$ where $|V(G)| = n$, the reference [14] proposes to define the communicability between two arbitrary vertices $v_i, v_j \in V(G)$ by the matrix exponential function applied to the weighted (normalized) adjacency matrix, denoted by $W(G)$, of the network G. This matrix is defined by the condition

$$W(G) := D^{-0.5}(G)\, A(G)\, D^{-0.5}(G), \quad (14)$$

where $D(G) \in \mathbb{R}^{\{n \times n\}}$ is the degree matrix. For the network G, this matrix has the form $D(G) = diag(k_i)$. Note that from the fact the degree matrix $D(G)$ is diagonal and positive, it can be easily observed that its reciprocal square root $D^{-0.5}(G)$ is just the diagonal matrix whose diagonal elements are the reciprocals of the square roots of the diagonal entries of $D(G)$. Based on the concept of the communicability between two arbitrary nodes in the network G defined *via* the normalized adjacency matrix, the reference [4] proposes the centrality measure which was shown to be useful in mining the brain connectivity complex networks. It follows that for any complex network G, the normalized adjacency matrix $W(G)$ given by the equation (14) coincides with the Randić matrix $R(G)$ defined in [12]. The conceptual motivation which leads the authors of the reference [14] to evaluate the exponential on the weighted adjacency matrix in order to define the communicability between nodes can be summarized as follows: For a complex network $G = (V(G), E(G))$ with the vertex set $V(G)$ where $|V(G)| = n$, the entry $\left[a^l\right]_{ij}$ of the l-th power of the adjacency matrix $A(G)$ given by the condition

$$\left[a^l\right]_{ij} = \sum_{r_1=1}^{n} \sum_{r_2=1}^{n} \ldots \sum_{r_k=1}^{n} a_{i,r_1}\, a_{r_1,r_2}\, a_{r_2,r_3} \ldots a_{r_{l-1},r_l}\, a_{r_l,j} \tag{15}$$

counts the number of walks of length l that originate at the vertex v_i and finish at the vertex v_j. From the fact that the binary adjacency matrix $A(G)$ considers only the presence (given by the entry equal to one) or the absence (given by the entry equal to zero) of a link between each pair of nodes in the network G, it follows that it is impossible to express the connectivity strength of this link solely on the basis of $A(G)$. In turn, many empirical network datasets contain the connectivity information in the form of real-valued, non-negative weights, where a larger weight of the link $v_i v_j \in E(G)$ indicates that the nodes v_i and v_j are more strongly connected. Consequently, in such cases, the equation (15) remains valid but now the term $a_{i,r_1}\, a_{r_1,r_2}\, a_{r_2,r_3} \ldots a_{r_{l-1},r_l}\, a_{r_l,j}$ in this equation does not give the dichotomous contribution depending on whether the walk $i \to r_1 \to r_2 \to r_3 \to \ldots \to r_l \to j$ takes place. Instead, the equation (15) contributes the product of the weights along all edges in the walk and simultaneously decreases the contribution stemmed from longer walks. The references [4,14] contain several empirical evidences that the above summarized methodology gives rise to the more realistic results.

The present paper proposes to define two novel centrality measures based on the exponential and resolvent functions evaluated on the product connectivity matrix $R(G)$ of the network G. The first ranking algorithm that we propose is the *Randić-Katz centrality measure* (alternatively termed as the *weighted* or *normalized Katz centrality measure*), denoted by RKC. For a complex network $G = (V(G), E(G))$ with the vertex set $V(G)$ where $|V(G)| = n$, the RKC index of the vertex $v_i \in V(G)$ is given by the subsequent expression

$$RKC_i := \sum_{j=1}^{n} \left[\left(I - \alpha\, R(G)\right)^{-1} \right]_{ij}, \tag{16}$$

where I is the identity matrix of order n and α is the *weighted Katz parameter* which satisfies the condition $0 < \alpha < \frac{1}{\rho R(G)}$. Since $\rho(R(G)) = 1$ for any complex network [31], it follows that $0 < \alpha < 1$. Note that the bounds imposed on α ensure that the matrix $I - \alpha\, R(G)$ is invertible and the power series expansion of the resolvent in the equation (16):

$$\left(I - \alpha\, R(G)\right)^{-1} = I + \alpha\, R(G) + \alpha^2\, R^2(G) + \ldots + \alpha^l\, R^l(G) + \ldots = \\ = \sum_{l=0}^{\infty} \alpha^l\, R^l(G) \tag{17}$$

is convergent to its inverse. In our computer simulation experiments presented in Section 5, we will test the performance of the newly proposed Randić-Katz centrality measure calculated with four different values of α. Namely, by analogy with the values of the Katz parameter given by the equations (5)–(8), we will test the RKC index calculated with following four weighted Katz parameters

$$\alpha_{0.5} = \frac{1}{2\rho(R(G))} = \frac{1}{2}, \tag{18}$$

$$\alpha_{0.85} = \frac{0.85}{\rho(R(G))} = 0.85, \tag{19}$$

$$\alpha_{rk} = \frac{1}{\|R(G)\|_\infty + 1}, \tag{20}$$

and

$$\alpha_{\min} = \frac{1 - \exp^{-\rho(R(G))}}{\rho(R(G))} = 1 - \exp(-1) = 1 - 0.3679 = 0.6321, \tag{21}$$

where $\rho(R(G))$ is the Randić spectral radius (always equal to one [31]) and rk is the Randić degree. Note that in undirected networks, the ∞-norm of the product connectivity matrix in the equation (20) is equal to the largest Randić degree of this network.

The second ranking method that we suggest is the *Randić total communicability centrality measure* (alternatively termed as the *weighted* or *normalized total communicability centrality measure*), denoted by RTC. For a complex network $G = (V(G), E(G))$ with the vertex set $V(G)$ where $|V(G)| = n$, the Randić total communicability centrality measure of the vertex $v_i \in V(G)$ is given by the following formula

$$RTC_i := \sum_{j=1}^{n} \left[\exp^{R(G)}\right]_{ij}. \tag{22}$$

Note that the measure defined by the equation (22) is different from the centrality index proposed in [4] which is also based on the term $\exp^{R(G)}$. Namely, the centrality index introduced in [4] can be viewed as a vertex-deleted measure, i.e., it quantifies the reduction in the global communicability of the network G if a single vertex is removed.

The conceptual justification for using in the equations (16) and (22) the normalized adjacency matrix $R(G)$ instead of the dichotomous adjacency matrix $A(G)$ is the same as the reasoning included in [14]. Namely, from the fact that the Randić matrix of any complex network contains the connectivity information reflecting the strength of connections between its nodes, it can be hypothesized that the centrality measures defined by the equations (16) and (22) will be more effective in identifying influential nodes in complex networks than their $A(G)$-based counterparts. A similar methodological approach in defining new centrality measures was adopted by the authors of [42] who claimed that "*[b]y dichotomizing the network, much of the information contained in a weighted network dataset is lost, and consequently, the complexity of the network topology cannot be described to the same extent or richly*".

Since the matrix resolvent $(I-\alpha\, M(G))^{-1}$ and the matrix exponential $\exp^{M(G)}$, where $M(G) \in \{A(G), R(G)\}$ can be viewed as two alternative network communicability functions, we will collectively refer to the ranking algorithms given by the equations (3), (13), (16) and (22) as the *communicability-based centrality indices*. Moreover, the ranking methods given by the equations (3) and (13) are collectively referred to as the *adjacency-based centrality indices* (shortly, as the $A(G)$-*based measures*) or as the *original communicability-based centrality measures*, whereas the ranking algorithms given by the equations (16) and (22) are collectively referred to as the *Randić matrix-based centrality measures* (shortly, as the $R(G)$-*based measures*) or as the *weighted (normalized) communicability-based centrality measures*.

In order to illustrate how the newly defined indices work, consider the small graph G in Figure 1.

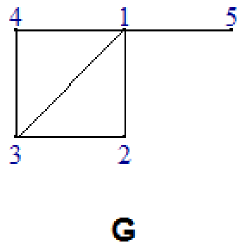

Fig. 1. The sample network G with five nodes

The product connectivity matrix of the graph G has the form

$$R(G) = \begin{bmatrix} & 1 & 2 & 3 & 4 & 5 \\ 1 & 0 & 0.3536 & 0.2887 & 0.3536 & 0.5 \\ 2 & 0.3536 & 0 & 0.4082 & 0 & 0 \\ 3 & 0.2887 & 0.4082 & 0 & 0.4082 & 0 \\ 4 & 0.3536 & 0 & 0.4082 & 0 & 0 \\ 5 & 0.5 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

The largest Randić degree of G is equal to 1.4958. Therefore, the α_{rk} parameter is equal to 0.4007 and the resolvent $(I - \alpha_{0.4007} R(G))^{-1}$ is identified with the following two-dimensional array

$$(I - \alpha_{0.4007} R(G))^{-1} = \begin{bmatrix} & 1 & 2 & 3 & 4 & 5 \\ 1 & 1.1211 & 0.1902 & 0.1919 & 0.1902 & 0.2246 \\ 2 & 0.1902 & 1.0605 & 0.2054 & 0.0605 & 0.0381 \\ 3 & 0.1919 & 0.2054 & 1.0894 & 0.2054 & 0.0384 \\ 4 & 0.1902 & 0.0605 & 0.2054 & 1.0605 & 0.0381 \\ 5 & 0.2246 & 0.0381 & 0.0384 & 0.0381 & 1.0450 \end{bmatrix}.$$

By applying the row sum operation to the above resolvent, we obtain the following scores of the RKC (α_{rk}) index: RKC_1 $(\alpha_{0.4007}) = 1.9181$, RKC_2 $(\alpha_{0.4007}) = 1.5548$, RKC_3 $(\alpha_{0.4007}) = 1.7306$, RKC_4 $(\alpha_{0.4007}) = 1.5548$, RKC_5 $(\alpha_{0.4007}) = 1.3843$. Similarly, the RKC index can be calculated for other values of α. In turn, the exponential of $R(G)$ has the form

$$\exp^{R(G)} = \begin{bmatrix} & 1 & 2 & 3 & 4 & 5 \\ 1 & 1.3409 & 0.4762 & 0.4929 & 0.4762 & 0.5542 \\ 2 & 0.4762 & 1.1705 & 0.5136 & 0.1705 & 0.1059 \\ 3 & 0.4929 & 0.5136 & 1.2508 & 0.5136 & 0.1033 \\ 4 & 0.4762 & 0.1705 & 0.5136 & 1.1705 & 0.1059 \\ 5 & 0.5542 & 0.1059 & 0.1033 & 0.1059 & 1.1316 \end{bmatrix}.$$

Then, by applying the row sum operation, we obtain the subsequent scores of the RTC index: $RTC_1 = 3.3403$, $RTC_2 = 2.4366$, $RTC_3 = 2.8742$, $RTC_4 = 2.4366$, $RTC_5 = 2.0008$.

It is worth noting that the values of the TC index calculated on the empirical complex networks are in an unreasonable numerical range. For instance, the minimum and maximum values of the TC measure for the *US airport 2010* network (cf. Table 2) are equal to $7.6077e + 41$ and to $1.2388e + 49$, respectively. In turn, the minimum and maximum values of the newly suggested RTC measure for the same network are equal to 1.1644 and 7.4177, respectively. Therefore, it is apparent that the novel RTC index is more handy in prioritizing nodes in real-world complex networks than its unweighted counterpart.

4. The datasets and computational methodology

Our numerical simulation experiments testing the performance of the communicability-based centrality measures are carried out on four real-world complex networks. The *Roget* network (the semantic network) is based on *Roget's Thesaurus* of English [9], the *E-mail* network (the communication network) is the network of e-mail exchanges between members of the University Rovira and Virgili [21], the *US political blogs* network (the web network) is the network of hyperlinks between weblogs on US politics [1], the *US airport 2010* network (the transportation network) is the network of flights among all commercial airports in US in 2010 derived from the U.S. Bureau of Transportation Statistics Transtats site [41]. All networks are treated as undirected and unweighted. If the network is disconnected, then only its giant component is used in the simulation experiments. The basic statistical parameters of these actual networks are listed in Table 2. Recall that many real-world networks are claimed to be *scale-free* [17, 38]. Therefore, we decided to carry out our simulation experiments on four actual networks whose degree sequences follow the power law [16, 46]. This means that the probability of finding a node having k links to other nodes decays as a negative power of the degree, i.e., $P(k) \sim k^{-\gamma}$. The applicability of the original communicability-based indices as well as their newly suggested normalized counterparts is not restricted to a specific domain. Consequently, to validate their performance we selected the exemplary networks from the recent *network science* literature and from different disciplines. Thus, the semantic network *Roget* belongs to a wider class of conceptual graphs; the communication network *E-mail* can be classified as a social network [38]; the web graph *US political blogs* can be viewed as an informational network, whereas the transportation network *US airport 2010* can be reckoned as an infrastructural network. These complex networks possess a wide variety of statistical properties (cf. Table 2) and together constitute a rather heterogeneous sample of empirical graphs. Moreover, these networks were frequently used to test the effectiveness of other recently proposed centrality measures [16, 24, 25, 46]. Besides the real-world networks, we will also validate the performance of the communicability-based centrality indices on two typical generative models: the *Erdős-Rényi random graph model*, denoted by *ER*, and the *Barabási-Albert random graph model*, denoted by *BA* [15, 17]. The *ER* model is constructed using the $G(n, m)$ scheme where n is the number of nodes and m is the number of links in the resulting graph. In turn, to construct the *BA* model, we start with a single node in the first time step. Then, we add one vertex in each time step and the new vertex initiates m edges to old vertices according to the preferential attachment mechanism. In this paper, we set m to 7. The basic statistical parameters of the *ER* and *BA* networks are also contained in Table 2. In the deliberate attack simulation experiments, we will use a single realization of each type of random network. This is motivated by the fact that the authors of [26] observed "[...] *that there is very little variance in the values of the V-index* [i.e., the robustness index] *obtained from different network realizations. Thus, the robustness results obtained from a single realization of a given type of network provide a true picture of the general robustness of networks of that type*". In Subsection 5.1, we will use the ensemble composed of 100 random undirected trees generated according to the *random citation growing graph model* [15]. This ensemble is denoted by *RCGG* and its basic statistical parameters are listed in Table 2.

Moreover, in our comparative analyses, we will use as null models four randomly rewired model networks corresponding to four empirical networks. These model networks have the degree sequences identical with their real-world counterparts and are denoted by *roget.rand*, *email.rand*, *blogs.rand* and *airport.rand*. In order to generate these models, we employed the algorithm proposed by F. Viger

Table 2

The statistical parameters of four empirical networks and three generative model networks. In the case of the RCGG model, the results are averages based on 100 simulation trials

Index	*Roget*	*E-mail*	*US political blogs*	*US airport 2010*	*ER*	*BA*	*RCGG*
n	994	1133	1222	1572	1000	1000	100
m	3640	5451	16714	17214	6972	6972	99
$\langle k \rangle$	7.3239	9.6222	27.3552	21.9008	13.944	13.944	1.98
$k_{\max}$	28	71	351	314	27	452	7.54
V_k	0.6638	0.9706	1.4038	2.0309	0.263	2.5343	0.6806
G_k	0.3557	0.4911	0.622	0.743	0.1486	0.4515	0.3264
γ	1.4831	1.4584	1.0694	1.0464	–	0.9536	1.9133
I_{TOP}	9891.4	11436.8	12346.5	15860.4	9965.8	9965.8	622.7
$\overline{orb}$	1.003	1.0244	1.0489	1.1973	1	1	1.2
C	0.1338	0.1663	0.226	0.3841	0.0125	0.0439	0
$Q\,(w)$	0.4638	0.5307	0.4253	0.2255	0.1891	0.1001	0.7401

and M. Latapy. This algorithm is implemented in [15] and described in [51, p. 24]. The basic statistical parameters of the model networks are included in Table 3. In Tables 2 and 3, the symbols $\langle k \rangle$, $k_{\max}$, V_k, G_k, γ, I_{TOP}, $\overline{orb}$, C and $Q\,(w)$ refers to the average degree, the maximum degree, the coefficient of variation (i.e., the ratio of the standard deviation to the mean) of the degree distribution [8], the Gini index of the degree distribution [8], the power law exponent of the degree distribution [8], total topological information content, the average number of nodes *per* one orbit of $Aut\,(G)$, the global clustering coefficient (i.e., the ratio of triangles and connected triples) and the modularity with respect to the walktrap community finding algorithm [15].

