A *k*-shell decomposition method for weighted networks

Antonios Garas[1,3], Frank Schweitzer[1] and Shlomo Havlin[2]

[1] ETH Zurich, Weinbergstrasse 58 CH-8092 Zurich, Switzerland
[2] Minerva Center and Department of Physics, Bar-Ilan University, 52900 Ramat Gan, Israel
E-mail: agaras@ethz.ch

Abstract. We present a generalized method for calculating the k-shell structure of weighted networks. The method takes into account both the weight and the degree of a network, in such a way that in the absence of weights we resume the shell structure obtained by the classic k-shell decomposition. In the presence of weights, we show that the method is able to partition the network in a more refined way, without the need of any arbitrary threshold on the weight values. Furthermore, by simulating spreading processes using the susceptible-infectious-recovered model in four different weighted real-world networks, we show that the weighted k-shell decomposition method ranks the nodes more accurately, by placing nodes with higher spreading potential into shells closer to the core. In addition, we demonstrate our new method on a real economic network and show that the core calculated using the weighted k-shell method is more meaningful from an economic perspective when compared with the unweighted one.

[3] Author to whom any correspondence should be addressed.

IOP Institute of Physics Φ DEUTSCHE PHYSIKALISCHE GESELLSCHAFT

Contents

1. Introduction

The continuously growing interest in complex network science has resulted over the last few years in novel methods of analysis for a large number of complex systems in various scientific fields [1–7]. The fundamental view of this interdisciplinary approach is that large complex systems can be described as complex networks (or graphs in mathematics terminology) where the nodes (or vertices) represent the system's interacting elements and the links (or edges) represent their interactions. This unified view was used in the analysis of social [7–9], biological [10–13], physiological [14], technological [15, 16], climate [17–19], economic [20–23] and financial systems [24, 25]. In combination with the technological advances that made enormously detailed data available, we are now able to understand and model the evolution of dynamical processes, such as epidemic outbreaks and information spreading [26–30].

Even the earliest empirical works in this field made it clear to researchers that the topology of a network affects its properties. For example, networks with broad degree distributions are more robust to random failures, but are fragile under intentional attacks [31–35]. Nowadays, there is a growing body of literature trying to understand global properties of a network by focusing on the properties of individual nodes, and their connectivity patterns [36]. Of course, the role of individual nodes has a profound relation with the evolution of any dynamical process, and with the evolution of the network itself. For example, very popular individuals in a social network (i.e. individuals with a large number of connections) usually attract more attention and increase their connectivity even more. While it is clear that such processes affect the evolution of the network topology, we can imagine that such individuals could assume key roles in the case of disease spreading, etc.

It is clear that questions such as 'who are the most important nodes in the network?' are natural to ask. Such questions can be addressed using centrality measures, which are the most frequently used measures when it comes to quantitative network analysis. However, there is a variety of centrality measures aiming to address the question of node 'importance'. For example, there is the *degree centrality* (or just the degree of a node, i.e. the number of its links), the *eigenvector centrality* [37], the *betweenness centrality* [38], the *closeness centrality* [39], etc. In this paper, we focus on a centrality measure based on the notion of k-cores which is a fundamental concept in graph theory [40] when it comes to ranking the centrality of nodes in a

complex network. Such a ranking was applied in many real networks [21, 41–48], allowing a thorough investigation of their structure, while highlighting the role of various topology-dependent processes.

One major limitation of most centrality measures, including the k-core decomposition method, is their design to work on unweighted graphs. However, in practice, real networks are weighted, and their weights describe important and well-defined properties of the underlying systems. In a weighted network, nodes have (at least) two properties that can characterize them, their degree and their weight. However, since weights are properties of the network's links, the node's weight is calculated as the sum over all link weights passing through a particular node. These two properties, even though in some cases they are correlated, are in general independent. As a result, nodes with high degree can have small weight (i.e. they have many connections to other nodes but the links of these connections have small weights), while there could also be nodes with small degree and high weight. Situations where the weights play an important role occur, for example, in economic or trade networks. In such networks, the weights are related to some measured property (such as trade flow, capital flow, etc), and in many cases one wishes to focus on nodes with high weights that are (usually) the most important players. Thus, in such systems the presence of nodes with high degree and relatively small weights may influence the results obtained by methods that are based only on the degree. In such cases, two main approaches have been used, with both having their own drawbacks. Under the first approach, one completely neglects the weights and performs the analysis on the unweighted network, but doing so one chooses to neglect an important property of the network. The second approach would be to consider only links with weights above some—(usually) arbitrary chosen—threshold value and filter out the rest. The drawback of this approach is the selection of a proper cutoff value, which may remove important high degree nodes with links of low weights (below the threshold), and as we will discuss later, this could have significant impact on the results. Additionally, by neglecting links below a threshold, the network becomes sparser with some nodes getting disconnected and not considered by the applied method afterwards.

In this paper, we aim to overcome these limitations by introducing a generalized method for calculating the k-shell structure of weighted networks. The paper is organized as follows. First we discuss the standard k-shell decomposition method, and right after that we introduce our generalized version. Next we apply both methods on real networks and present their results. Subsequently, we compare in more detail the performance of both methods in ranking nodes according to their importance when it comes to spreading processes, and finally we present the conclusions.

