Nested subgraphs of complex networks

Nested subgraphs of complex networks

Bernat Corominas-Murtra[1]**, José F F Mendes**[2] **and Ricard V Solé**[1,3]

[1] ICREA-Complex Systems Lab, Universitat Pompeu Fabra, Dr Aiguader 80, 08003 Barcelona, Spain
[2] Departamento de Física da Universidade de Aveiro, 3810-193 Aveiro, Portugal
[3] Santa Fe Institute, 1399 Hyde Park Road, New Mexico 87501, USA

E-mail: bernat.corominas@upf.edu

Abstract

We analytically explore the scaling properties of a general class of nested subgraphs in complex networks, which includes the K-core and the K-scaffold, among others. We name such a class of subgraphs K-nested subgraphs since they generate families of subgraphs such that $\dots S_{K+1}(\mathcal{G}) \subseteq S_K(\mathcal{G}) \subseteq S_{K-1}(\mathcal{G}) \dots$. Using the so-called *configuration model* it is shown that any family of nested subgraphs over a network with diverging second moment and finite first moment has infinite elements (i.e. lacking a percolation threshold). Moreover, for a scale-free network with the above properties, we show that any nested family of subgraphs is self-similar by looking at the degree distribution. Both numerical simulations and real data are analyzed and display good agreement with our theoretical predictions.

PACS numbers: 05.65.+b, 05.40.−a, 05.50.+q

1. Introduction

The internal organization of most complex systems displays some sort of nestedness associated with some type of hierarchical organization. Such patterns can be detected by using appropriate theoretical tools which help us in understanding the system's structure in terms of a network [1–7]. Furthermore, the structure of such communities can provide us with valuable information about invariant properties and potential universals. In this work we will define a general class of network substructures which we called the K-nested subgraph. Such a class of subgraphs includes the K-core, the K-scaffold or the random deletion of nodes. But it also includes any other substructure you can define, if a small set of probabilistic restrictions holds. We develop a general, unified framework that enables us to study generic properties of such K-nested subgraphs. As we should see, the most common class of real networks, with connectivity patterns following a power-law distribution $P(k) \propto k^{-\alpha}, 2 > \alpha > 3$, has very interesting properties when looking at subgraph nestedness. In this context, theoretical studies on the

resilience of both *K*-cores [4, 8, 9] and *K*-scaffolds [6, 10] suggest that arbitrarily large scale-free networks contain infinite, asymptotically self-similar, *K*-cores and *K*-scaffolds, indicating that such subgraphs are highly robust against the random deletion of nodes. Metaphorically, it has been suggested that the structure of complex nets is similar to a Russian doll [4].

These results are consistent with the mounting evidence indicating that scale-free networks exhibit general self-similar properties [4, 11–14]. From the physical point of view, the asymptotical invariance of the degree distribution of scale-free nets under nesting operations is one of their most salient properties. At the theoretical level, the conservation of $P(k)$ the degree distribution implies self-similarity, as far as most of the properties of a random graph are determined by its degree distribution [15]. Of course, real nets are not exactly random graphs, but such an approach revealed surprisingly adequate to study real systems [16]. Furthermore, self-similar properties and scaling laws might be an indication that such objects are organized near criticality [17, 18].

In this work we generalize previous approaches, showing that *any* nested family of subgraphs of a given scale-free network has an infinite percolation threshold, i.e., there is an infinite set of *Russian dolls* for such networks. Moreover, it can be shown that such families are self-similar. We develop such concepts under the framework of the so-called *configuration model* [19], which works on an ensemble of arbitrarily large, sparse and uncorrelated graphs with specific properties.

The remainder of this paper is organized as follows: first, we formally define the concept of *K-nested* subgraph and show how the above-mentioned examples hold the required conditions. Then, we derive the general percolation properties and the final generic form of an arbitrary nested subgraph of a given net. From the developed formalism, we apply our results to specific network topologies.

2. Nested subgraphs

Formally, a complex network is topologically described by a graph $\mathcal{G}(V, \Gamma)$ where V is the set of nodes and $\Gamma : V \rightarrow V$ is the set of edges connecting the nodes of V. If $P(k)$ is the probability that a randomly chosen node is connected to k other nodes, then

$$\langle k \rangle = \sum_{k}^{\infty} k P(k) \langle k^2 \rangle = \sum_{k}^{\infty} k^2 P(k)$$

is the average connectivity of $\mathcal{G}$ and the second moment of the distribution, respectively.