In order to evaluate the *resolution* (i.e., the *granularity*) of the tested ranking methods, the *monotonicity*, denoted by M, and the *percentage of uniquely defined metric*, denoted by PUD, of a ranking vector R are used. Both resolution metrics are described in [51, p. 20]. Recall that for any ranking vector R, $M(R) \in [0,1]$ and $PUD(R) \in [0\,\%, 100\,\%]$. The higher the values of both metrics are, the more granular the ranking vector R is.

One way of evaluating the performance of a vertex significance ranking algorithm is based on removing highest ranked (with respect to this algorithm) nodes and observing the remaining network [2, 23–26, 51]. In the *network science* litera-

Table 3
The statistical parameters of four randomly rewired model networks with the preserved degree distribution. The results are averages based on 50 simulation trials

Index	*roget.rand*	*email.rand*	*blogs.rand*	*airport.rand*
I_{TOP}	9893.5	11458.5	12477.1	16288.4
$\overline{orb}$	1.0019	1.0158	1.0205	1.1067
C	0.0126	0.0281	0.1492	0.2343
$Q\,(w)$	0.2173	0.122	0.028	0.0283

ture, this procedure is known as the *deliberate* (i.e., *intensional*, *targeted*) *attack* on the network. Accordingly, we will quantify the performance of the tested ranking methods in identifying influential nodes in complex networks by deleting one by one the top 20 % of the nodes from six exemplary networks according to the importance ranking lists generated by these methods. Thus, in each step of the attack, we will delete only one node from the network according to the centrality-induced ranking list and then we will calculate on the post-attack network the following parameters

1) the number of isolated nodes,

2) the number of connected components,

3) the relative size of the giant component,

4) the decline rate of the network efficiency.

These evaluation parameters, after the removal of the top-k most important vertices, are denoted by $Iso_k(G)$, $C_k(G)$, $r(k)$ and $\varepsilon(k)$, respectively. The mathematical definitions of the $r(k)$ and $\varepsilon(k)$ parameters are contained in [51, pp. 21–22]. We will repeat the procedure of removing the top-k most significant vertices for k ranging from 1 to about 20 % of the node in each exemplary network and after each step of the attack, we will record the above four evaluation parameters. For the sake of comparison, after all steps of the attack, the values of the $Iso_k(G)$ and $C_k(G)$ parameters are summed up. Accordingly, the quantities $\sum_{k=1}^{k} Iso_k\,(G)$ and $\sum_{k=1}^{k} C_k\,(G)$ represent the accumulation of the network damages caused by the attack strategy after k steps of the attack (cf. Tables 5, 6, 9 and 10). Undoubtedly, the higher both cumulative quantities are, the more efficient the attack scenario is. In the *network science* literature, the plots of the functions $r(k)$ and $\varepsilon(k)$ are known as the *robustness curve* and the *efficiency decline curve*, respectively. Such plots corresponding to the attack schemes guided by the tested centrality measures are included in Figures 3–6. Besides the visual comparison of the $r(k)$ and $\varepsilon(k)$ plots generated by the ranking algorithms under consideration, we will also calculate the areas under the robustness and efficiency decline curves corresponding to each attack protocol (cf. Tables 7, 8, 11 and 12). Undoubtedly, the lower (higher) the area under the robustness (the efficiency decline curve) is, the better the ranking method is.

In order to estimate the ranking consistency of the centrality indices under study, we will use the Kendall τ rank correlation coefficient. The mathematical definition of this metric can be found in [51, pp. 22-23] with the appropriate ranges indicating the strength of the relationship between two rankings.

In order to evaluate the robustness of the node importance ranking algorithms considered in the present work in the face of three types of the random link errors, we will follow the methodology used in [33, 39]. In the present study, we will focus on three random link errors models. In a nutshell, each model relies on the following procedure: We start with the *true* (i.e., *original*, *error-free*) network and calculate the centrality index on it. This centrality index is referred to as the *true* (i.e., *original*, *error-free*) *centrality index* and is denoted by C_{true}. Then, we perturb (i.e., distort) the network and calculate the new centrality index on it. This new centrality index is referred to as the *noisy* (i.e., *perturbed*, *distorted, erroneous*) *centrality index* and is denoted by C_{noisy}. Finally, we measure the Kendall τ rank correlation coefficient between the true and noisy centrality scores, i.e., $\tau\left(C_{true}, C_{noisy}\right)$. Thus, in the current paper, we will identify the robustness of the vertex significance ranking algorithm C with the values of $\tau\left(C_{true}, C_{noisy}\right)$. Undoubtedly, the higher the values of $\tau\left(C_{true}, C_{noisy}\right)$ are, the more robust the centrality measure C is. In our simulation experiments, we will perturb the network according to three models (mechanisms). In the first model, we will delete from the original network k randomly chosen edges where k ranges from 0.5 % to 10 % (with the step of 0.05 %). This mechanism is called the *random edge removal model.* In the second model, we will add to the network k randomly chosen edges where k ranges from 0.5 % to 10 % (with the step of 0.05 %). This mechanism is called the *random edge addition model.* In the third model, we will randomly rewire edges with the probability p ranging from 0.005 to 0.1 with the step of 0.005. This mechanism is called the *random rewiring model.* In all cases, we will also measure the Pearson correlation coefficient r between the intensity of the perturbing factor and the variation in the values of $\tau\left(C_{true}, C_{noisy}\right)$.

All computations included in the present paper were conducted in the R programming language [7, 13, 15, 27, 28, 36, 44, 49].

5. The results and discussion

In this section, we will comparatively juxtapose the original and weighted communicability-based centrality measures with respect to their resolution, efficiency in identifying crucial nodes in complex networks, rank correlations, robustness and assortativeness. In our comparative analyses, the classical DC measure will be used as a reference metric. Our analyses are focused on the centrality indices defined by the equations (3), (13), (16) and (22). The Katz index is calculated with respect to the values of the attenuation parameter given by the equations (5)–(8), whereas the Randić-Katz measure is calculated according to the values of the weighted attenuation parameter given by the equations (18)–(21). These values are listed in Tables 1 and 2 in [52]. Thus, we will consider four numerical instances of the KC index as well as four numerical instances of the RKC index.

5.1. The resolution of the tested ranking methods

The resolution quantified by the M and PUD metrics of the communicability-based centrality indices as well as the DC measure evaluated on the empirical and model networks is summarized in Table 4. The detailed measurements are contained in Tables 3–12 in [52].

These data indicate that the granularity of the original communicability-based indices is (approximately) equal to the granularity of their normalized counterparts. Moreover, in all cases, the resolution of all communicability-based ranking methods is higher in the rewired datasets than in the actual datasets. This property is especially noticeable when the granularity is estimated by the PUD metric. In turn, in both generative models, all communicability-based centrality indices are perfectly monotonous. The values of the Katz parameter given by the equations (5)–(8) and its weighted variant given by the equations (18)–(21) have no influence on the resolution of the resolvent-based measures (cf. Tables 3–10 in [52]). Moreover, it can be observed that, in comparison to the resolution attained by the communicability-based centrality indices, the granularity of the DC measure

Table 4
The numerical ranges of the monotonicity and uniquely defined metric quantifying the granularity of the tested ranking methods. The symbol − denotes the dash

Index	Empirical networks		Rewired networks		*ER*		*BA*	
	M	*PUD*	*M*	*PUD*	*M*	*PUD*	*M*	*PUD*
$A(G)$-based	> 0.99	78.12 −99.6	> 0.999	85.96 −99.78	1	100	1	100
$R(G)$-based	> 0.99	79.2 −99.6	> 0.999	86.76 −99.71	1	100	1	100
DC	0.8486 −0.9328	0.4 −4.77	0.8486 −0.9328	0.4 −4.77	0.8544	0.1	0.4774	3.7

is very low (cf. Tables 11 and 12 in [52]). The above regularities are noticed in all network datasets.

In order to try to find some correlations between the structure of a complex network and the resolution of the communicability-based centrality indices defined on that network, we propose some general *heuristic* guidelines. Namely, we hypothesize that the granularity of a centrality measure defined on the network G may depend on the partition of the node set $V(G)$ of G. The most fundamental method of inducing the partition of the vertex set $V(G)$ is based on automorphic mappings of the vertices of the network G. Recall that an automorphism of a complex network G is an isomorphism of G onto itself (for the definition of an isomorphism cf. [50, p. 26]). The set of all automorphisms of a network constitutes an automorphism group, denoted by $Aut(G)$, of this network. The occurrence of an automorphism depends on the presence of equivalent nodes which can be mapped automorphically onto each other, that is, they can interchange preserving the adjacency relations within the network. Accordingly, automorphically equivalent nodes cannot be distinguished only in terms of the network structure. Hence, it can be claimed that automorphically equivalent nodes should be regarded to fulfill the same role within the network. A subset of $V(G)$ formed by all mutually equivalent nodes is called an *orbit* of the network nodes. The so-called *total topological information content,* denoted by $I_{TOP}(G)$, of a complex network $G = (V(G), E(G))$ with the vertex set $V(G)$ where $|V(G)| = n$ is defined by the following condition [48]:

$$I_{TOP}(G) := n \log_2 n - \sum_{i=1}^{g} |orb_i| \log_2 |orb_i|, \tag{23}$$

where $|orb_i|$ is the number of automorphically equivalent vertices of i-th type (i.e., the cardinality of the i-th orbit) and $\{orb_i \mid 1 \leq i \leq g\}$ is the collection of all orbits of $Aut\,(G)$. The higher the value of $I_{TOP}\,(G)$ is, the less nodes within the network G can be regarded as topologically equivalent. Moreover, we propose to consider the average number of nodes *per* one orbit of $Aut\,(G)$, denoted by $\overline{orb}\,(G)$, as a new graph-theoretical invariant. Formally, for a complex network $G = (V\,(G)\,, E\,(G))$ whose vertex set $V\,(G)$ is partitioned into the collection of orbits $\{orb_i \mid 1 \leq i \leq g\}$, this descriptor is defined by the formula

$$\overline{orb}\,(G) := \frac{1}{g}\sum_{i=1}^{g} |orb_i|\,. \tag{24}$$

The values of $\overline{orb}\,(G)$ range from 1 to n. If G is an *identity graph*, then $\overline{orb}\,(G) = 1$, i.e., each orbit of $Aut\,(G)$ is a singleton and, consequently, all nodes in G are topologically distinguishable from each other. On the other hand, if G is a *symmetric graph* on n vertices, then $\overline{orb}\,(G) = n$, i.e., all nodes in G belong to the same orbit of $Aut\,(G)$ and, consequently, they are topologically equivalent. Accordingly, the higher the value of $\overline{orb}\,(G)$ is, the more vertices within the network G can be considered as topologically equivalent. The first observation that has to be made is that the values of the newly suggested descriptor $\overline{orb}\,(G)$ in the actual datasets (in which the resolution of the communicability-based measures is lower) are higher than in the rewired datasets (in which the resolution of the communicability-based indices is higher). In turn, the values of $I_{TOP}\,(G)$ are lower in the empirical datasets than in the rewired datasets. Thus, it can be conjectured that, at least for some networks and some centrality measures, the resolution of a ranking algorithm is negatively (positively) correlated with the values of $\overline{orb}\,(G)$ (of $I_{TOP}\,(G)$) estimated on the underlying network. To our best knowledge, in the *network science* literature, there are no attempts to elucidate the differential resolution of vertex ranking algorithms in a *causal* way. Therefore, we first decided to test our hypothesis on simple synthetic networks. Namely, we estimated the granularity of the TC and $KC\;(\alpha_k)$ indices as well as their normalized counterparts on the ensemble of 100 undirected trees generated according to the random citation growing graph model (cf. Section 4). Figure 2 displays the relationship (with the corresponding values of the Pearson correlation coefficients r) between the resolution of the TC and $KC\;(\alpha_k)$ measures quantified on the $RCGG$ model and the values of the $I_{TOP}\,(G)$ and $\overline{orb}\,(G)$ descriptors evaluated on this model.

The monotonicity of the TC and $KC\;(\alpha_k)$ indices in the $RCGG$ model ranges from 0.9847 to 0.9964 and from 0.9803 to 0.9964 (respectively), whereas their

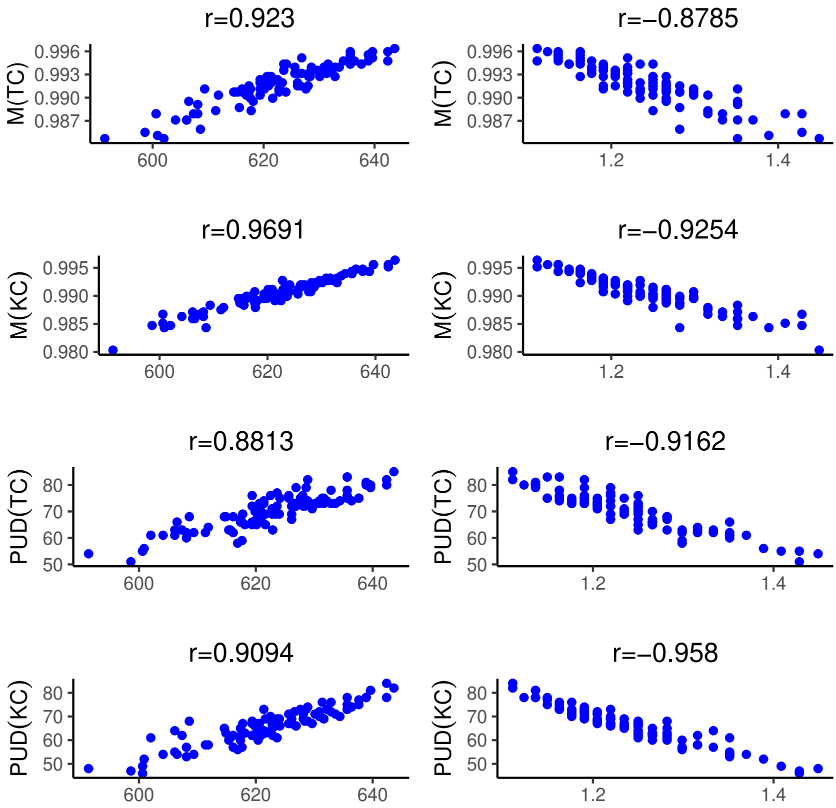

Fig. 2. The relationship between the resolution of the TC and KC centrality indices estimated on the $RCGG$ model and the values of the I_{TOP} and $\overline{orb}$ descriptors of this model. The resolution is quantified by the M and PUD metrics. The KC centrality index is calculated with respect to the attenuation parameter α_k. r is the Pearson correlation coefficient. In all subplots in the left panel, the x axis corresponds to the values of I_{TOP}, whereas in all subplots in the right panel, the x-axis corresponds to the values of $\overline{orb}$.