2. The unweighted *k*-shell decomposition method

The k-core/k-shell decomposition method partitions a network into sub-structures that are directly linked to centrality [49]. This method assigns an integer index, k_s, to each node that is representative of the location of the node in the network, according to its connectivity patterns. Nodes with low/high values of k_s are located at the periphery/center of the network. This way, the network is described by a layered structure (similar to the structure of an onion), revealing the full hierarchy of its nodes. The innermost nodes belong to a structure called the core or 'nucleus' of the network, while the remaining nodes are placed into more external layers (k-shells).
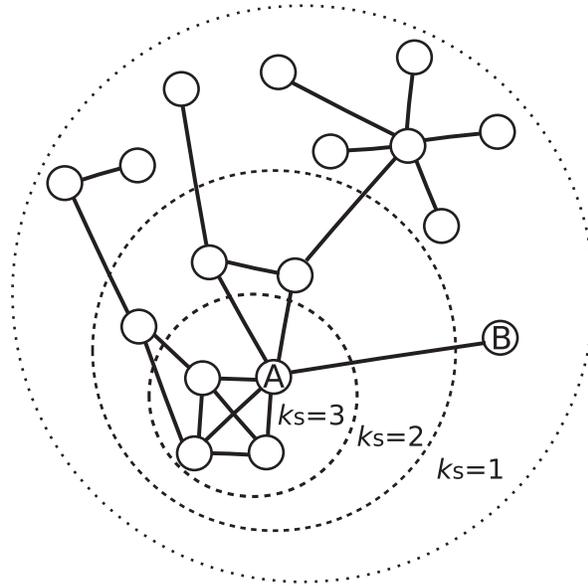
Figure 1. Illustration of the layered structure of a network, obtained using the k-shell decomposition method. The nodes between the two outer rings compose shell 1 ($k_s = 1$), while the nodes between the two inner rings compose shell 2 ($k_s = 2$). The nodes within the central ring constitute the core, in this case $k_s = 3$.

A more detailed description of how a network is divided into this k-shell structure is as follows (see figure 1). First we remove recursively from the network all nodes with degree $k = 1$, and we assign the integer value $k_s = 1$ to them. This procedure is repeated iteratively until there are only nodes with degree $k \geqslant 2$ left on the network. Subsequently, we remove all nodes with degree $k = 2$ and assigns to them the integer value $k_s = 2$. Again, this procedure is repeated iteratively until there are only nodes with degree $k \geqslant 3$ left on the network, and so on. This routine is applied until all nodes of the network have been assigned to one of the k-shells. This is how the original k-shell decomposition method works, which, as described above, does not consider the weights of the links at all; therefore, from now on we will call it the unweighted k-shell decomposition method ($U_{k\text{-shell}}$).

3. The weighted *k*-shell decomposition method

Here we propose a generalization of the k-shell decomposition method, which we call the weighted k-shell decomposition method ($W_{k\text{-shell}}$). This method applies the same pruning routine that was described earlier, but is based on an alternative measure for the node degree. This measure considers both the degree of a node and the weights of its links, and we assign for each node a *weighted degree*, k'. The weighted degree of a node i is defined as

$$k_i' = \left[k_i^\alpha \left(\sum_j^{k_i} w_{ij} \right)^\beta \right]^{\frac{1}{\alpha+\beta}}, \tag{1}$$

where k_i is the degree of node i and $\sum_j^{k_i} w_j$ is the sum over all its link weights. In this paper, we discuss only the case when $\alpha = \beta = 1$, which treats the weight and the degree equally. The full exploration of the parameter space is beyond our scope and is left for future work. Therefore, for what follows, $k'_i = \sqrt{k_i \sum_j^{k_i} w_{ij}}$.

Using the above approach in the case of unweighted networks, where $w_{ij} = 1$, the weighted degree is equivalent to the node degree ($k' \equiv k$), and we resume the same network partitioning as with the $U_{k\text{-shell}}$ decomposition method. However, in order that a typical weighted link be regarded as of unit weight before we calculate k' using equation (1), we perform the following steps. First, we normalize all the weights with their mean value $\langle w \rangle$, next we divide the resulting weights with their minimum value, and we discretize them by rounding to the closest integer; this way the minimum link weight is equal to 1.[4]

In figure 1, we illustrate schematically the layered structure obtained by applying the $U_{k\text{-shell}}$ decomposition method in a graph. In order to highlight the weaknesses of the unweighted method, let us suppose that the network is weighted. For simplicity, we assume that all link weights are equal to 1, except for the weight of the link between nodes A and B, which is $w_{AB} = 3$. As illustrated in figure 1, the node B is located at the periphery of the network, even though it is strongly connected to one of the core nodes. In real networks such a strong link (three times the capacity of other links) means that this particular node is of more importance for the core, but this is not depicted in the layered structure calculated by the classical unweighted approach, since this node will be placed in the outermost shell ($k_s = 1$). However, if we apply the $W_{k\text{-shell}}$ decomposition method, then node B is assigned to $k_s = 2$ that is one shell away from the core of the network, highlighting its actual importance.

4. Application to real networks

In order to compare the results obtained from the $U_{k\text{-shell}}$ and the $W_{k\text{-shell}}$ decomposition method, we used as case studies the following four real networks:

 (i) *Corporate ownership network* (CON). This is an economic network linking 206 different countries. It is constructed [21] using the 616 000 direct or indirect subsidiaries of the 4000 world corporations with the highest turnover, based on the 2007 version of the ORBIS database obtained from the Bureau van Dijk Electronic Publishing (BvDEP)[5]. The network is weighted, and its weights represent the business ties among countries [21].

 (ii) *The collaboration network of scientist working in network science* (SCIE). This network contains the co-authorship relations of scientists working on network theory and experiment, as compiled by Newman [50]. The network is weighted, and its weights are assigned as described in [51].