We will say that $S(A, \Gamma_A)$ is an induced subgraph of $\mathcal{G}(V, \Gamma)$ if $A \subseteq V$ and $\Gamma_A \subseteq \Gamma$, being Γ_A a mapping $\Gamma_A : A \rightarrow A$. We can define many subgraphs from a given graph. Here we are interested in a special set of subgraphs, hereafter *K-nested subgraphs*, which includes, as special cases, the family of successive *K*-cores or *K*-scaffolds and the so-called $\widehat{v}$-deletion graph, obtained by deleting a fraction $\widehat{v}$ of nodes. A *K*-nested family of subgraphs $\mathcal{N}$ is a collection of subgraphs of a given graph $\mathcal{G}$, $\mathcal{N} = \{S_1(\mathcal{G}), S_2(\mathcal{G}), \ldots, S_i(\mathcal{G}), \ldots\}$ such that

$$\ldots S_{K+1}(\mathcal{G}) \subseteq S_K(\mathcal{G}) \subseteq S_{K-1}(\mathcal{G}) \ldots \tag{1}$$

For every family of *K*-nested subgraphs we associate a *nesting function* $\varphi_K(k)$, namely the probability for a randomly chosen node with degree k to belong to S_K. If $U \subseteq \mathbf{R}$ is a set that depends on the nature of the nesting, $\varphi_K(k)$ is such that

$$\varphi_K(k) : U \times N \rightarrow [0, 1]. \tag{2}$$

It is easy to see that for a function to be a nesting function, it has to fulfil the following logical conditions:

$$(\varphi_K(k') > \varphi_K(k)) \Rightarrow (k' > k), \tag{3}$$

$$(\varphi_{K'}(k) > \varphi_K(k)) \Rightarrow (K' < K), \tag{4}$$

$$(\forall \varphi_K)(\exists \lambda_{S_K} \in (0, 1]) | (\lim_{k \to \infty} \varphi_K(k)) = \lambda_{S_K}), \tag{5}$$

where λ_{S_K} is a scalar whose value will depend on the explicit form of S_K. In short, $\varphi_K(k)$ is a non-decreasing function on k (equation (3)) and a non-increasing function on K (equation (4)). Note that such a function implies that all the nodes satisfying the conditions are taken into account: our subgraphs are maximal under the conditions imposed by φ_K. Furthermore, note that, for a fixed K, $\varphi_K(k)$ has a horizontal asymptote at $\varphi_K(k) = \lambda_{S_K}$ (equation (5)). Thus

$$\lim_{k \to \infty} (\varphi_K(k+1) - \varphi_K(k)) = 0. \tag{6}$$

From (3)–(6) we can see that, for a fixed K, and $0 < \delta < 1$ there exists a k^* such that

$$(\forall k_i, k_j > k^*) \Rightarrow (||\varphi_K(k_i) - \varphi_K(k_j)|| < \delta), \tag{7}$$

and we can conclude that the sequence $(\varphi_K(k)) = \varphi_K(1), \varphi_K(2), \ldots, \varphi_K(i), \ldots$ is a Cauchy sequence. As we should see, this property will be useful in the following sections. Let us now explore some relevant nesting functions.

(a) *K-core subgraphs.* The K-core is the largest induced subgraph whose minimal connectivity is K (see figures 1(*b*) and 2(*b*)). Intuitively, it is clear that a collection of K-cores from a given graph $\mathcal{G}$ defines a nested family of subgraphs. Within the configuration model, we can informally identify the probability for a given node of $\mathcal{G}$ to belong to the giant K-core with the probability of belonging to an infinite $(K-1)$-ary subtree of $\mathcal{G}$ [4, 9, 20]. Therefore, the probability for a given node to belong to the K-core equals the probability of belonging to an infinite $(K-1)$-ary subtree. Let R be the probability that a given end of an edge is not the root of an infinite $(K-1)$-ary subtree. The associated nesting function for the K-core is $\varphi_K(k) = 0$, if $k < K$ and

$$\varphi_K(k) = \sum_{i=K}^{k} \binom{k}{i} R^{k-i} (1-R)^i \tag{8}$$

otherwise. It is straightforward to check that such a function follows (3)–(5).

(b) *K-scaffold subgraphs.* The K-scaffold of a given graph is the subgraph obtained by choosing all the nodes whose $k \geqslant K$ and the nodes that, despite their connectivity being $k < K$, they are connected to a node e' whose $k' \geqslant K$ [6, 10] (see figures 1(*a*) and 2(*c*)). The nesting function for the K-scaffold is $\varphi_K(k) = 1$, if $k \geqslant K$ and

$$\varphi_K(k) = 1 - \left(\sum_{k' < K} \frac{k' P(k')}{\langle k \rangle} \right)^k \tag{9}$$

otherwise. Note that, for both the K-nested families of K-scaffolds and K-cores, $\lambda_{S_K} = 1$. A variety of subgraphs can be defined from the K-scaffold, such as the *naked K-scaffold* (a subgraph obtained by cutting all the nodes whose degree is $k = 1$ in the K-scaffold).