weighted counterparts from 0.9803 to 0.9964 and from 0.9831 to 0.996, respectively. In turn, the granularity quantified by the PUD metric of the TC and KC (α_k) algorithms ranges from 51 % to 85 % and from 46 % to 84 % (respectively), whereas their normalized variants from 47 % to 85 % and from 47 % to 83 %, respectively. From Figure 2, it can be inferred that the resolution of both ranking methods estimated on the $RCGG$ model is positively highly or very highly correlated with the $I_{TOP}\,(G)$ invariant. In turn, the newly proposed descriptor $\overline{orb}\,(G)$ is negatively highly or very highly correlated with the granularity of the

TC and $KC\ (\alpha_k)$ measures. The linear correlations between the graph-theoretical descriptors given by the equations (23) and (24) and the resolution of the RTC and $RKC\ (\alpha_{rk})$ algorithms are very similar, i.e., the values of r for the linear correlations between $I_{TOP}\,(G)$ and the $M\,(RTC)$), $PUD\,(RTC)$, $M\,(RKC\,(\alpha_{rk}))$ and $PUD\,(RKC\,(\alpha_{rk}))$ values are equal to 0.9529, 0.9243, 0.9487 and 0.9185, respectively. In turn, the values of r for the linear correlations between $\overline{orb}\,(G)$ and the $M\,(RTC)$ and $PUD\,(RTC)$, $M\,(RKC\,(\alpha_{rk}))$ and $PUD\,(RKC\,(\alpha_{rk}))$ values are equal to -0.897, -0.9598, -0.8936 and -0.9583, respectively. Accordingly, it can be uttered that, at least for some complex networks and some ranking algorithms, the more topologically equivalent nodes a network contains, the lower the granularity of the centrality measure estimated on that network is. Moreover, in both cases, it can be observed that $I_{TOP}\,(G)$ is a better predictor of the granularity quantified by the M metric, while $\overline{orb}\,(G)$ is a better predictor of the ganularity quantified by the PUD metric. Therefore, the introduction of the $\overline{orb}\,(G)$ descriptor seems to be justified. Note that the above *heuristic* rules relating the number of topologically equivalent nodes within the network with the resolution of the node significance ranking algorithms estimated on that network does not refer to any other graph-theoretical invariants (like, for instance, the connectivity) and, consequently, these *heuristic* guidelines attempting to elucidate the granularity of the communicability-based centrality measures are, in substance, *purely algebraic.* The above ideas establishing a bridge between *algebraic* and *quantitative graph theory* will be further explored in a separate work.

5.2. The intensional attack simulation experiments

In the deliberate attack simulation experiments carried out in the current subsection, we will juxtapose the communicability-based centrality indices derived from the adjacency matrices with their weighted counterparts. All simulations are carried out on four empirical networks and on two generative model networks presented in Table 2.

In the first series of our simulation experiments, we will analyse the performance of the $A\,(G)$-based centrality measures. Table 5 lists the cumulative values of the $Iso_k(G)$ parameter for the attack scenarios guided by the tested ranking algorithms.

From this table, it follows that, in all network datasets, all numerical instances of the KC index give (approximately) the same results. Also, it is straightforward to observe that, in all cases with the exception of the ER model, the KC measure considerably outperforms the TC measure. In the empirical datasets, with respect

Table 5
The accumulation of the damages in six exemplary networks (quantified by the number of isolated nodes in the post-attack graphs) triggered by the attacks based on the original communicability-based centrality indices. The most aggressive scenarios are marked in bold

Index	*Roget*	*E-mail*	*US political blogs*	*US airport 2010*	*ER*	*BA*
TC	887	11774	19618	103569	0	15388
KC ($\alpha_{\min}$)	915	15241	**30525**	**130598**	0	**28217**
KC (α_k)	**917**	**15245**	**30525**	**130598**	0	**28217**
KC ($\alpha_{0.5}$)	**917**	**15245**	**30525**	**130598**	0	**28217**
KC ($\alpha_{0.85}$)	915	15244	**30525**	**130598**	0	**28217**

to the cumulative values of the $Iso_k(G)$ parameter, the improvement attained by the KC index over the TC index is in the range from 3.27 % (the *Roget* network) to 35.73 % (the *US political blogs* network). In the *ER* and *BA* models, this improvement is equal to 0 % and 45.47 %, respectively.

Table 6 includes the cumulative values of the $C_k(G)$ parameter for the attack schemes driven by the tested ranking methods.

Table 6
The accumulation of the damages in six exemplary networks (quantified by the number of connected components in the post-attack graphs) triggered by the attacks based on the original communicability-based centrality indices. The most aggressive scenarios are marked in bold

Index	*Roget*	*E-mail*	*US political blogs*	*US airport 2010*	*ER*	*BA*
TC	1485	12585	21205	109673	200	17103
KC ($\alpha_{\min}$)	1542	16158	**32300**	**137909**	200	32818
KC (α_k)	**1546**	**16162**	**32300**	**137909**	200	**32819**
KC ($\alpha_{0.5}$)	**1546**	**16162**	**32300**	**137909**	200	**32819**
KC ($\alpha_{0.85}$)	1542	16161	**32300**	**137909**	200	**32819**

Also in this case, all numerical variants of the KC metric give (approximately) the same results. Furthermore, it is evident that, in all cases with the exception of the *ER* model, the KC metric yields the better results than the TC metric. In the actual datasets, with respect to the cumulative values of the $C_k(G)$ parameter, the improvement achieved by the KC index over the TC index ranges from 3.95 %

(the *Roget* network) to 34.35 % (the *US political blogs* network). In the *ER* and *BA* models, this betterment is equal to 0 % and 47.89 %, respectively.

Now, we will study the impact of removing the most vital nodes on the relative size of the giant component. Figure 3 presents the relationship between the number of nodes removed from the network and the relative size of the giant component.

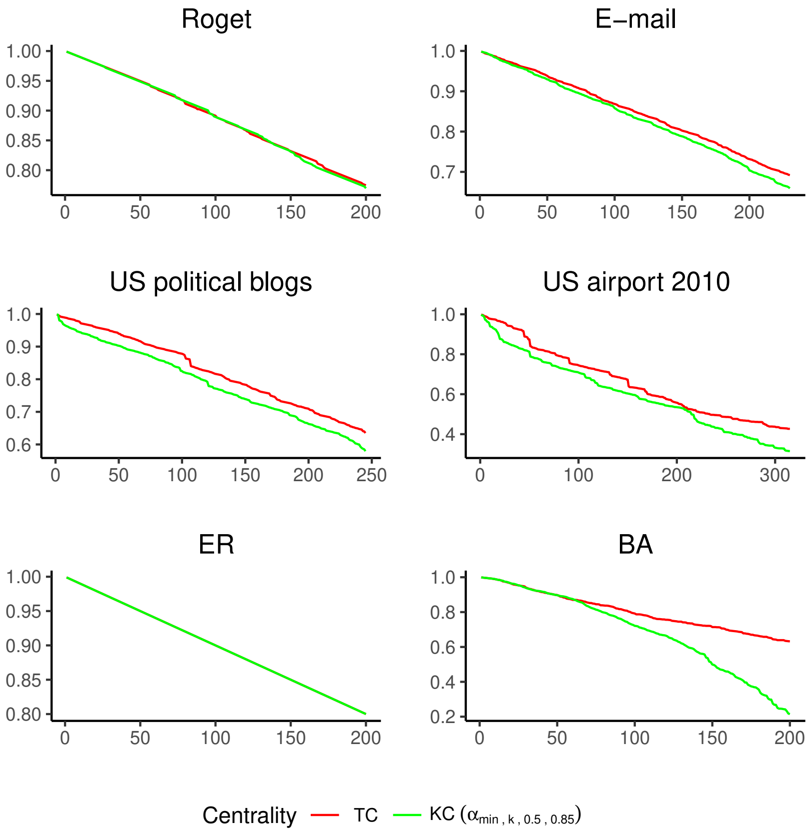

Fig. 3. The relative size of the giant component as a function of the number of nodes removed from six exemplary networks subjected to the attacks guided by the TC index as well as by four numerical instances of the KC index. In all subplots, the x axis corresponds to the number of nodes removed from the network and the y axis corresponds to the reletavie size of the giant component

In turn, Table 7 contains the areas under the robustness curves corresponding to each attack protocol from Figure 3.

These data indicate that, in all network datasets, the attack scenarios based on four numerical versions of the KC index are (approximately) equally destructive

Table 7
The accumulation of the damages in six exemplary networks (quantified by the area under the robustness curve) triggered by the attacks based on the original communicability-based centrality indices. The most aggressive scenarios are marked in bold

Index	*Roget*	*E-mail*	*US political blogs*	*US airport 2010*	*ER*	*BA*
TC	176.922	194.2908	201.1227	207.2452	**179.0005**	160.0215
KC ($\alpha_{\min}$)	176.7757	191.0327	**191.91**	**188.2634**	**179.0005**	137.509
KC (α_k)	**176.7696**	**191.0291**	**191.91**	**188.2634**	**179.0005**	**137.507**
KC ($\alpha_{0.5}$)	**176.7696**	**191.0291**	**191.91**	**188.2634**	**179.0005**	**137.507**
KC ($\alpha_{0.85}$)	176.7757	191.03	**191.91**	**188.2634**	**179.0005**	**137.507**

to the network connectivity. In Figure 3, the robustness curves corresponding to these attack scenarios are visually indistinguishable from each other. Moreover, it can be seen that, in all cases with the exception of the *ER* model, the attacks guided by the KC index are more effective in dismantling the networks than the attacks driven by the TC index. In the *ER* model, all attacks protocols are equally deleterious to the network integrity and, consequently, all robustness curves are visually indistinguishable from each other. In the real-world datasets, with respect to the area under the robustness curve, the betterment attained by the KC measure over the TC measure ranges from 0.09 % (the *Roget* network) to 4.58 % (the *US political blogs* network). In the *ER* and *BA* models, this improvement is equal to 0 % and 14.07 %, respectively.

Figure 4 shows the relationship between the number of vertices deleted from the network and the decline rate of the network efficiency. In turn, Table 8 includes the areas under the efficiency decline curves corresponding to each attack schemes from Figure 4.

These data indicate that, in all datasets, the attack scenarios based on four numerical instances of the KC metric are (approximately) equally aggressive to the network connectivity and, consequently, it is impossible to visually distinguish between the corresponding efficiency decline curves. Furthermore, these data demonstrate that the attack schemes driven by the KC metric produce the better results than the attacks driven by the TC metric. The above regularity is noticed in all six network datasets. In the real-world datasets, with respect to the area under the efficiency decline curve, the refinement attained by the KC index over the TC index is in the range from 14.79 % (the *US airport* network)

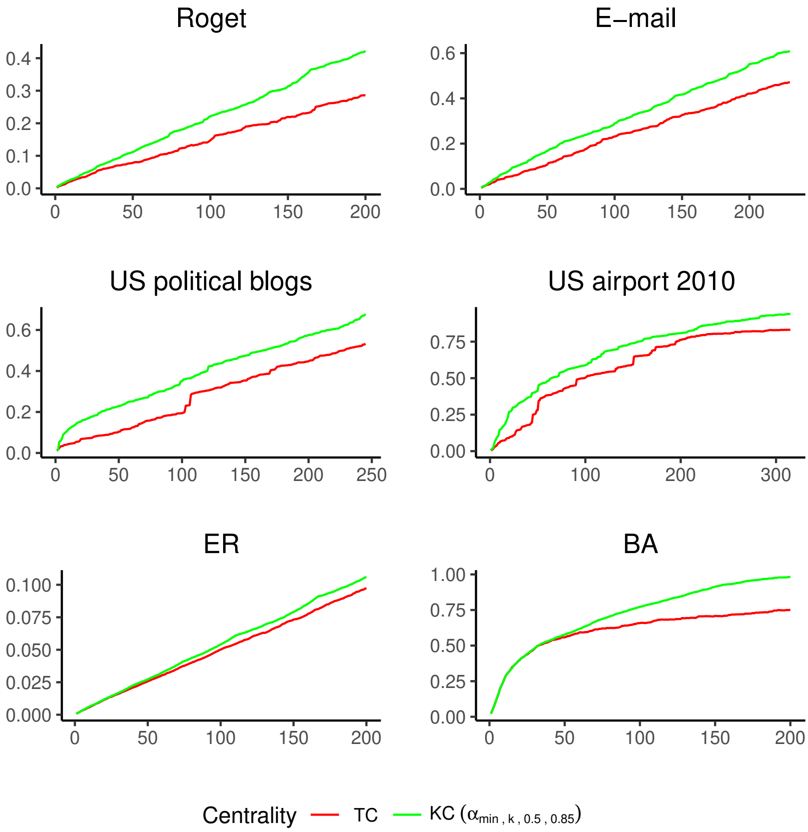

Fig. 4. The decline rate of the network efficiency as a function of the number of nodes removed from six exemplary networks subjected to the attacks guided by the TC index as well as by four numerical instances of the KC index. In all subplots, the x axis corresponds to the number of nodes removed from the network and the y axis corresponds to the decline rate of the network efficiency

to 30.95 % (the *Roget* network). In the ER and BA models, this improvement is equal to 8.06 % and 15.69 %, respectively.

All in all, it can be claimed that the intensional attack simulation experiments presented in the current subsection unambiguously demonstrated that, in all network datasets, the choices of the Katz parameter given by the equations (5)–(8) do not affect the effectiveness of the KC index in dismantling the networks. These experiments also showed that, in the actual network datasets and in the BA model, the KC vertex significance ranking algorithm is more effective in attacking the networks than the TC algorithm. In turn, in the ER model, with respect to

Table 8
The accumulation of the damages in six exemplary networks (quantified by the area under the efficiency decline curve) triggered by the attacks based on the original communicability-based centrality indices. The most aggressive scenarios are marked in bold

Index	*Roget*	*E-mail*	*US political blogs*	*US airport 2010*	*ER*	*BA*
TC	29.8577	57.0609	68.9312	182.5766	9.8824	121.1165
KC ($\alpha_{\min}$)	43.2322	74.9148	**96.6421**	**214.2714**	10.7488	143.6549
KC (α_k)	43.2417	**74.9212**	**96.6421**	**214.2714**	10.7487	**143.6558**
KC ($\alpha_{0.5}$)	**43.2422**	**74.9212**	**96.6421**	**214.2714**	10.7487	**143.6558**
KC ($\alpha_{0.85}$)	43.2323	74.9203	**96.6421**	**214.2714**	**10.7489**	**143.6558**

three evaluation criteria, i.e., the $Iso_k(G)$, $C_k(G)$ parameters and the area under the robustness curve, all attack strategies guided by the communicability-based centrality indices are equally hazardous to the network coherence, while with respect to the area under the efficiency decline curve, the KC centrality measure outperforms the TC measure.

In the second series of our simulation experiments, we will analyse the performance of the $R(G)$-based centrality measures. As a reference centrality measure, we will use the DC index.

Table 9 contains the cumulative values of the $Iso_k(G)$ parameter for the attack protocols guided by the tested weighted ranking methods as well as by the DC measure.