(iii) *The neural network of the nematode C. elegans* (CEL). This network was compiled by Watts and Strogatz [52] using the original experimental data of White *et al* [53]. It is a weighted representation of the neural network of *C. elegans*.

(iv) *The US air transportation network* (AIR). This is a weighted network obtained by considering the 500 US airports with the largest amount of traffic from publicly available

---

[4] We also tested the effect of the normalization by dividing it by the minimum weight, and the results we obtained in terms of node positioning with or without the normalization were similar.
[5] Bureau van Dijk Electronic Publishing (BvDEP), http://www.bvdep.com/.

IOP Institute of Physics ⬥ DEUTSCHE PHYSIKALISCHE GESELLSCHAFT

Table 1. Statistical properties of the networks used in our analysis. Here N_N is the number of nodes, N_E is the number of edges, $\langle k \rangle$ is the average degree of the network nodes, d is the diameter, C is the clustering coefficient [52] and B is the network's betweenness [38, 55]. If the original network is disconnected, we only consider its LCC.

Network	N_N	N_E	$\langle k \rangle$	d	C	B
CON	206	2886	28.0	4	0.38	94.6
SCIE	379	914	4.82	17	0.43	952.9
CEL	297	2345	15.8	5	0.18	215.4
AIR	500	2980	11.92	7	0.35	496.7

Table 2. Comparison of the network hierarchies obtained by the $U_{k\text{-shell}}$ and $W_{k\text{-shell}}$ decomposition method. Here s^U and s^W is the total number of k-shells, while n_c^U and n_c^W are the total number of nodes in the cores obtained using the $U_{k\text{-shell}}$ and the $W_{k\text{-shell}}$, respectively. N_C is the number of common nodes in both cores, N_{UW} is the fraction of nodes of the core obtained by the $U_{k\text{-shell}}$ that also belong to the core obtained by the $W_{k\text{-shell}}$ and N_{WU} is the fraction of nodes of the core obtained by the $W_{k\text{-shell}}$ that also belong to the core obtained by the $U_{k\text{-shell}}$.

Network	s^U	s^W	n_c^U	n_c^W	N_C	N_{UW}	N_{WU}
CON	28	87	41	11	11	0.27	1
SCIE	8	10	9	13	9	1	0.69
CEL	10	21	119	26	26	0.22	1
AIR	29	257	35	31	28	0.8	0.9

data [54]. Nodes represent US airports and edges represent air travel connections among them. It reports the anonymized list of connected pairs of nodes and the weight associated with the edge, expressed in terms of the number of available seats on the given connection on a yearly basis.

In table 1, we provide some detailed statistical properties of the above networks. For our analysis, if not stated otherwise, when we talk about the network we refer to the largest connected component (LCC), and whenever we discuss network properties these are calculated from the LCC.

In table 2, we compare the network hierarchies obtained by applying the $U_{k\text{-shell}}$ and the $W_{k\text{-shell}}$ decomposition method. We observe that the $W_{k\text{-shell}}$ method yields a more refined partitioning (a larger number of k-shells) of the networks. This means that by applying this method we obtain more detailed information about the networks' internal structure, which is similar to using a high-resolution microscope to observe small-size structures of a larger system.

Furthermore, for three out of the four studied networks the core obtained with the $W_{k\text{-shell}}$ contains a smaller number of nodes, while these nodes are almost entirely part of the core obtained by the $U_{k\text{-shell}}$. This means that the weighted method in most cases is able to split the
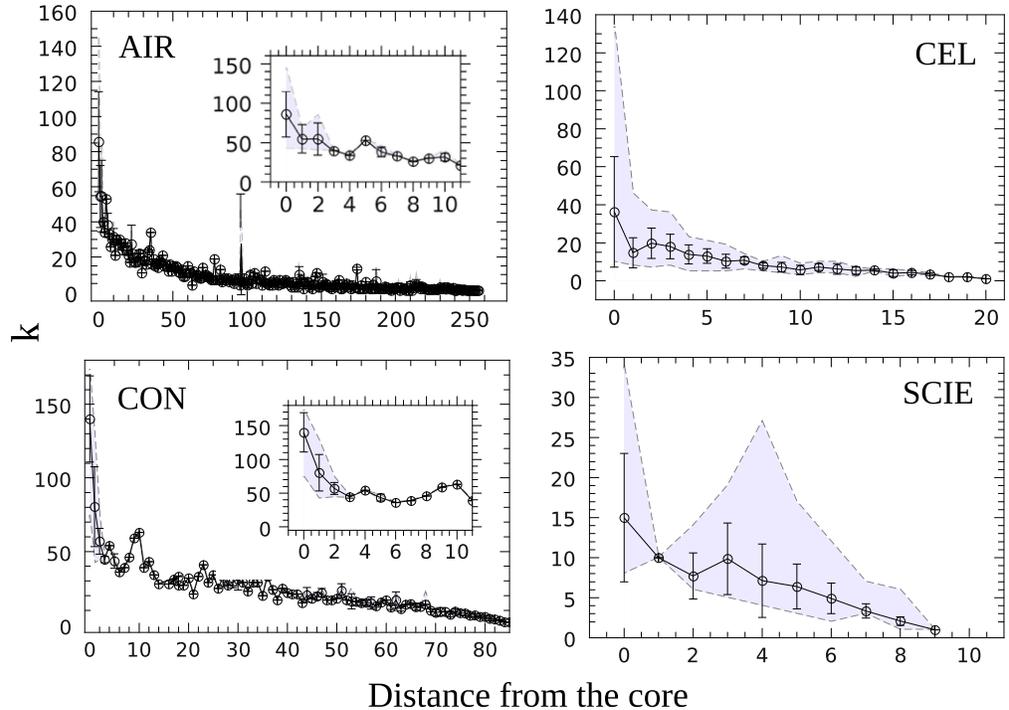
Figure 2. The average degree of all nodes in each shell, obtained using the $W_{k\text{-shell}}$ decomposition method. The shaded area highlights the full range of the degree values in each shell. The shells are ranked according to their distance from the core, and the error bars are showing the standard deviation. Insets: zoom to distances closer to the core for networks with a large number of shells.

cores obtained by the unweighted method further and to identify which are *most central of the central nodes*.

In figure 2, we plot the degrees of the nodes according to the k-shell they belong to (expressed as the distance from the core of the network). The node ranking is obtained using the $W_{k\text{-shell}}$ method for all the four different networks described above. As shown in figure 2, the degree is highly (and nonlinearly) correlated with the position of the node in the k-shell structure, but there are particular cases where the trend is not monotonic. This means that there are nodes with high degree that may not be as central to the network as one would expect; this is in line with our discussion for the example network of figure 1.

4.1. A detailed example: analysis of the core of an economic network

Next we compare the core of the $U_{k\text{-shell}}$ and the $W_{k\text{-shell}}$ decomposition methods applied to the global CON studied in [21]. The CON connects 206 countries around the globe, using as links the ownership relations within large companies. If companies listed in country A have subsidiary corporations in country B, there is a link connecting these two countries directed from country A to country B. The weight of the link, w_{AB}, equals the number of the subsidiary corporations in country B controlled by companies of country A.
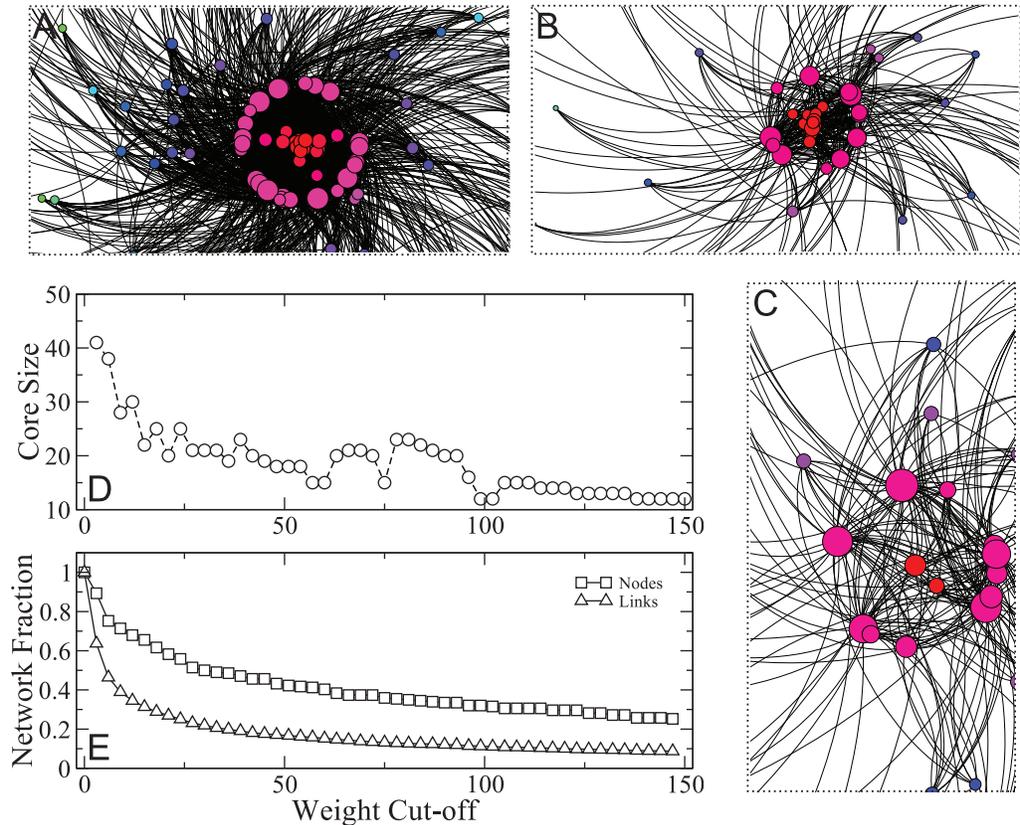
Figure 3. Changes in the CON network structure when using different weight cutoff values w_c. Panels (A)–(C) show the network snapshots around the central region for $w_c = 3$, 75 and 150, respectively. The size of the nodes is proportional to their degree. (D) Evolution of the core size as a function of w_c (after Garas *et al* [21]). (E) Fraction of nodes and links of the original network that remain for different w_c values.

Using the $U_{k\text{-shell}}$ decomposition method, as shown in table 2 and figure 3, we identify a core of 41 countries. However, we expect that in the current state of the global economy, a smaller set of countries are the major players (G8, G20, etc). In order to reduce the size of the core and to highlight which are the potentially more important nodes of this network by using the classic k-shell decomposition method, a cutoff value of $w_c = 100$ was assumed by Garas *et al* [21]. It was shown that the remaining network after filtering the links with $w_c < 100$ contains only 66 out of the original 206 nodes. However, a core formed by the following 12 countries: the United States of America (US), the United Kingdom (GB), France (FR), Germany (DE), the Netherlands (NL), Japan (JP), Sweden (SE), Italy (IT), Switzerland (CH), Spain (ES), Belgium (BE) and Luxembourg (LU), was identified. In figure 3 the evolution of the core and network size of the CON is shown, as a function of the weight cutoff value w_c.