Table 9
The accumulation of the damages in six exemplary networks (quantified by the number of isolated nodes in the post-attack graphs) triggered by the attacks based on the weighted communicability-based centrality indices and on the DC index. The most aggressive scenarios are marked in bold

Index	*Roget*	*E-mail*	*US political blogs*	*US airport 2010*	*ER*	*BA*
RTC	2979	21401	41815	162064	0	28314
RKC ($\alpha_{\min}$)	2365	19626	38580	157437	0	**28392**
RKC (α_{rk})	**3141**	**22208**	**44151**	**165180**	0	28321
RKC ($\alpha_{0.5}$)	2743	20762	40607	160239	0	28379
RKC ($\alpha_{0.85}$)	1617	17385	34334	148893	0	28324
DC	946	15432	30613	130930	0	28174

From this table, it follows that, in all empirical datasets, the RKC (α_{rk}) index outperforms all other centrality indices. In addition, it can be observed that, in the actual network datasets, the attack scenarios based on the RTC index have the second best attack effect. Moreover, it is evident that, in all datasets with the exception of the *ER* model, the weighted communicability-based ranking algorithms give the better or significantly better results than the DC index. On the other hand, the scores from Table 9 indicate that, among the communicability-based centrality measures, when the nodes are deleted from the network according to the importance ranking lists produced by the RKC $(\alpha_{0.85})$ index, then the network has the worst fragmentation effect. This regularity was observed in all real-world network datasets. In these networks, with respect to the cumulative values of the $Iso_k(G)$ parameter, the improvement achieved by the RKC (α_{rk}) index over the RTC, KC and DC indices is in the range from 1.89 % (the *US airport 2010* network) to 5.29 % (the *US political blogs* network) and from 20.94 % (the *US airport 2010* network) to 70.81 % (the *Roget* network) as well as from 20.73 % (the *US airport 2010* network) to 69.88 % (the *Roget* network), respectively. In turn, the improvement attained by the RTC index over its unweighted version is in the range from 36.09 % (the *US airport 2010* network) to 70.22 % (the *Roget* network). In the *ER* model, all attack schemes are equally aggressive to the network unity and the performance of the weighted communicability-based centrality indices is the same as the performance of their original forms. In the *BA* model, the improvement achieved by the best normalized resolvent-based measure over the best original resolvent-based measure and over the DC index is equal to 0.62 % and 0.77 %, respectively. In turn, the refinement attained by the RTC measure over its unweighted counterpart is equal to 45.65 %.

Table 10 presents the cumulative values of the $C_k(G)$ parameter for the attack schemes guided by the considered weighted centrality metrics as well as by the DC metric.

The scores from this table indicate that, in all empirical networks, the attack strategies guided by the RKC (α_{rk}) measure are the most effective in degrading the networks. On the other hand, the attack strategies driven by the RTC measure possess the second best attack effect. Moreover, in all cases with the exception of the *ER* model, the normalized communicability-based centrality measures outperform or remarkably outperform the DC measure. In the empirical datasets, among the indices defined by the matrix functions, the RKC $(\alpha_{0.85})$ measure performs the worst. In these datasets, with respect to the cumulative values of the $C_k(G)$ parameter, the betterment attained by the RKC (α_{rk}) measure over the RTC, KC and DC measures ranges from 1.72 % (the *US airport 2010* network)

Table 10
The accumulation of the damages in six exemplary networks (quantified by the number of connected components in the post-attack graphs) triggered by the attacks based on the weighted communicability-based centrality indices and on the *DC* index. The most aggressive scenarios are marked in bold

Index	*Roget*	*E-mail*	*US political blogs*	*US airport 2010*	*ER*	*BA*
RTC	4558	22832	44207	171435	**200**	32990
RKC ($\alpha_{\min}$)	3763	21012	40809	166520	**200**	32977
RKC (α_{rk})	**4731**	**23683**	**46688**	**174439**	**200**	32987
RKC ($\alpha_{0.5}$)	4246	22170	42955	169563	**200**	**33028**
RKC ($\alpha_{0.85}$)	2684	18620	36343	156566	**200**	32946
DC	1588	16367	32408	138279	**200**	32716

to 5.31 % (the *US political blogs* network) and from 20.94 % (the *US airport 2010* network) to 67.32 % (the *Roget* network) as well as from 20.73 % (the *US airport 2010* network) to 66.43 % (the *Roget* network), respectively. In turn, the betterment attained by the *RTC* measure over its unweighted instance ranges from 36.03 % (the *US airport 2010* network) to 67.42 % (the *Roget* network). In the *ER* model, all attack protocols are equally destructive to the network connectivity and the efficiency of the weighted communicability-based indices is the same as the efficiency of their original variants. In the *BA* model, the betterment attained by the best normalized resolvent-based index over the best original resolvent-based index and over the *DC* index is equal to 0.63 % and 0.94 %, respectively. In turn, the betterment attained by the *RTC* index over the *TC* index is equal to 48.16 %.

Next, we will analyse the influence of deleting the most crucial nodes on the relative size of the giant component. Figure 5 presents the relationship between the number of nodes removed from the network and the relative size of the giant component. In turn, Table 11 lists the areas under the robustness curves corresponding to each attack scenario from Figure 5.

These data demonstrate that, in all actual networks, the attack strategies based on the *RKC* (α_{rk}) index are the most harmful to the network connectivity. Consequently, in all empirical datasets, the *RKC* (α_{rk}) index outperforms all other node significance ranking algorithms. Moreover, in the real-world networks, the attack protocols based on the *RTC* index have the second best attack effect. It can be observed that, in all actual datasets, the attack schemes driven by the centrality indices defined *via* the matrix functions are decisively more deleterious

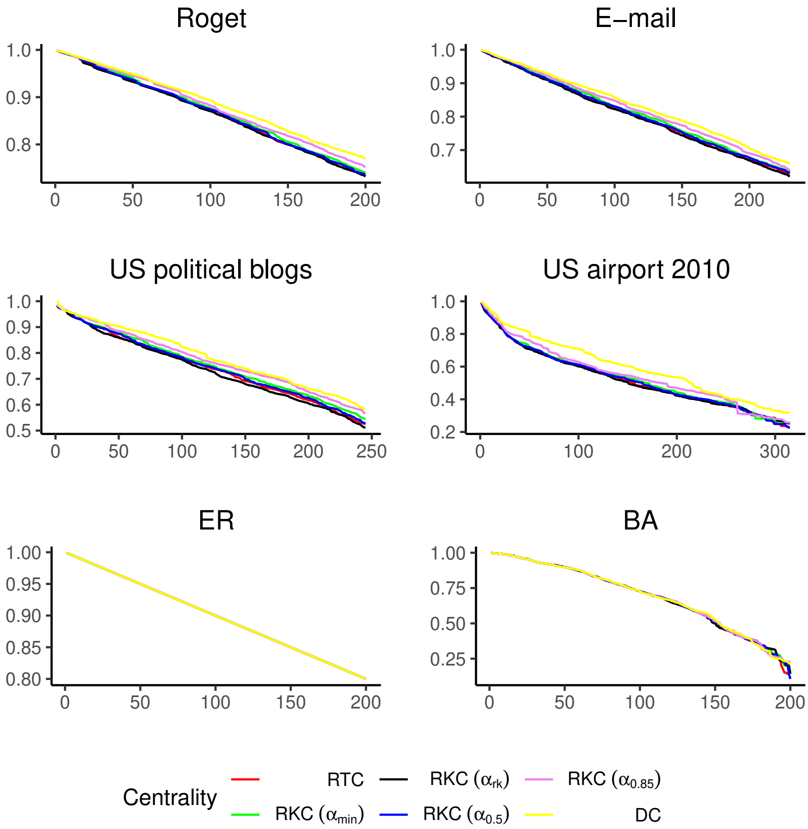

Fig. 5. The relative size of the giant component as a function of the number of nodes removed from six exemplary networks subjected to the attacks guided by the weighted communicability-based centrality indices as well as by the *DC* index. In all subplots, the x axis corresponds to the number of nodes removed from the network and the y axis corresponds to the relative size of the giant component

to the network connectivity than the attacks driven by the *DC* index. Among the newly proposed normalized communicability-based centrality metrics, the RKC ($\alpha_{0.85}$) index produces the poorest results. The above regularity was recorded in all real-world network datasets. In these datasets, with respect to the area under the robustness curve, the refinement achieved by the RKC (α_{rk}) metric over the RTC, KC and DC metrics ranges from 0.11 % (the *Roget* network) to 1.23 % (the *US political blogs* network) and from 2.44 % (the *Roget* network) to 13.38 % (the *US airport 2010* network) as well as from 2.31 % (the *Roget* network) to 13.24 % (the *US airport 2010* network), respectively. In turn, the refinement achieved by

Table 11
The accumulation of the damages in six exemplary networks (quantified by the area under the robustness curve) triggered by the attacks based on the weighted communicability-based centrality indices and on the *DC* index. The most aggressive scenarios are marked in bold

Index	*Roget*	*E-mail*	*US political blogs*	*US airport 2010*	*ER*	*BA*
RTC	172.6413	184.6161	181.4992	164.0795	**179.0005**	**136.921**
RKC ($\alpha_{\min}$)	173.6434	186.1311	184.5098	166.7856	**179.0005**	137.4195
RKC (α_{rk})	**172.4552**	**183.8027**	**179.2709**	**163.0808**	**179.0005**	137.169
RKC ($\alpha_{0.5}$)	173.0468	185.2048	182.5912	164.9545	**179.0005**	137.232
RKC ($\alpha_{0.85}$)	175.082	188.4131	188.3822	172.4517	**179.0005**	137.3675
DC	176.7153	190.8314	191.7971	187.965	**179.0005**	137.5485

the *RTC* metric over its unweighted form ranges from 2.42 % (the *Roget* network) to 20.83 % (the *US airport 2010* network). In the *ER* model, all attack scenarios are equally detrimental to the network coherence and, consequently, the robustness curves corresponding to these attack schemes are visually indistinguishable from each other. Moreover, the $R(G)$-based centrality measures are equally effective in degrading the networks as the $A(G)$-based centrality measures. In the *BA* model, all numerical variants of the *RKC* index are (approximately) equally effective in attacking the networks. Moreover, the *RKC* measure give (approximately) the same results as the *DC* measure and its unweighted variant. In turn, the refinement attained by the *RTC* ranking algorithm over its unweighted version and over the *DC* algorithm is equal to 14.44 % and 0.46 %, respectively.

Figure 6 presents the relationship between the number of vertices deleted from the network and the decline rate of the network efficiency. In turn, Table 12 includes the areas under the efficiency decline curves corresponding to each attack scheme from Figure 6.

These data show that, in the empirical datasets, all weighted communicability-based ranking methods considerably outperform their original variants as well as the *DC* algorithm. Namely, in these datasets, with respect to the area under the efficiency decline curve, the improvement attained by the normalized resolvent-based centrality measures over their unweighted forms and over the *DC* measure is in the range from 8.51 % (the *US airport 2010* network) to 15 % (the *US political blogs* network) and from 8.41 % (the *US airport 2010* network) to 14.82 % (the *US political blogs* network), respectively. In turn, the improvement achieved by the *RTC* measure over its unweighted variant is in the range from 22.05 % (the *US*

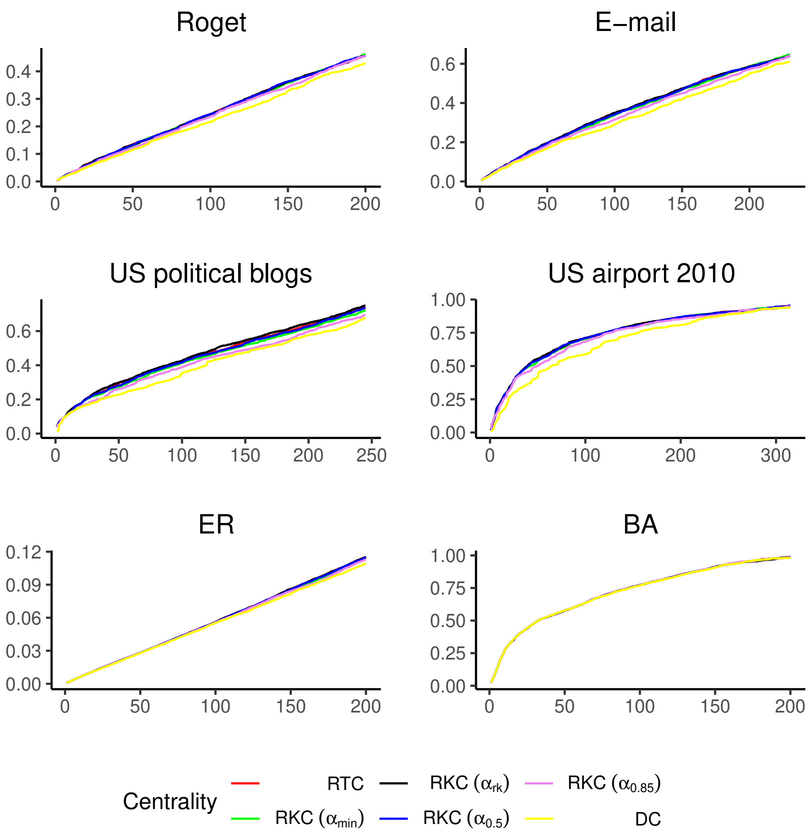

Fig. 6. The decline rate of the network efficiency as a function of the number of nodes removed from six exemplary networks subjected to the attacks guided by the weighted communicability-based centrality indices as well as by the *DC* index. In all subplots, the x axis corresponds to the number of nodes removed from the network and the y axis corresponds to the decline rate of the network efficiency

airport 2010 network) to 38.55 % (the *Roget* network). Furthermore, in the *ER* model, the refinement attained by the best normalized resolvent-based centrality measure over the best original resolvent-based measure and over the *DC* index is equal to 5.69 % and 3.66 %, respectively. In turn, the improvement achieved by the *RTC* index over its unweighted form and over the *DC* algorithm is equal to 13.29 % and 3.66 %, respectively. In the *BA* model, all numerical instances of the *RKC* index produce (approximately) the same results as the *DC* index and their original versions. Moreover, the betterment achieved by the *RTC* algorithm

Table 12

The accumulation of the damages in six exemplary networks (quantified by the area under the eficiency decline curve) triggered by the attacks based on the weighted communicability-based centrality indices and on the *DC* index. The most aggressive scenarios are marked in bold

Index	*Roget*	*E-mail*	*US political blogs*	*US airport 2010*	*ER*	*BA*
RTC	48.5872	84.1488	111.4528	**234.2109**	11.3966	**143.8439**
RKC ($\alpha_{\min}$)	48.3705	83.2272	108.175	233.1285	11.3135	143.8333
RKC (α_{rk})	48.5667	**84.4713**	**113.6907**	234.056	**11.3973**	143.7122
RKC ($\alpha_{0.5}$)	**48.6052**	83.8585	110.3517	233.9009	11.3679	143.8386
RKC ($\alpha_{0.85}$)	47.065	80.3485	102.9144	229.0105	11.1934	143.8198
DC	44.0658	75.3853	96.84	214.5179	10.9796	143.5713

over its unweighted form is equal to 15.8 %, whereas the betterment over the *DC* ranking method is negligible.