Using the $W_{k\text{-shell}}$ decomposition method, we obtain the layered structure of the network including all the 206 nodes, without using any arbitrary cutoff parameter. The core of the network obtained with this method consists of the following 11 counties: US, GB, FR, DE, NL, JP, Canada (CA), IT, CH, ES and BE. Comparing these two cores we find a striking similarity.

The only two differences are the presence of CA in the core calculated using our new weighted k-shell approach, while SE and LU have moved to the second innermost layer. These differences can be well understood considering that CA is a major economy; it is part of G7, and all the other six members of G7 are already part of the core. Furthermore, CA outperforms SE and LU in terms of population and other macroeconomic indicators, such as total import/exports and GDP. It is thus natural to conclude that the core obtained using the $W_{k\text{-shell}}$ decomposition method is more meaningful from an economic perspective, since it groups together some of the largest (developed) global economies.

5. Dynamics: shell positioning and spreading potential

In recent years, models such as the susceptible infectious recovered (SIR) model [56] have been used extensively in network research in order to explore epidemic spreading [27, 56–58], economic crisis spreading [21] as well as information and rumor spreading [26, 28] in social processes. In such processes the topology of the network is not the only thing that matters; the position of the node where the spreading begins plays an important role as well. In the recent work of Kitsak *et al* [48], it was shown that the spreading power of a node cannot be predicted solely based on its degree. A better measure is its actual position in the network, as it is described by the k-shell where it belongs.

Using this perspective, it is reasonable to assume that a k-shell partitioning method provides us with a more accurate node ranking for representing the nodes' spreading power. In addition, since the individual nodes are grouped in k-shells, it is reasonable to assume that every k-shell should contain nodes with similar spreading power. In what follows, we will use these assumptions to evaluate and compare the performance of the $U_{k\text{-shell}}$ and $W_{k\text{-shell}}$ decomposition methods.

We modeled the spreading process by applying the SIR model on all the networks described above. However, since we are interested in the weights of the network, we used a version of the SIR model which takes into account the weight of the links that mediate the spreading. This model was originally introduced to simulate the spreading of an economic crisis [21]; for this model the probability of infection is different for every link and is calculated by

$$p_{ij} \propto m \cdot w_{ij}/\tilde{w}_j, \tag{2}$$

where w_{ij} is the weight of the link that connects the origin node i with the destination node j, and $\tilde{w}_j$ is the total weight ($\tilde{w}_j = \sum_i w_{ij}$) of the destination node j. The factor m is a free amplification parameter that can determine, for example, the severity of a crisis, how infectious a virus is, the importance of a rumor, etc. In what follows, we will call this model *weighted SIR* (W-SIR).

The modeling procedure of the W-SIR is the following. Initially, we assign all nodes to be susceptible (S) to an infection. Next, one node, i, is chosen and is assumed to be infected (I). This node will infect all its neighboring nodes with probability p_{ij} during the first time step. This causes all infected nodes to switch their status from S to I, while the node that initiated this process changes to the recovered state (R), and can no longer infect other nodes or become infected. At every consecutive time step the process is repeated, and all the infected nodes are trying to infect their susceptible (S) neighbors in the network. The process lasts until there are no infected nodes left in the network.
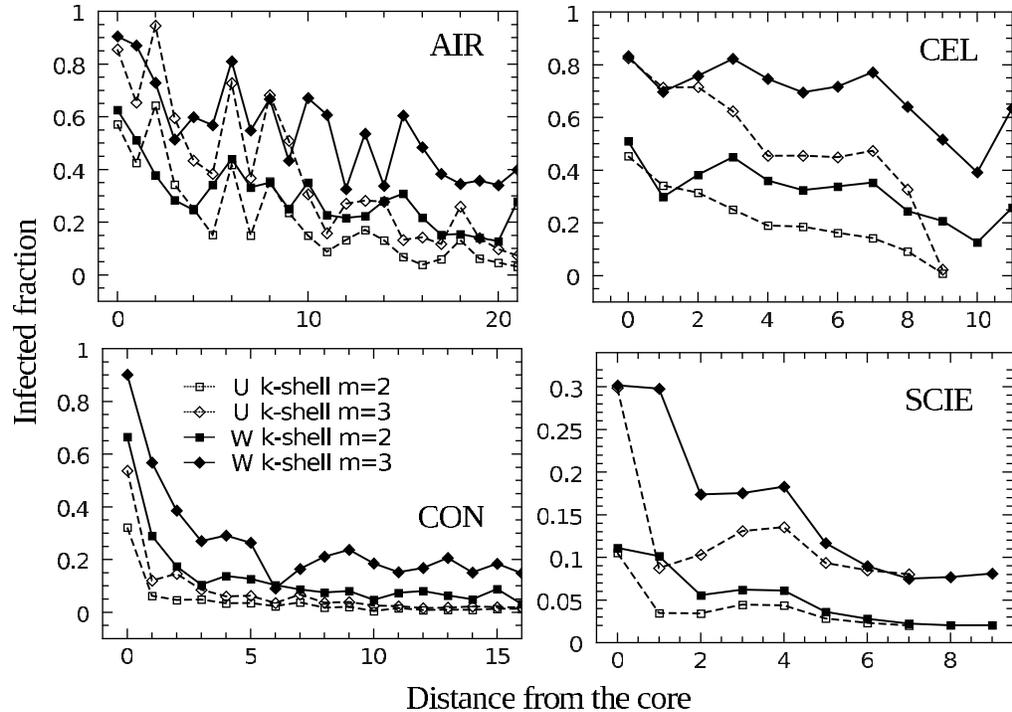
Figure 4. Average infected fraction of a k-shell versus the shell's distance from the core of the network.