All in all, from the above results, it can be inferred that the values of the weighted Katz parameter given by the equations (18)–(21) exert the significant influence on the performance of the newly conceptualized ranking algorithms. Following the references [26, 38], we use the term *percolation* to cover any specific procedure for the node removal from the network. As is well known, for the random uniform deletion of vertices from the network or for the targeted removal of vertices in decreasing order of their degrees, the percolation processes on complex networks were carefully studied and many elegant analytical results were obtained (cf. [23,26,38] and the references cited therein). In our contribution, we study the effect on the network structure of the intensional removal of nodes driven by several more composite node ranking algorithms than simply the *DC* index. Similarly as in [26], where the deliberate attacks guided by the degree, *betweenness*, *closeness* and *eigenvector* centrality measures were analysed, it can be claimed that "[...]*since these centrality measures are subtle non-local measures of a vertex's significance it seems unrealistic to anticipate any all-embracing analytical theory of the corresponding percolation process, and hence our present work is computational in nature*". Consequently, to attempt to elucidate why, in most cases, the *KC* index is more effective in dismantling the network than the *TC* index or why, in most cases, the normalized communicability-based measures considerably outperform their original variants we propose some *heuristic* guidelines. Namely, we conjecture that the superiority of the *KC* measure over the *TC* measure recorded in the actual datasets is mainly dictated by the heterogeneous degree distribution of these network datasets. As previously stated, four real-world graphs used in the present study can be considered as scale-free. Such networks possess few nodes with many edges and many nodes with few edges. In the *network science* literature, to quantify the *heterogeneity* of the degree distribution (shortly, the *degree heterogeneity*), many different invariants were proposed [8]. Nevertheless, as pointed out in [8], only two descriptors, i.e., the coefficient of variation and the Gini coefficient are applicable to all distributions and have the desired properties of consistency over alternations in a distribution arising from the transfer, addition and replication principles. From the comparison of the values of the V_k and G_k indices corresponding to the degree sequences of the exemplary networks (cf. Table 2), it follows that, among the actual networks, the degree heterogeneity of the *Roget* network is the lowest. This fact juxtaposed with the results of the deliberate attack simulation experiments indicating that, with respect to three evaluation parameters (i.e., the $Iso_k(G)$ and $C_k(G)$ indices and the area under the robustness curve), the improvement attained by the KC algorithm over the TC algorithm is also the lowest, can be considered as a premise that this improvement depends on the degree heterogeneity of the underlying graphs. To substantiate this hypothesis, we carried out the intensional attacks on two generative model networks of the same size and density. These models are generated according to the Erdős-Rényi $G(n, m)$ procedure (the *ER* model) and according to the preferential attachment procedure of Barabási-Albert (the *BA* model). As is widely known, the *ER* model is characterized by the *Poisson* degree distribution. Intuitively, such a distribution means that most vertices possess a degree close to the average and, consequently, can be viewed as homogeneous. From Table 2, it follows that, among all datasets, the degree heterogeneity of the *ER* model quantified by the V_k and G_k invariants is the lowest. In turn, the topology of the *BA* model is scale-free and, consequently, the degree heterogeneity of this model is high. From the targeted attack experiments, it can be observed that, in the *ER* model, the performance of all tested centrality metrics (i.e. the TC, KC, RTC, RKC and DC indices) quantitatively assessed by the three evaluation parameters (i.e., the $Iso_k(G)$ and $C_k(G)$ indices and the area under the robustness curve) is the same which implies that the percolation of vital nodes identified by each ranking method exert no significant impact on degrading the network structure. The similar effect was noted in [24] where the authors juxtaposed the *density* centrality measure with the *DIL*, *gravity* and closeness centrality indices. In the *ER* models, the results of the deliberate attacks guided by these ranking algorithms were summarized by the authors of [24] by the following assertion: "[...] [t]*he reason* [for this effect] *is mainly due to the homogeneous degree distribution (the degrees are not so different) of such* *networks*". Thus, it can be postulated that, owning to the homogeneity of the *ER* network, there is no substantial difference whether vertices are removed from this graph according to the *TC*, *KC*, *RTC*, *RKC* or *DC* measures. Thus, it can be inferred that, in the *ER* model, all tested ranking methods identify nodes which are equally significant to the network connectivity. On the other hand, in the *ER* model, when the performance of the considered ranking methods is quantified by the area under the efficiency decline curve, the *KC* index is superior to the *TC* index as well as the weighted communicability-based measures are superior to their original counterparts and to the *DC* measure. Hence, in the *ER* model, the post-attack graphs in the attack scenarios guided by all studied centrality metrics have the same size (i.e., the same number of nodes) but are not isomorphic. This feature cannot be observed from the robustness curve but is visible from the efficiency decline curve (cf. Figures 5 and 6). The similar phenomenon was recorded in [24] where, with respect to the decline rate of the network efficiency, the density centrality index was superior to its competitors. Therefore, it can be concluded that, with respect to the network overall communicability, the *KC* index outperforms the *TC* index as well as the normalized exponential-based and resolvent-based ranking algorithms outperform their original counterparts. In turn, in the *BA* model with the heterogeneous degree distribution, the effectiveness of the attack protocols based on the *KC* and *TC* indices vary considerably. Since the *ER* and *BA* models used in the current work possess the same size and density, it can be claimed that the *heuristic* assumption attempting to explain the superiority of the *KC* index over the *TC* index by the heterogeneity of the degree sequences of the underlying graphs is substantiated. It should be emphasized that our results unambiguously showcased that, in the actual networks, the newly proposed weighted communicability-based centrality indices are decisively more effective in attacking the networks than the widely used classical *DC* algorithm. In turn, in the *BA* model, the novel ranking algorithms are only slightly better than the *DC* measure. To shed some light on these results, let us recall the excerpt from [26] where the authors uttered that "[...] [i]*t is rather striking that degree which is a purely local centrality measure provides a more effective means of targeting vertices than any of the other centrality measures, which are non-local in nature and can account for the global structure of the network. We believe that degree centrality will prove in general to be superior to other centrality measures at exposing the vulnerability under simultaneous targeted attack of any network which lacks certain specific structural properties that would favor the efficacy of other centrality measures.* [...] *In the absence of any particular structural properties the best estimator of the vulnerability of a vertex under simultaneous targeted* *attack appears to be simply the number of neighbors that the vertex has*". Thus, it can be concluded that the *BA* model used in our study, although it was generated according to the preferential attachment mechanism, is essentially random in nature and, consequently, lack any specific structural properties (for instance, the modularity or the global clustering, cf. Table 2) that would allow other vertex significance ranking algorithms to be superior to the *DC* algorithm. Accordingly, in the random scale-free networks, the newly proposed weighted communicability-based centrality measures are only slightly more successful in recognizing crucial nodes than their original counterparts and the classical *DC* measure. In turn, in the real-world networks, all normalized communicability-based ranking methods remarkably outperform their $A(G)$-based variants as well as the *DC* index.

In summary, it should be emphasized that our results unquestionably showcased that, by substituting the dichotomous adjacency matrix by its weighted version in the definitions of the centrality indices developed by L. Kac [29] as well as by M. Benzi and C. Klymko [10], we arrive at the conceptualization of two novel ranking algorithms which are able to prioritize vertices in networks more correctly. Therefore, it can be concluded that the proposed modification of the original *TC* and *KC* measures seems to be justified. Furthermore, it should be indicated that our *heuristic* guidelines attempting to elucidate the superiority of one communicability-based index over another or over the *DC* measure should be read in the light of the reference [18]. Namely, the authors of [18] defined four new communicability functions (two based on networks of quantum harmonic oscillators and two on networks of classical harmonic oscillators) and compared their efficacy in several different scenarios. They remarked that "[i]*t is important to note that it is not a matter of deciding which index is the 'correct' one to indicate the communication. There is indeed no standard that we can refer to in judging the 'correctness' of an index. It is a matter of which index is more appropriate to a specific problem than others. In judging the appropriateness, we will have to resort to our intuition and experience. Typically, we would make predictions from various indices and compare them with the results of analyzing actual datasets or sometimes even with a plausible guess*". Later on, the authors of [18] claimed that "[...] *there is no systematic way of selecting one communicability function for a particular problem; the use of one or another of these functions relies on the particular problem under study*". Accordingly, the remarks of E. Estrada, N. Hatano and M. Benzi can be also applied to the communicability functions and the centrality indices based on them which are conceptualized and analysed in our contribution.

5.3. The rank correlations between the tested ranking methods

In this subsection, we will scrutinize the rank correlations quantified by the Kendall τ rank correlation coefficient between the centrality measures defined *via* the matrix functions estimated on the real-world networks and on their rewired analogues as well as on two generative models. For a comparative purpose, we will record the rank correlations between the communicability-based centrality indices and the classical DC measure. In our analyses, we will characterize the exemplary empirical networks with respect to their Kendall centrality correlation profiles conditioned by the tested centrality metrics. This profile is identified with a specific pattern of rank correlations between different centrality measures [45,51]. In our studies, we will follow the methodology of null models with the preserved degree sequences. Recall that the degree distribution has become accepted as the most fundamental network characteristic (apart from its size). Therefore, it is a standard to compare network invariants to a null model in which the degrees of the graph are fixed and everything else is random. Accordingly, when a randomly rewired model network with the identical degree sequence as the real-world (original) graph is used as a reference (null) object, then it is possible to obtain information about the structure of the empirical network apart from what comes from its degree distribution. To check if the Kendall centrality correlation profiles of the exemplary complex networks determined by the communicability-based centrality indices depend on their degree distribution, we have generated for each empirical graph the ensemble of 50 randomly rewired model networks which possess the same degree sequences as the original graphs but are less structured, i.e., they have the significantly lower values of C and $Q\,(w)$ (cf. Table 3). In order to quantitatively compare the rank correlation coefficients evaluated on the real-world and rewired networks, we will calculate the corresponding z-scores. The mathematical definition of the z-score statistics in included in [51, pp. 39–40]. Generally, it is assumed that a result is statistically significant if its corresponding z-score is above 2 [51]. Thus, in our context, we presuppose that a given rank correlation between two centrality measures does not depend on the degree distribution of the underlying network if its z-score is below -2 or above 2. The detailed measurements are contained in Tables 13–27 in [52].

Table 13 summarizes the rank correlations between the original communicability-based centrality indices and their weighted counterparts.

These data indicate that, in all network datasets, the rank correlations between the exponential-based measure and its normalized counterpart are lower

Table 13
The numerical ranges of the rank correlations between the original communicability-based centrality indices and their weighted counterparts. The symbol − denotes the dash

Correlation	Empirical networks	Rewired networks	*ER*	*BA*
$\tau\,(TC, RTC)$	0.2289 −0.6652	0.6326 −0.8087	0.7351	0.0318
$\tau\,(KC, RKC)$	0.5751 −0.9127	0.7263 −0.9323	0.824 −0.8645	0.3823 −0.4813

than the correlations between the resolvent-based indices and their weighted variants (cf. Tables 13 and 14 in [52]). Moreover, all rank correlations are higher in the rewired datasets than in the real-world datasets. Thus, all z-scores are negative (cf. Table 15 in [52]). Furthermore, the values of the z-score statistics are always above 2. Consequently, it can be conjectured that the rank correlations between the communicability-based centrality indices derived from the adjacency matrices of the empirical networks and their newly proposed weighted versions do not depend on the degree distribution of the underlying graphs. From the fact that the rank correlations between the original communicability-based centrality indices and their modified counterparts rarely exceed 0.9, it can be uttered that the novel normalized communicability-based ranking algorithms do not duplicate the structural information contained in the ranking methods developed in [10,29].

Table 14 summarizes the rank correlations between the communicability-based measures and the classical DC index.

These data show that, in the actual and rewired datasets as well as in the ER model, all numerical instances of the KC index are very highly correlated with the DC measure, whereas in the BA model the correlations $\tau\,(DC, KC)$ are high (cf. Tables 16 and 17 in [52]). Thus, the KC and DC indices induce the very similar ranking lists. This phenomenon is not surprising, since the Katz algorithm can be perceived as a generalization of the classical degree centrality measure. Furthermore, this similarity can explain the fact that the attack protocols guided by both these ranking methods are (approximately) equally effective in dismantling the networks (cf. Subsection 5.2). On the other hand, in the empirical and rewired datasets as well as in the ER model, all numerical instances of the RKC algorithm are highly or very highly correlated with the DC algorithm, whereas in the BA model this correlation is moderate or high (cf. Tables 19 and 20 in [52]).

Table 14
The numerical ranges of the rank correlations between the communicability-based centrality indices and the *DC* index. The symbol − denotes the dash

Correlation	Empirical networks	Rewired networks	*ER*	*BA*
$\tau\,(DC, TC)$	0.5544 −0.814	0.7989 −0.9138	0.8668	0.3641
$\tau\,(DC, KC)$	0.96 −0.9829	0.9598 −0.9828	0.9614	0.8312
$\tau\,(DC, RTC)$	0.6883 −0.847	0.6606 −0.9411	0.9227	0.7491
$\tau\,(DC, RKC)$	0.8636 −0.9102	0.8219 −0.9708	0.9226 −0.9614	0.6958 −0.8299

Moreover, among the weighted resolvent-based centrality indices estimated on the real-world and synthetic networks, the rank correlations $\tau\,(DC, RKC\,(\alpha_{rk}))$ are always the lowest, whereas the correlations $\tau\,(DC, RKC\,(\alpha_{0.85}))$ are always the highest. Thus, nodes which are crucial with respect the *DC* measure are, in general, also crucial according to the *RKC* $(\alpha_{0.85})$ algorithm. On the other hand, the rankings based on the *DC* and *RKC* (α_{rk}) measures are, to a considerable extent, divergent. These phenomena can elucidate the fact that, among the node ranking methods analysed in Subsection 5.2, the attack schemes driven by the *DC* index and all numerical instances of the *KC* algorithm as well as by the *RKC* $(\alpha_{0.85})$ measure are the least effective in degrading the networks, whereas the attack protocols guided by the *RKC* (α_{rk}) index are the most deleterious to the network connectivity. Therefore, it is apparent that the above rank correlation analyses are in accordance with the results of the deliberate attack simulation experiments from Subsection 5.2. Furthermore, the values of the z-score statistics for the rank correlations between the weighted communicability-based centrality indices and the *DC* measure are always above 2 (cf. Table 21 in [52]). This fact exemplifies that there is no trivial relationship between the distributions of the newly proposed centrality algorithms and the classical degree ranking algorithm.

Finally, Table 15 summarizes the rank correlations between all pairs of communicability-based centrality measures estimated on the empirical and synthetic networks.

Table 15
The numerical ranges of the pairwise rank correlations between the communicability-based centrality indices. The symbol − denotes the dash

Correlation	Empirical networks	Rewired networks	*ER*	*BA*
$A\,(G)$-based	0.5916 − 1	0.8278 − 1	0.8978 − 1	0.6033 − 1
$R\,(G)$-based	0.8172 −0.9858	0.8570 −0.9884	0.9568 −0.9982	0.8627 −0.9888

From these data, it follows that the rank correlations between all pairs of the original communicability-based indices range from moderate to perfect in the actual datasets and in the *ER* model and from high to perfect in the rewired datasets and in the *BA* model (cf. Tables 22 and 23 in [52]). On the other hand, the correlations between all pairs of normalized communicability-based centrality measures range from high to very high in the empirical and rewired datasets as well as in the *BA* model, whereas in the *ER* model these correlations are always very high (cf. Tables 25 and 26 in [52]). Moreover, in the vast majority of cases, the values of the corresponding z-score statistics are above 2 (cf. Tables 24 and 27 in [52]). Therefore, it can be concluded that the rank correlations between the communicability-based ranking algorithms are (generally) independent on the degree distribution of the underlying graphs.

In order to further test if the overall Kendall centrality correlation profiles of the empirical networks induced by the rank correlations between the communicability-based centrality metrics depend on the degree sequences of the underlying graphs, we carry out the Principal Component Analysis (PCA). In the first series of our measurements, each network (empirical or rewired) is identified with the 10-dimensional feature vector comprising all pairwise Kendall τ correlation coefficients between the original communicability-based centrality measures. In turn, in the second series of our measurements, each network (empirical or rewired) is also represented by the 10-dimensional feature vector consisting of all pairwise Kendall τ correlation coefficients between the weighted communicability-based centrality measures. Figures 7 and 8 display four two-dimensional projections of the ensembles composed of 51 graphs (i.e., one real-world network and 50 rewired networks) from the Kendall centrality correlation profile spaces determined by the tested vertex significance ranking algorithms.

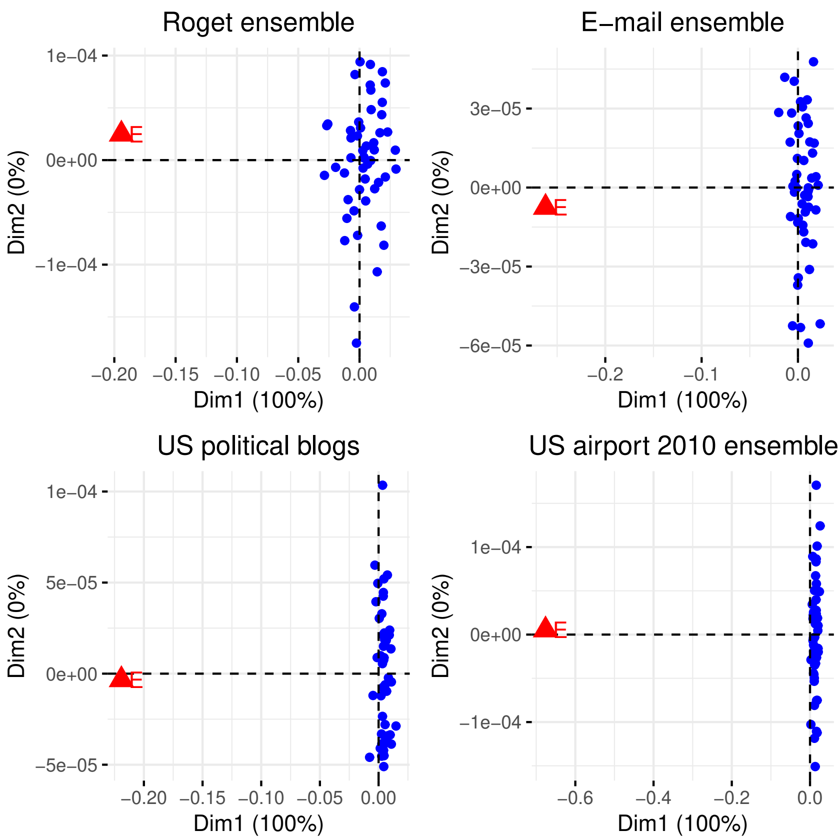

Fig. 7. The PCA projections of the Kendall centrality correlation profile spaces determined by the original communicability-based centrality indices estimated on the empirical and randomly rewired model networks. In all subplots, the real-world network is denoted by the letter E and is marked by the red triangle symbol

From Figures 7 and 8, it can be seen that, in all eight two-dimensional projections, the real-world networks are always located outside the region of the Kendall correlation profile space spanned by the randomly rewired model networks with the conserved degree distribution. Thus, it can be uttered that the PCA outcomes unquestionably revealed that the overall Kendall centrality correlation profiles conditioned by the communicability-based node importance ranking algorithms are able to discriminate between the real-world complex networks and their randomly rewired counterparts with the preserved degree distribution. Therefore, it can be concluded that the Kendall centrality correlation profiles induced by the centrality

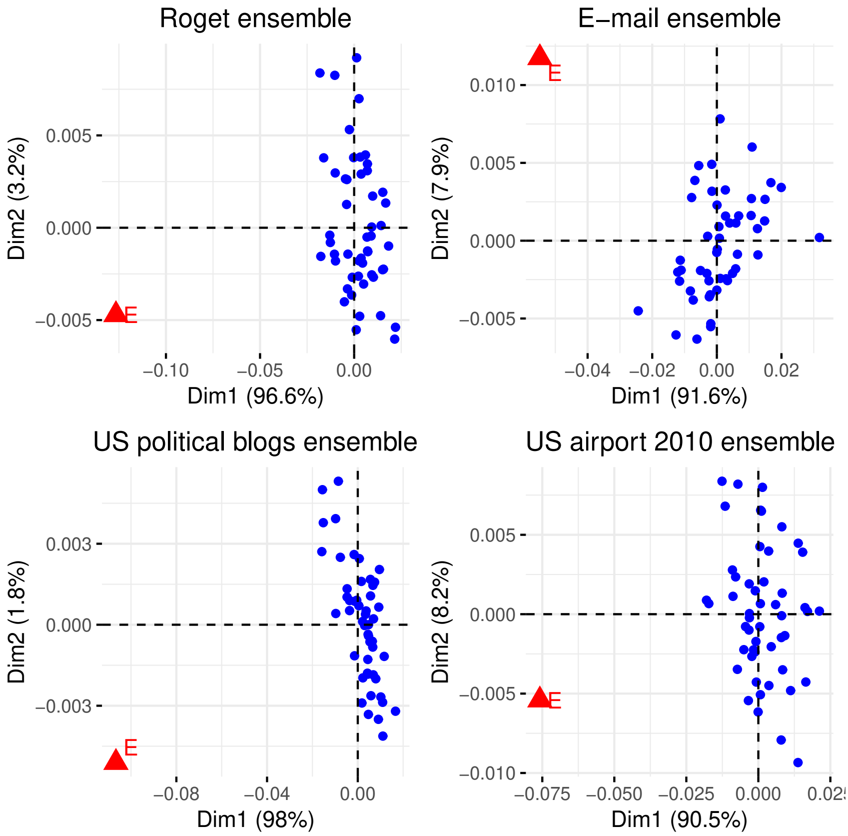

Fig. 8. The PCA projections of the Kendall centrality correlation profile spaces determined by the weighted communicability-based centrality indices estimated on the empirical and randomly rewired model networks. In all subplots, the real-world network is denoted by the letter E and is marked by the red triangle symbol

indices under study can characterize the empirical complex networks and that this characterization can not be inferred from the degree sequences of these networks.

All in all, it can be concluded that the above rank correlation analyses clearly showcased that there is no trivial relationship between the novel centrality measures derived from the product connectivity matrices and their original forms. Thus, it is obvious that the newly suggested weighted communicability-based node significance ranking algorithms do not duplicate the structural information contained in the indices proposed by L. Katz [29] as well as by M. Benzi and C. Klymko [10].

5.4. The robustness of the tested ranking algorithms

Most contemporary network studies rely on the *observed* (i.e., *measured*) networks that may differ from the *underlying* (i.e., *true*) networks which are often obfuscated by different measurement errors. Therefore, the issue of measurement errors in network data is a key problem in the realm of *applied network science*, as virtually all real-world network datasets are plagued by some kind of measurement errors. Several previous works have revealed that observational errors often possess a significant influence on the accuracy of network measures calculated from *erroneous* (i.e., *noisy*) data (cf. [33, 39] and the references cited therein).

In the present subsection, we will study the robustness of the communicability-based centrality indices against three types of random link errors introduced in Section 4. The detailed measurements are included in Tables 28–33 in [52].

Figures 9 and 10 present the relationship between the number of edges deleted from the network or added to the network and the robustness quantified by $\tau\left(C_{true}, C_{noisy}\right)$ of the communicability-based centrality indices estimated on four empirical networks.

From these data, several conclusions can be drawn. Namely, it can be observed that in the random edge removal model, all communicability-based centrality indices are more robust than in the random edge addition model. This means that when links are added to the network, the effect on $\tau\left(C_{true}, C_{noisy}\right)$ is different when the same amount of links are deleted from the network. The similar asymmetrical phenomenon was recorded in [39]. Namely, the authors of the reference [39] observed that, in many cases, the robustness of several common centrality measures was lower in the random edge addition model than in the random edge removal model. Moreover, the intensity of the above asymmetrical phenomenon significantly depends on the network as well as on the node importance ranking algorithm. Furthermore, it can be noticed that when edges are deleted from the network, the robustness of the $A\left(G\right)$-based centrality indices is slightly higher than the robustness of their normalized counterparts. On the other hand, when links are added to the network, the robustness of the $A\left(G\right)$-based measures is considerably higher than the robustness of the $R\left(G\right)$-based ranking algorithms. Moreover, it can be inferred from Figure 9 that when links are deleted from the network or added to the network all numerical instances of the KC index are equally robust. Thus, the choices of the Katz parameter given by the equations (5)–(8) do not affect the robustness of the resolvent-based centrality measures. On the other hand, in both models, it can be noted that, among the $R\left(G\right)$-based indices, the RKC $\left(\alpha_{0.85}\right)$ centrality measure is the most robust, whereas the RKC $\left(\alpha_{rk}\right)$ index is

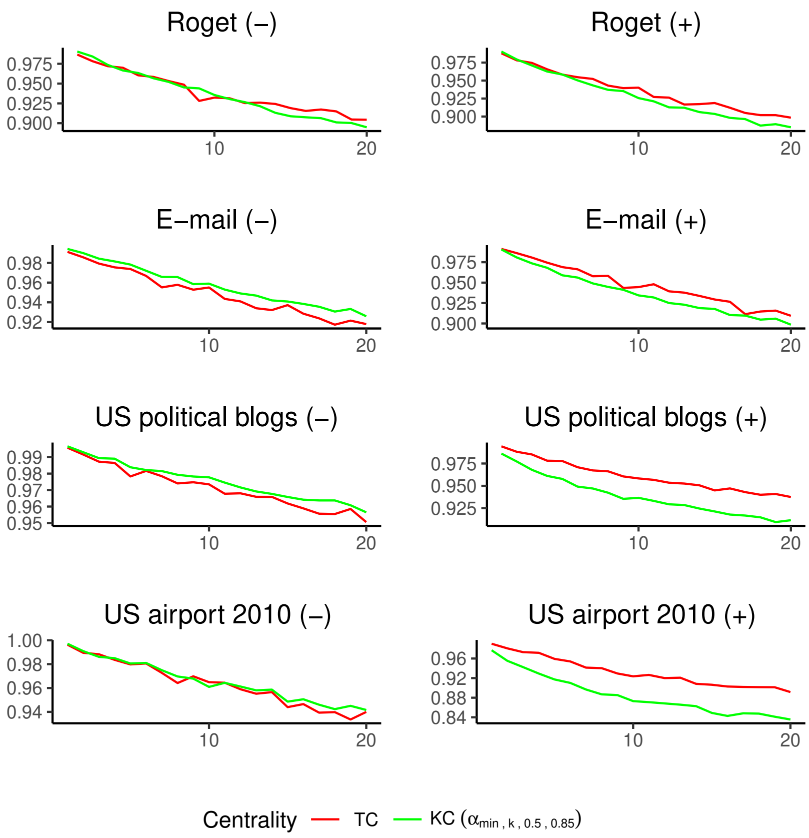

Fig. 9. The robustness of the TC index and four numerical instances of the KC index estimated on four empirical networks in the random edge removal model ($-$) as well as in the random edge addition model ($+$). In all subplots, the x axis corresponds to the number of edges deleted from the network or added to the network and the y axis corresponds to the robustness $\tau\left(C_{true}, C_{noisy}\right)$. The results are averages based on 50 simulation trials

the least robust. The effect of the values of the weighted Katz parameter given by the equations (18)–(21) on the robustness of the normalized resolvent-based vertex significance ranking algorithms is moderate in the random edge removal model and considerable in the random edge addition model.

Next, we will study the robustness of the communicability-based centrality indices with respect to the random edge rewiring. Figures 11 and 12 present the relationship between the rewiring probability p and $\tau\left(C_{true}, C_{noisy}\right)$. Figure 11 shows that all numerical instances of the KC index are (approximately) equally

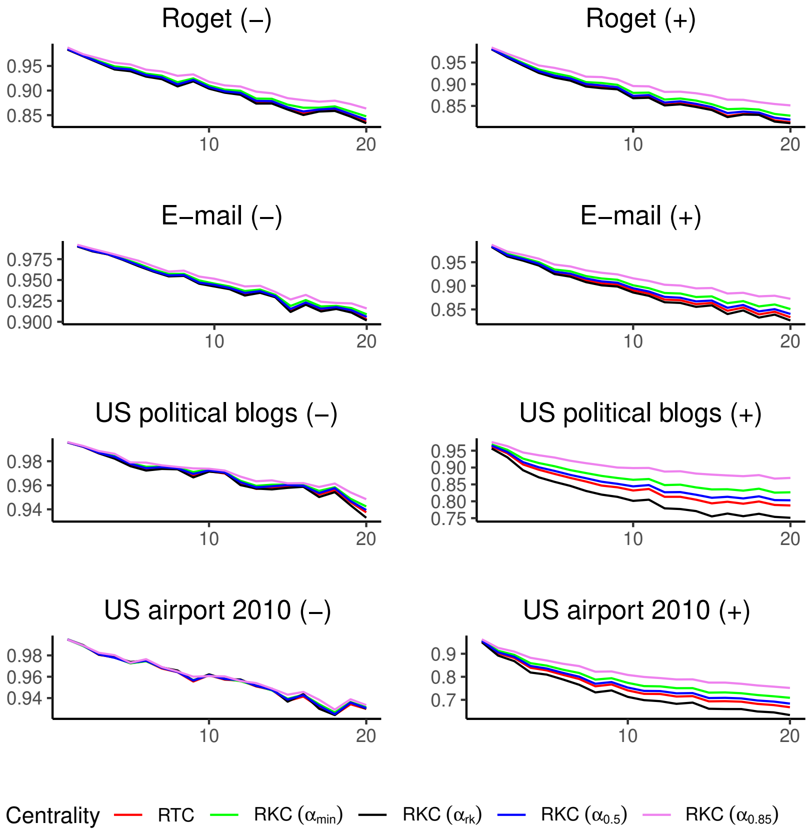

Fig. 10. The robustness of the weighted communicability-based centrality indices estimated on four empirical networks in the random edge removal model $(-)$ as well as in the random edge addition model $(+)$. In all subplots, the x axis corresponds to the number of edges deleted from the network or added to the network and the y axis corresponds to the robustness $\tau\,(C_{true}, C_{noisy})$. The results are averages based on 50 simulation trials

robust in the random edge rewiring model. Moreover, it can be seen that, in the overwhelming number of cases, the robustness of the $A\,(G)$-based centrality measures is higher than the robustness of their normalized variants. In turn, the choices of the weighted Katz parameter given by the equations (18)–(21) strongly affect the robustness of the normalized communicability-based ranking algorithms. Namely, among the $R\,(G)$-based centrality indices, the $RKC\ (\alpha_{rk})$ index is the least robust, whereas the $RKC\ (\alpha_{0.85})$ index is the most robust in the face of the random edge rewiring. Thus, the $RKC\ (\alpha_{rk})$ ranking algorithm which is the