For each individual node we performed 100 realizations of the W-SIR model, and we calculated the average infected fraction of the network for different values of $m \in [0, 10]$. This fraction is used as score in order to rank the nodes according to their spreading potential. We restricted ourselves to values of m in this interval, as for much larger m values the role of individual nodes is no longer important, and an epidemic outbreak emerges no matter where the infection starts. Next, we partitioned the network using the $U_{k\text{-shell}}$ and the $W_{k\text{-shell}}$ decomposition methods, and ranked the obtained k-shells according to their distance from the core. By calculating the average infected fraction that results from an epidemic starting separately from all nodes of every individual k-shell, we estimated the shell's spreading potential.

In figure 4, we study how the average infected fraction changes versus the distance of each k-shell from the core of the network for both methods. We find that, in general, the central k-shells obtained by the $W_{k\text{-shell}}$ method are more able to initiate a severe outbreak in comparison with the central k-shells obtained using the $U_{k\text{-shell}}$ method. This result is robust for all networks used in this study, and for different values of the parameter m. The above finding means that the $W_{k\text{-shell}}$ decomposition method positions the nodes with the higher average spreading potential in shells closer to the core.

Next, we tested how homogeneous are the obtained k-shells with respect to the spreading potential of their containing nodes. In order to do so, we calculated the standard deviation, σ, of a node's infected fraction (spreading potential) for every k-shell for a given value of the parameter m. Next we calculated the average value over all the shells, $\langle\sigma\rangle$, and we plotted it versus m (figure 5). We find that the average standard deviation of the spreading potential
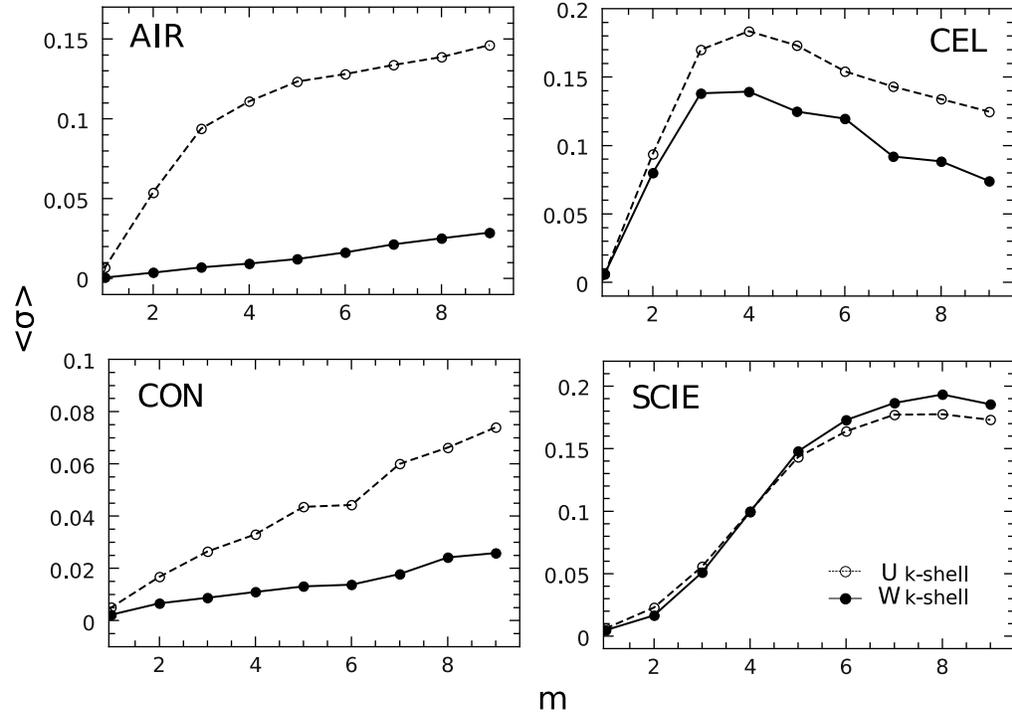
Figure 5. Average value of the spreading potential of nodes within a k-shell over all shells, $\langle \sigma \rangle$, versus m.

using W-SIR is always lower when we partition the network using the $W_{k\text{-shell}}$ method, with respect to partitioning using the $U_{k\text{-shell}}$ method. This means that the $W_{k\text{-shell}}$ method gives more *homogeneous k-shells*, where all nodes in the shell have similar importance for the dynamical process in question.

As a final step and given that the $W_{k\text{-shell}}$ method performs better in positioning the nodes according to their W-SIR spreading potential in weighted graphs, it is interesting to further explore the role of the weights in this process. To do so, we created ten realizations of the CON network with shuffled weights, and we performed 100 runs of the W-SIR model on every one of these ten networks. Next, we calculated the average spreading potential per k-shell using the infected fraction obtained by the implementation of W-SIR on the network with shuffled weights. As shown in figure 6, in the shuffled case the k-shells are becoming significantly more inhomogeneous, and their $\langle \sigma \rangle$ is always larger that the $\langle \sigma \rangle$ obtained by the original, unshuffled network. This procedure highlights the role of the weights in the process, since in the case when the weights do not to play any role these two curves should collapse into one.