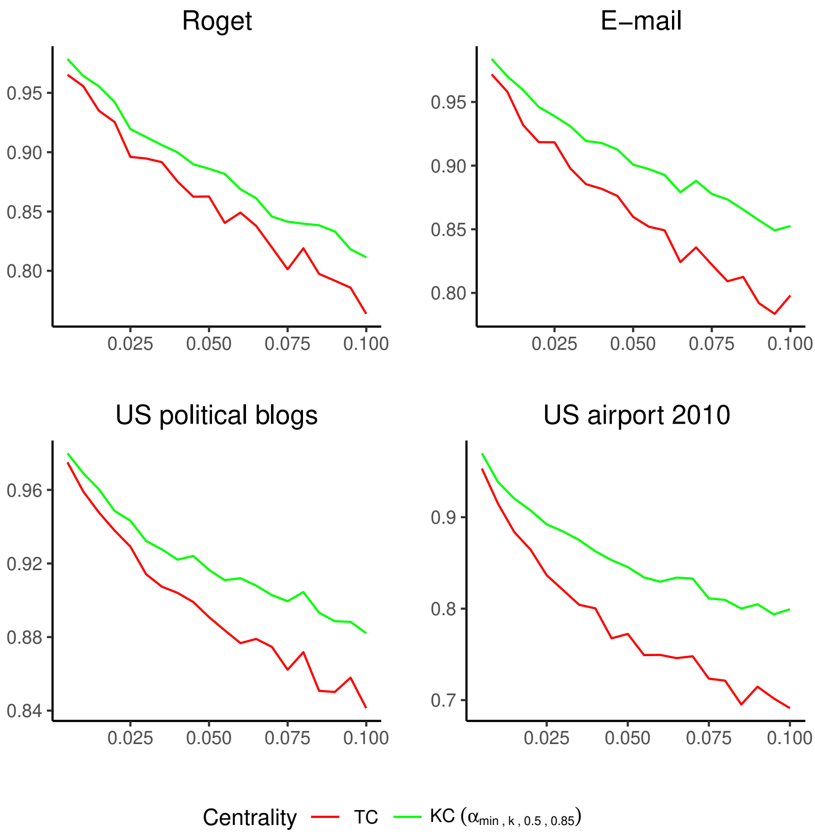

Fig. 11. The robustness of the TC index and four numerical instances of the KC index estimated on four empirical networks in the random edge rewiring model. In all subplots, the x axis corresponds to the rewiring probability (p) and the y axis correspond to the robustness $\tau(C_{true}, C_{noisy})$. The results are averages based on 50 simulation trials

most effective in dismantling the empirical networks is simultaneously the least robust against three types of measurement errors considered in the present work. Therefore, it can be hypothesized that, at least for some node importance ranking algorithms, there is a connection between the efficacy in attacking the networks and the robustness against some random link errors.

To compare the robustness of the communicability-based ranking algorithms to some reference metric, we will assess the robustness against three types of measurement errors of the classical DC measure. These results are included in Table 16.

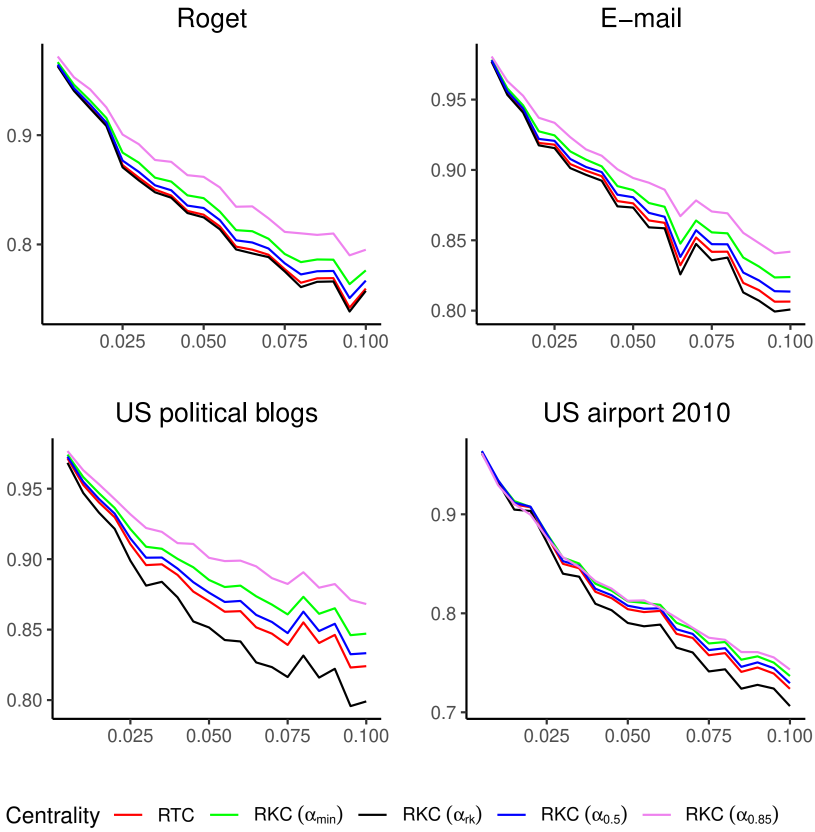

Fig. 12. The robustness of the weighted communicability-based centrality indices estimated on four empirical networks in the random edge rewiring model. In all subplots, the x axis corresponds to the rewiring probability (p) and the y axis correspond to the robustness $\tau\left(C_{true}, C_{noisy}\right)$. The results are averages based on 50 simulation trials

In this table, the robustness of the DC index is quantified when $\frac{m}{10}$ links are deleted from the network with m edges or added to the network with m edges as well as when links in the network are rewired with $p = 0.1$. These data indicate that the robustness of the DC measure is higher than the robustness of the communicability-based ranking algorithms. Moreover, it can be noticed that the robustness of the DC index is higher in the random edge removal model than in the random edge addition model. Nevertheless, similarly as in the case of the communicability-based ranking methods, the intensity of this asymmetrical phenomenon substantially hinges on the network.

Table 16
The robustness of the classical degree centrality measures estimated on four empirical networks

Type of error	*Roget*	*E-mail*	*US political blogs*	*US airport 2010*
Removal	0.9298	0.9517	0.9717	0.965
Addition	0.9206	0.9276	0.9242	0.8449
Rewiring	0.86	0.8927	0.9047	0.8448

Another key finding of the current contribution is that, in all cases, the intensity of the considered measurement errors is (approximately) linearly correlated with the variation in the values of $\tau\left(C_{true}, C_{noisy}\right)$. The Pearson correlation coefficient between the intensity of the perturbing factor and the alternations in the values of $\tau\left(C_{true}, C_{noisy}\right)$ is always above 0.94. The implication of this finding is that, at least in principle, if the rate and type of some measurement errors in network datasets are known, then it is possible to construct some confidence intervals around centrality scores.

5.5. The assortativeness of the tested ranking algorithms

In this subsection, we will study the assortativeness of the real-world networks and their rewired models with respect to the assortativity indices given by the equations (1) and (2). In order to compare the degree and centrality assortativity coefficients calculated on the empirical networks with some reference quantities, we will follow the methodology of null models described in Subsection 5.3. Table 17 includes the values of the $A_k(G)$ coefficient estimated on the real-world and rewired networks with the corresponding z-scores.

From this table, it can be easily observed that all empirical and rewired networks are degree-neutral or degree-weakly disassortative. In three cases, the values of the $A_k(G)$ index calculated on the real-world networks are higher than the values of this index estimated on the model networks. In all cases, the absolute values of the corresponding z-scores are above 2. Therefore, it can be uttered that the $A_k(G)$ coefficient of four exemplary networks does not depend on their degree distribution. Tables 18 and 19 contain the values of the $A_C(G)$ index where C is one of the communicability-based centrality measures estimated on the real-world networks and on their rewired models, respectively.

Table 17
The assortativity with respect to the degree centrality index of four empirical and four randomly rewired model networks with the corresponding z-score statistics. In the case of the rewired networks, the results are averages based on 50 simulation trials

Index	*Roget*	*E-mail*	*US political blogs blogs*	*US airport 2010 2010*
$A_k(G)$	0.174	0.0782	−0.2213	−0.1134
Index	*roget.rand*	*email.rand*	*blogs.rand*	*airport.rand*
$A_k(G)$	−0.0087	−0.0193	−0.1296	−0.2632
z-score	12.4972	8.943	−20.9301	37.8161

Table 18
The assortativity with respect to the communicability-based centrality indices estimated on four empirical networks

Index	*Roget*	*E-mail*	*US political blogs*	*US airport 2010*
TC	0.475	0.344	0.1111	0.1509
KC ($\alpha_{\min}$)	0.1941	0.0865	−0.2206	−0.1118
KC (α_k)	0.1823	0.0806	−0.2212	−0.1128
KC ($\alpha_{0.5}$)	0.184	0.0823	−0.221	−0.1126
KC ($\alpha_{0.85}$)	0.191	0.0852	−0.2207	−0.112
RTC	−0.0819	−0.112	−0.3079	−0.3194
RKC ($\alpha_{\min}$)	−0.0194	−0.0655	−0.3023	−0.3118
RKC (α_{rk})	−0.0956	−0.1302	−0.3096	−0.3186
RKC ($\alpha_{0.5}$)	−0.0589	−0.0952	−0.3073	−0.3188
RKC ($\alpha_{0.85}$)	0.0747	0.0019	−0.2807	−0.2707

In sum, Tables 18 and 19 include 40 measurements of $A_C(G)$ on the real-world networks and 40 measurements of $A_C(G)$ on the model networks. The corresponding z-scores comparing the measurements carried out on the empirical and generated networks are included in Table 34 in [52]. These data show that the measurements of the $A_C(G)$ coefficient classify the real-world networks into three different assortativity levels (cf. Table 1 in Section 2): centrality-neutral (60 %), centrality-weakly assortative (5 %) and centrality-weakly disassortative (35 %). In turn, the same measurements classify the rewired networks into two different assortativity levels: centrality-neutral (75 %) and centrality-weakly disassortative

Table 19
The assortativity with respect to the exponential-based and resolvent-based centrality indices derived from the adjacency and product connectivity matrices of four randomly rewired model networks. The results are averages based on 50 simulation trials

Index	*roget.rand*	*email.rand*	*blogs.rand*	*airport.rand*
TC	0.1848	0.0863	−0.1166	−0.2557
KC ($\alpha_{\min}$)	0.0078	−0.014	−0.1295	−0.2632
KC (α_k)	−0.0025	−0.0179	−0.1296	−0.2632
KC ($\alpha_{0.5}$)	−0.0005	−0.0166	−0.1296	−0.2632
KC ($\alpha_{0.85}$)	0.0053	−0.0148	−0.1295	−0.2632
RTC	−0.1472	−0.1119	−0.1536	−0.2798
RKC ($\alpha_{\min}$)	−0.1061	−0.0832	−0.1486	−0.2768
RKC (α_{rk})	−0.1622	−0.1281	−0.1576	−0.2816
RKC ($\alpha_{0.5}$)	−0.1325	−0.1015	−0.1519	−0.279
RKC ($\alpha_{0.85}$)	−0.0544	−0.0481	−0.1420	−0.2713

(25 %). Furthermore, the $A_C(G)$ index calculated with respect to the original communicability-based centrality measures was in 80 % higher in the empirical networks than in the model networks. In turn, the $A_C(G)$ index calculated according to the weighted forms of these node significance ranking algorithms was in 45 % higher in the real-world networks than in the synthetic networks. From Table 34 in [52], it follows that the measurements of the $A_C(G)$ index induced by the original communicability-based centrality measures possess always the values of the z-scores above 2. On the other hand, the measurements of the same index calculated according to the weighted counterparts of these measures have in 75 % of the recorded cases the values of the z-scores above 2. Consequently, it can be claimed that, in the overwhelming majority of instances, the centrality-centrality correlation coefficients induced by the vertex importance ranking algorithms defined by the matrix functions under consideration does not depend on the degree distribution of the underlying networks. To corroborate the hypothesis from Section 2 stating that the assortativity coefficients based on the centrality measures other than the DC index convey some useful information on the structural properties of the networks, we juxtaposed the scores included in Tables 18 and 19. From this comparison, it can be inferred that the degree and centrality assortativity coefficients estimated on the actual networks have in 25 % of the recorded cases the opposite sign and in 20 % of the recorded cases they classify the networks into different assortativity levels. In turn, the same coefficients calculated on the

rewired model networks possess in 10 % of the recorded instances the opposite sign and in all recorded instances they classify the networks into the same assortativity level. Therefore, it can be concluded that, in many cases, the centrality-centrality correlation coefficients conditioned by the centrality measures defined *via* the matrix functions under study do not duplicate the structural information contained in the degree-degree correlation coefficient and give some new structural insights about the networks.

Thus, from the above facts and computational results included in [3,6,20,32,34, 35,53], it follows that, in many cases, the assortativity indices calculated according to the equations (1) and (2) often vary considerably. Therefore, we propose to define the so-called *centrality assortativity profile* of a complex network. This profile consists of the values of the centrality-centrality correlation coefficients induced by the node significance ranking algorithms of interest and can be perceived as a way to characterize the networks. Since the assortativeness of the empirical complex networks calculated according to the equations (1) and (2) is (generally) different, it seems especially justifiable to consider a specific pattern of the centrality assortativity coefficients induced by several different vertex importance ranking methods in order to gain a more accurate picture of the structural relationships within the networks.

In order to further test if the centrality assortativity profiles of the exemplary complex networks induced by the considered communicability-based centrality measures depend on the degree distribution of the underlying graphs, we will carry out the Principal Component Analysis using the randomly rewired model networks with the preserved degree sequences as null models. In the first series of our measurements, each network (real-world or rewired) is identified with the 5-dimensional feature vector comprising five centrality-centrality correlation indices calculated with respect to the $A(G)$-based centrality metrics. In the second series of our measurements, each network (real-world or rewired) is identified with the 5-dimensional feature vector consisting of five centrality-centrality correlation indices calculated with respect to the $R(G)$-based centrality indices. Figures 13 and 14 illustrate four two-dimensional projections of the ensembles composed of 51 networks (i.e., one empirical network and 50 model networks) from the centrality assortativity profile spaces conditioned by the tested centrality metrics.