6. Conclusion

In summary, we presented a generalized k-shell decomposition method ($W_{k\text{-shell}}$) that considers the link weights of networks, without applying any arbitrary cutoff threshold on their value. The method resumes the same shell structure obtained by the classic k-shell decomposition in the absence of weights, but when weights are present, it is able to partition the network in a more refined way. In its general formulation, our method allows us to vary the importance
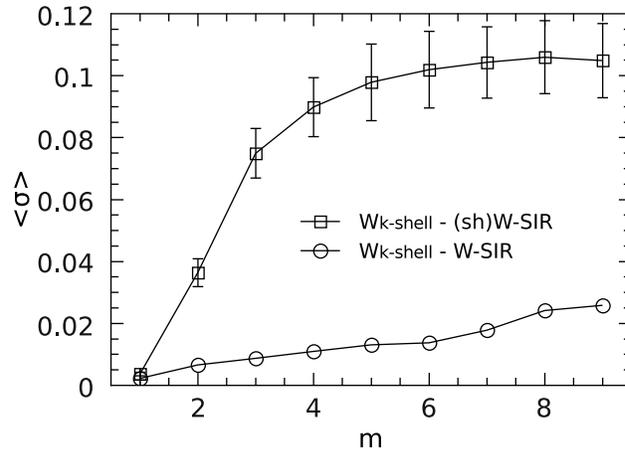
Figure 6. Comparison of $\langle \sigma \rangle$ versus m for two different configurations of the CON. $W_{k\text{-shell}}$—W-SIR is the original case (also shown in figure 5) where the nodes' spreading potential is obtained by applying the W-SIR in the original network. $W_{k\text{-shell}}$—(Sh)W-SIR is a case when we calculated the nodes' spreading potential by applying the W-SIR on the ten realizations of the CON with shuffled weights.

assigned to either the node weights or the node degree, by adjusting the exponents α and β of equation (1). While in this paper we did not fully explore the parameter space, we would like to stress that this additional flexibility provides a more accurate ranking for various applications. Here, using $\alpha = \beta = 1$ we showed that the partitioning obtained by the $W_{k\text{-shell}}$ method is particularly meaningful in terms of the spreading potential of the nodes. We demonstrated the weighted version of the SIR model in four different networks, and showed that nodes with higher spreading potential were positioned in the core or in shells closer to the core better in comparison with the $U_{k\text{-shell}}$ method.

Acknowledgments

References

[1] Albert R and Barabási A L 2002 Statistical mechanics of complex networks *Rev. Mod. Phys.* **74** 47–97
[2] Dorogovtsev S N and Mendes J F F 2003 *Evolution of Networks: From Biological Nets to the Internet and WWW* (Oxford: Oxford University Press)
[3] Caldarelli G 2007 *Scale-Free Networks: Complex Webs in Nature and Technology* (Oxford: Oxford University Press)
[4] Barrat A, Barthélemy M and Vespignani A 2008 *Dynamical Processes on Complex Networks* (Cambridge: Cambridge University Press)

[5]   Cohen R and Havlin S 2010 *Complex Networks: Structure, Robustness and Function* (Cambridge: Cambridge University Press)

[6]   Newman M E J 2010 *Networks: An Introduction* (Oxford: Oxford University Press)

[7]   Jackson M O 2010 *Social and Economic Networks* (Princeton, NJ: Princeton University Press)

[8]   Snijders T A B, Pattison P E, Robins G L and Handcock M S 2006 New specifications for exponential random graph models *Sociol. Methodol.* **36** 99–153

[9]   Borgatti S P 2006 Identifying sets of key players in a social network *Comput. Math. Organ. Theory* **12** 21–34

[10] Milo R, Shen-Orr S, Itzkovitz S, Kashtan N, Chklovskii D and Alon U 2002 Network motifs: simple building blocks of complex networks *Science* **298** 824–7

[11] Alon U 2003 Biological networks: the tinkerer as an engineer *Science* **301** 1866–7

[12] Khanin R and Wit E 2006 How scale-free are biological networks *J. Comput. Biol.* **13** 810–8

[13] Gallos L K, Makse H A and Sigman M 2012 A small world of weak ties provides optimal global integration of self-similar modules in functional brain networks *Proc. Natl Acad. Sci. USA* **109** 2825–30

[14] Bashan A, Bartsch R P, Kantelhardt J W, Havlin S and Ivanov P C 2012 Network physiology reveals relations between network topology and physiological function *Nature Commun.* **3** 702

[15] Reed D A, Kapur K C and Christie R D 2009 Methodology for assessing the resilience of networked infrastructure *Syst. J. IEEE* **3** 174–80

[16] Johansson J and Hassel H 2010 An approach for modelling interdependent infrastructures in the context of vulnerability analysis *Reliab. Eng. Syst. Saf.* **95** 1335–44

[17] Tsonis A A and Swanson K L 2008 Topology and predictability of El Niño and La Niña networks *Phys. Rev. Lett.* **100** 228502

[18] Donges J F, Zou Y, Marwan N and Kurths J 2009 The backbone of the climate network *Europhys. Lett.* **87** 48007

[19] Gozolchiani A, Havlin S and Yamasaki K 2011 Emergence of El Niño as an autonomous component in the climate network *Phys. Rev. Lett.* **107** 148501

[20] Schweitzer F, Fagiolo G, Sornette D, Vega-Redondo F, Vespignani A and White D R 2009 Economic networks: the new challenges *Science* **325** 422–5

[21] Garas A, Argyrakis P, Rozenblat C, Tomassini M and Havlin S 2010 Worldwide spreading of economic crisis *New J. Phys.* **12** 113043

[22] Vitali S, Glattfelder J B and Battiston S 2011 The network of global corporate control *PLoS One* **6** e25995