From Figures 13 and 14, it can be seen that, in all cases, the real-world networks are located outside the region of the centrality assortativity profile space spanned by the randomly rewired model networks with the conserved degree distribution. Thus, it can be concluded that the PCA outcomes clearly demonstrated that the centrality assortativity profiles determined by the communicability-based node

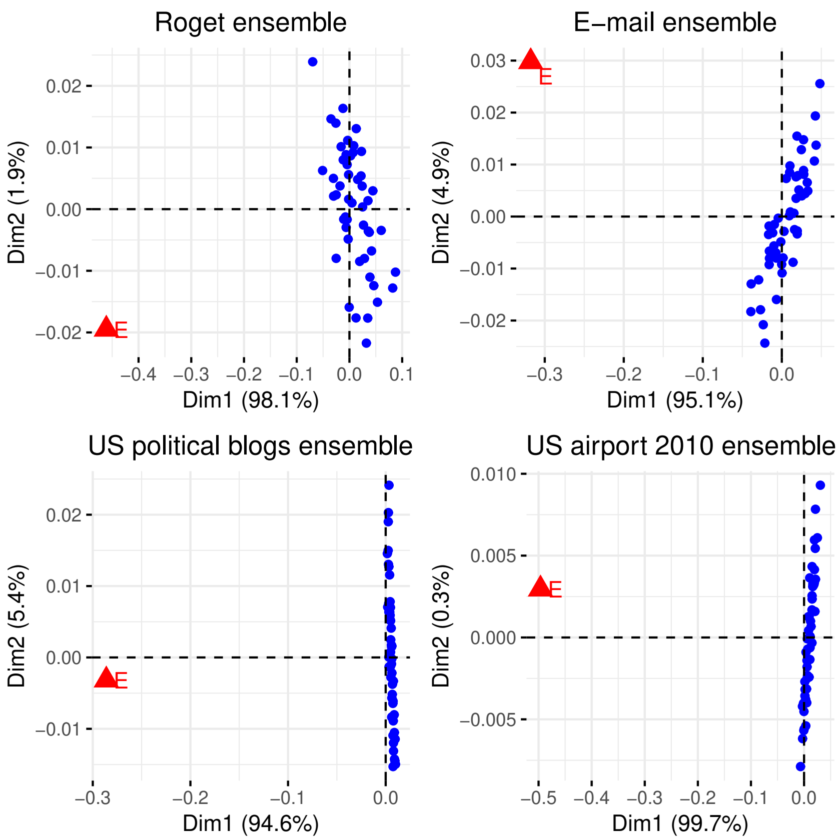

Fig. 13. The PCA projections of the centrality assortativity profile spaces determined by the original communicability-based centrality indices estimated on the empirical and randomly rewired model networks. In all subplots, the real-world network is denoted by the letter E and is marked by the red triangle symbol

importance rank algorithms can be used to distinguish between the empirical complex networks and their randomly rewired counterparts possessing the same degree sequences.

Therefore, it can be contended that the centrality assortativity profiles of the exemplary networks conditioned by the communicability-based centrality indices are independent on the degree distribution of the underlying graphs and can be regarded as new (*composite*) structural invariants characterizing the networks. Accordingly, it is evident that the hypothesis stating that the centrality-centrality correlation coefficients induced by the centrality measures other than the *DC* index do not duplicate the structural information contained in the degree-degree

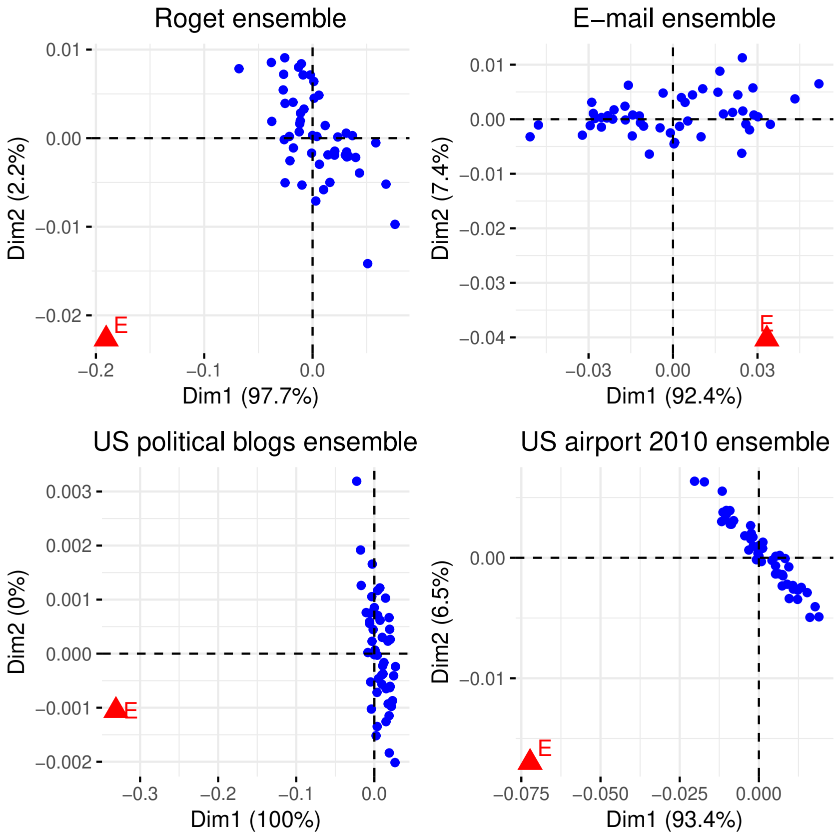

Fig. 14. The PCA projections of the centrality assortativity profile spaces determined by the weighted communicability-based centrality indices estimated on the empirical and randomly rewired model networks. In all subplots, the real-world network is denoted by the letter E and is marked by the red triangle symbol

correlation coefficient is empirically corroborated and, consequently, the centrality assortativity profile can be understood as a useful conceptual tool in complex network mining.

6. Conclusions

The concept of a centrality measure is at the forefront of *network science* for more than two decades. In the present article, we have proposed two novel vertex significance ranking algorithms – the Randić-Katz centrality measure and the

Randić total communicability centrality measure. These two newly introduced centrality metrics can be viewed as weighted (normalized) versions of the indices developed by L. Katz [29] as well as by M. Benzi and C. Klymko [10]. They are defined in terms of the resolvent and exponential functions applied to the product connectivity matrix of the network. The intensional attack simulation experiments included in the present paper unambiguously showcased that, by substituting the binary adjacency matrix by its weighted (normalized) version in the definitions of the indices proposed in [10, 29], we arrive at the conceptualization of two more accurate ranking algorithms. Furthermore, the rank correlation analyses unquestionably indicated that there is no trivial relationship between the novel centrality measures derived from the product connectivity matrix and their original counterparts. Consequently, it can be uttered that the newly introduced centrality indices give some novel structural insights about the networks. In our contribution, we defined two novel ranking methods by using the normalized adjacency matrix where each edge is weighted by its Randić weight. For future works, it would be interesting to consider the centrality measures whose definitions are based on other weighted adjacency matrices. For instance, the Randić weight could be replaced by its generalized (i.e., the quantity $(k_i k_j)^r$ where r is any real number) or additive (i.e., the quantity $(k_i + k_j)^{-0.5}$) counterparts.

We found that, at least for some networks and some ranking algorithms, the more topologically equivalent nodes a network contains, the lower the granularity of the centrality measure estimated on that network is. Thus, we established a bridge between *algebraic* and *quantitative graph theory*.

Moreover, the present paper proposes the notion of the *centrality assortativity profile* of a complex network. This novel conceptual tool in *network science* can be perceived as a new composite graph-theoretical invariant which globally characterizes the ways in which nodes with given centrality values are connected within the network. The computational studies illustrated that the *centrality assortativity profile* of a network does not duplicate the topological information contained in the classical degree assortativity coefficient and, therefore, can be regarded as a valuable theoretical concept in *network science*.

Acknowledgment

The author would like to express his thanks to the anonymous reviewer for his valuable comments and tips.

References

1. Adamic L.A., Glance N.: *The political blogosphere and the 2004 US Election.* In: Proceedings of the WWW-2005 Workshop on the Weblogging Ecosystem 2005.
2. Albert R., Jeong H., Barabási A.-L.: *Error and attack tolerance of complex networks.* Nature **406** (2000), 378–382.
3. Allen-Perkins A., Pastor J.M., Estrada E.: *Two-walks degree assortativity in graphs and networks.* Appl. Math. Comput. **311** (2017), 262–271.
4. Andreotti J., Jann K., Melie-Garcia L., Giezendanner S., Abela E., Wiest R., Dierks T., Federspiel A.: *Validation of network communicability metrics for the analysis of brain structural networks.* PloS ONE **9**, no. 12 (2014), e115503.
5. Aprahamian M., Higham D.J., Higham N.J.: *Matching exponential-based and resolvent-based centrality measures.* J. Complex Netw. **4**, no. 2 (2015), 157–176.
6. Arcagni A., Grassi R., Stefani S., Torriero A.: *Higher order assortativity in complex networks.* Eur. J. Oper. Res. **262** (2017), 708–719.
7. Auguie B.: *gridExtra: Miscellaneous functions for 'Grid' graphics. R package version 2.3* (2017). https://CRAN.R-project.org/package=gridExtra.
8. Badham J.M.: *Commentary: Measuring the shape of degree distribution.* Netw. Sci. **1**, no. 2 (2013), 213–225.
9. Batagelj V., Mrvar A.: *Pajek datasets* (2006). http://vlado.fmf.uni-lj.si/pub/networks/data.
10. Benzi M., Klymko C.: *Total communicability as a centrality measure.* J. Complex Netw. **1** (2013), 124–149.
11. Benzi M., Klymko C.: *A matrix analysis of different centrality measures.* arXiv:1312.6722v3 (2014).
12. Bozkurt S.B., Güngör A.D., Gutman I.: *Randić matrix and Randić energy.* MATCH Commun. Math. Comput. Chem. **64** (2010), 239–250.
13. Borchers H.W.: *pracma. Practical numerical math functions. R package version 2.1.1* (2017). https://CRAN.R-project.org/package=pracma.
14. Crofts J.J., Higham D.J.: *A weighted communicability measure applied to complex brain networks.* J. R. Soc. Interface **6** (2009), 411–414.
15. Csardi G., Nepusz T.: *The igraph software package for complex network research.* InterJournal Complex Syst. (2006), 1695.

16. Durón C.: *Heatmap centrality: A new measure to identify super-spreader nodes in scale-free networks.* PLoS ONE **15**, no. 7 (2000), e0235690.

17. Estrada E.: *The Structure of Complex Networks: Theory and Applications.* Oxford University Press, Oxford 2011.

18. Estrada E., Hatano N., Benzi M.: *The physics of communicability in complex networks.* Phys. Rep. **514** (2012), 89–119.

19. Estrada E., Rodrígues-Velázquez J.A.: *Subgraph centrality in complex networks.* Phys. Rev. E **71** (2005), 056103.

20. Goh K.I., Oh E., Kahng B., Kim D.: *Betweenness centrality correlation in social networks.* Phys. Rev. E **67** (2003), 017101.

21. Guimera R., Danon L., Diaz-Guilera A., Giralt F., Arenas A.: *Self-similar community structure in a network of human interactions.* Phys. Rev. E **68** (2003), 065103(R).

22. Higham N.J.: *Functions of Matrices. Theory and Computation.* Society for Industrial and Applied Mathematics, Philadelphia 2008.

23. Holme P., Kim B.J., Yoon C.N., Han S.K.: *Attack vulnerability of complex networks.* Phys. Rev. E **65** (2002), 056109.

24. Ibnoulouafi A., El Haziti M.: *Density centrality: identifying influential nodes based on area density formula.* Chaos Soliton Fract. **114** (2018), 69–80.

25. Ibnoulouafi A., El Haziti M., Cherifi H.: *M-centrality: identifying key nodes based on global position and local degree variation.* J. Stat. Mech.: Theory and Exp. **2018** (2018), 073407.

26. Iyer S., Killingback T., Sundaram B., Wang Z.: *Attack robustness and centrality of complex networks.* PLoS ONE **8**, no. 4 (2013), e59613.

27. Kassambara A.: *ggpubr: 'ggplot2' based publication ready plots. R package version 0.1.8* (2018). https://CRAN.R-project.org/package=ggpubr.

28. Kassambara A., Mundt F.: *factoextra: Extract and visualize the results of multivariate data analyses. R package version 1.0.5* (2017). https://CRAN.R-project.org/package=factoextra.

29. Katz L.: *A new status index derived from sociometric data analysis.* Psychometrika **18** (1953), 39–43.

30. Kirkland S.J., Neumann M.: *Group Inverses of M-Matrices and their Applications.* CRC Press, Boca Raton 2013.

31. Liu B., Huang Y., Feng J.: *A note on the Randić spectral radius.* MATCH Commun. Math. Comput. Chem. **68** (2012), 913–916.

32. Lv M., Guo X., Chen J., Lu Z.-M., Nie T.: *Second-order centrality correlation in scale-free networks*. Int. J. Mod. Phys. C **26**, no. 10 (2015), 1550116.

33. Martin C., Niemeyer P.: *Influence of measurement errors on networks: Estimating the robustness of centrality measures*. Netw. Sci. **7**, no. 2 (2019), 180–195.

34. Mayo M., Abdelzaher A., Ghosh P.: *Long-range degree correlations in complex networks*. Comput. Soc. Netw. **2** (2015), article number 4.

35. Meghanathan N.: *Assortativity analysis of real-world network graphs based on centrality metrics*. Computer and Information Science **9**, no. 3 (2016), 7–25.

36. Mueller L.A.J., Kugler K.G., Dander A., Graber A., Dehmer M.: *QuACN: an R package for analyzing complex biological networks quantitatively*. Bioinformatics **27** (2011), 140–141.

37. Newman M.E.J.: *Mixing patterns in networks*. Phys. Rev. E **67**, no. 2 (2003), 026126.

38. Newman M.E.J.: *Networks. An Introduction*. Oxford University Press Inc., New York 2010.

39. Niu Q., Zeng A., Fan Y., Di Z.: *Robustness of centrality measures against network manipulation*. Physica A **438** (2015), 124–131.

40. Noldus R., Van Mieghem P.: *Assortativity in complex networks*. J. Complex Netw. **3**, no. 4 (2015), 507–542.

41. Opsahl T.: *Why anchorage is not (that) important: Binary ties and sample selection*. http://wp.me/poFcY-Vw.

42. Opahl T., Agneessens J., Skvoretz J.: *Node centrality in weighted networks: Generalizing degree and shortest paths*. Soc. Netw. **32** (2010), 245–251.

43. Plemmons R.J.: *M-matrix characterizations. I-nonsingular M-matrices*. Linear Algebra Appl. **18** (1977), 175–188.

44. R Core Team: *R: a Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna 2017, http://www.R-project.org/.

45. Ronqui J.R.F., Travieso G.: *Analyzing complex networks through correlations in centrality measurements*. J. Stat. Mech.: Theory Exp. **2015** (2015), P05030.

46. Steyvers M., Tenenbaum J.B.: *The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth*. Cogn. Sci. **29** (2005), 41–78.

47. Takemoto K., Oosawa C.: *Introduction to complex networks: measures, statistical properties, and models.* In: Statistical and Machine Learning Approaches for Network Analysis. Dehmer M., Basak S.C. (eds.), John Wiley & Sons, New Jersey 2012.

48. Todeschini R., Consonni V.: *Molecular Descriptors for Chemoinformatics.* Wiley-VCH, Weinheim 2009.

49. Wickham H.: *ggplot2: Elegant Graphics for Data Analysis.* Springer-Verlag, New York 2009.

50. Wilczek P.: *Novel centrality measures and distance-related topological indices in network data mining.* Silesian J. Pure Appl. Math. **7** (2017), 21–63.

51. Wilczek P.: *Identifying influential nodes in the genome-scale metabolic networks.* Silesian J. Pure Appl. Math. **9** (2019), 9–47.

52. Wilczek P.: *The communicability-based centrality measures* (2020). https://repod.icm.edu.pl/dataset.xhtml?persistentId=doi:10.18150/J0G6U.

53. Zhou S., Cox I.J., Hansen L.K.: *Second-order assortative mixing in social networks.* In: Complex Networks VIII: Proceedings of the 8th Conference on Complex Networks. CompleNet 2017. Gonçalves B., Menezes R., Sinatra R., Zlatic V. (eds.), Springer Int. Publ., Cham, 2017.