[23] Harmon D, Stacey B and Bar-Yam Y 2010 Networks of economic market interdependence and systemic risk arXiv:1011.3707v2

[24] Bonanno G, Caldarelli G, Lillo F and Mantegna R 2003 Topology of correlation-based minimal spanning trees in real and model markets *Phys. Rev.* E **68** 046130

[25] Garas A, Argyrakis P and Havlin S 2008 The structural role of weak and strong links in a financial market network *Eur. Phys. J.* B **63** 265–71

[26] Daley D J and Kendall D G 1965 Stochastic rumours *IMA J. Appl. Math.* **1** 42–55

[27] Colizza V, Barrat A, Barthélemy M and Vespignani A 2006 The role of the airline transportation network in the prediction and predictability of global epidemics *Proc. Natl Acad. Sci. USA* **103** 2015–20

[28] Castellano C, Fortunato S and Loreto V 2009 Statistical physics of social dynamics *Rev. Mod. Phys.* **81** 591–646

[29] Yang R, Huang L and Lai Y-C 2008 Selectivity-based spreading dynamics on complex networks *Phys. Rev.* E **78** 026111

[30] Yang R, Zhou T, Xie Y-B, Lai Y-C and Wang B-H 2008 Optimal contact process on complex networks *Phys. Rev.* E **78** 066109

[31] Albert R, Jeong H and Barabási A L 2000 Error and attack tolerance of complex networks *Nature* **406** 378–82

[32] Cohen R, Erez K, Ben-Avraham D and Havlin S 2000 Resilience of the internet to random breakdowns *Phys. Rev. Lett.* **85** 4626–8

[33] Callaway D S, Newman M E, Strogatz S H and Watts D J 2000 Network robustness and fragility: percolation on random graphs *Phys. Rev. Lett.* **85** 5468–71

[34] Cohen R, Erez K, Ben-Avraham D and Havlin S 2001 Breakdown of the internet under intentional attack *Phys. Rev. Lett.* **86** 3682–5

[35] Gallos L, Cohen R, Argyrakis P, Bunde A and Havlin S 2005 Stability and topology of scale-free networks under attack and defense strategies *Phys. Rev. Lett.* **94** 188701

[36] Park J and Barabási A L 2007 Distribution of node characteristics in complex networks *Proc. Natl Acad. Sci. USA* **104** 17916–20

[37] Bonacich P 1987 Power and centrality: a family of measures *Am. J. Sociol.* **92** 1170–82

[38] Freeman L C 1977 A set of measures of centrality based on betweenness *Sociometry* **40** 35–41

[39] Newman M E J 2005 A measure of betweenness centrality based on random walks *Soc. Netw.* **27** 39–54

[40] Bollobás B 1984 *Graph Theory and Combinatorics: Proc. Cambridge Combinatorial Conf. in Honour of Paul Erdös* (New York: Academic)

[41] Seidman S B 1983 Network structure and minimum degree *Soc. Netw.* **5** 269–87

[42] Bader G D and Hogue C W V 2003 An automated method for finding molecular complexes in large protein interaction networks *BMC Bioinformatics* **4** 2

[43] Wuchty S and Almaas E 2005 Peeling the yeast protein network *Proteomics* **5** 444–9

[44] Dorogovtsev S N, Goltsev A V and Mendes J F F 2006 *k*-core organization of complex networks *Phys. Rev. Lett.* **96** 040601

[45] Carmi S, Havlin S, Kirkpatrick S, Shavitt Y and Shir E 2007 A model of internet topology using *k*-shell decomposition *Proc. Natl Acad. Sci. USA* **104** 11150–4

[46] Alvarez-Hamelin J, Dall'Asta L, Barrat A and Vespignani A 2008 K-core decomposition of Internet graphs: hierarchies, self-similarity and measurement biases *NHM* **3** 371–93

[47] Shao J, Buldyrev S, Braunstein L, Havlin S and Stanley H E 2009 Structure of shells in complex networks *Phys. Rev.* E **80** 036105

[48] Kitsak M, Gallos L K, Havlin S, Liljeros F, Muchnik L, Stanley H E and Makse H A 2010 Identification of influential spreaders in complex networks *Nature Phys.* **6** 888–93

[49] Batagelj V and Zaveršnik M 2011 Fast algorithms for determining (generalized) core groups in social networks *Adv. Data Anal. Classif.* **5** 129–45

[50] Newman M E J 2006 Finding community structure in networks using the eigenvectors of matrices *Phys. Rev.* E **74** 036104

[51] Newman M E J 2001 Scientific collaboration networks II. Shortest paths, weighted networks, and centrality *Phys. Rev.* E **64** 016132

[52] Watts D J and Strogatz S H 1998 Collective dynamics of 'small-world' networks *Nature* **393** 440–2

[53] White J G, Southgate E, Thomson J N and Brenner S 1986 The structure of the nervous system of the nematode *Caenorhabditis elegans Phil. Trans. R. Soc. B* **314** 1–340

[54] Colizza V, Pastor-Satorras R and Vespignani A 2007 Reaction–diffusion processes and metapopulation models in heterogeneous networks *Nature Phys.* **3** 276–82

[55] Freeman L C 1979 Centrality in social networks conceptual clarification *Soc. Netw.* **1** 215–39

[56] Anderson R M and May R M 1992 *Infectious Diseases of Humans: Dynamics and Control* (Oxford: Oxford Science Publications)

[57] Hethcote H W 2000 The mathematics of infectious diseases *SIAM Rev.* **42** 599

[58] Newman M E J 2002 Spread of epidemic disease on networks *Phys. Rev.* E **66** 016128