(c) *Random deletion of nodes.* Suppose we delete a fraction $\widehat{\nu} = 1 - \nu$ of nodes from our graph. Such an operation can also be formalized in terms of nesting functions. For the sake of simplicity, if we perform a random deletion of a fraction of nodes from $\mathcal{G}$, we will indicate the nesting function and the subgraphs as φ_ν and S_ν, respectively. The associated
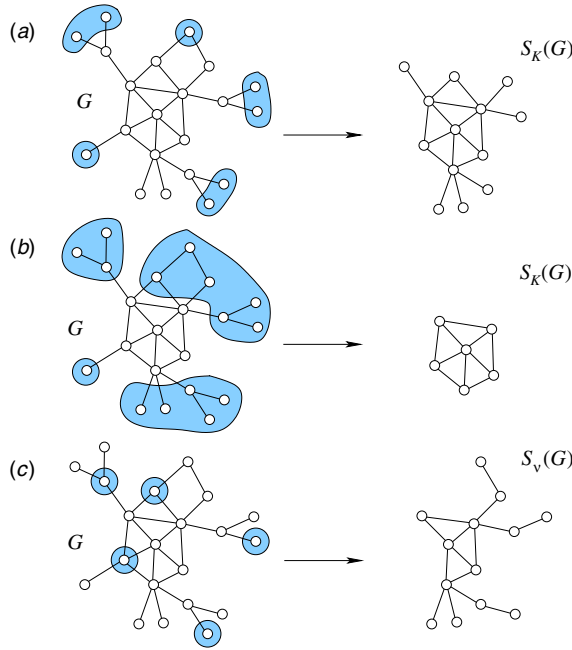
Figure 1. Some subgraphs samples that enable us to define a nested family of subgraphs. In the original graph (left) we shadowed the nodes that disappear under the operation of S_K. In the right-hand side, we display the giant component of the obtained graph, S_K. We find the K-scaffold, ($K = 3$) (*a*). The K-scaffold is the subgraph obtained by choosing all the nodes whose connectivity is equal or higher than K and all the nodes connected to them. Such a subgraph enables us to study the fundamental hub-connector structure of the complex networks. (*b*) The K-core ($K = 3$), the largest induced subgraph whose minimal connectivity is equal to K. (*c*) A subgraph obtained by randomly deleting a fraction ($\widehat{v} = 5/21$) of nodes (commonly referred by the literature as *random failures*.)

nesting function is, simply,

$$(\forall k)(\varphi_v(k) = v). \tag{10}$$

For mathematical purposes, let us introduce an additional class of subgraphs, $S_{K\gamma}$, of a given subgraph S_K. The main feature of such subgraphs is that $S_{K\gamma} \subseteq S_K$. We name such subgraphs *minor subgraphs* of S_K. To characterize such subgraphs, we say that $\gamma_K(k)$ is a minor nesting function of $\varphi_K(k)$ if $(\gamma_K(k) < \varphi_K(k))$ for all k. Given an arbitrary $\varphi_K(k)$, we can build a minor nesting function as follows: let k' be the minimum k such that $\varphi_K(k') \neq 0$ (it could be $k' = 1$). Then find an $\epsilon > 0$ such that $\epsilon < \varphi_K(k')$. Thus

$$\gamma_K(k) = \begin{cases} 0 & \text{if } k < k' \\ \epsilon & \text{if } k \geqslant k'. \end{cases} \tag{11}$$

This trivial way to define a minor subgraph from a given subgraph S_K is enough, since both $\gamma_K(k)$ and $\varphi_K(k)$ verify (3)–(5). Moreover, it is clear that[4] ($S_{K\gamma} \subseteq S_K$) for all K.

---

[4] Let us suppose a graph $\mathcal{G}$ and its two subgraphs, S_v, $S_{v'}$, obtained by deleting at random $\widehat{v} = 1 - v$ and $\widehat{v}' = 1 - v'$, with $v' > v$. Clearly, we cannot conclude that S_v is an induced subgraph of $S_{v'}$. But it is true that the properties of S_v will be, with high probability, the properties of some induced subgraph of $S_{v'}$ obtained by deleting at random $\widehat{v}$ nodes of $\mathcal{G}$. Furthermore, it can be shown that the probability of finding a diverging value decays exponentially with the size of the system—recall that we are working with an ensemble formalism.
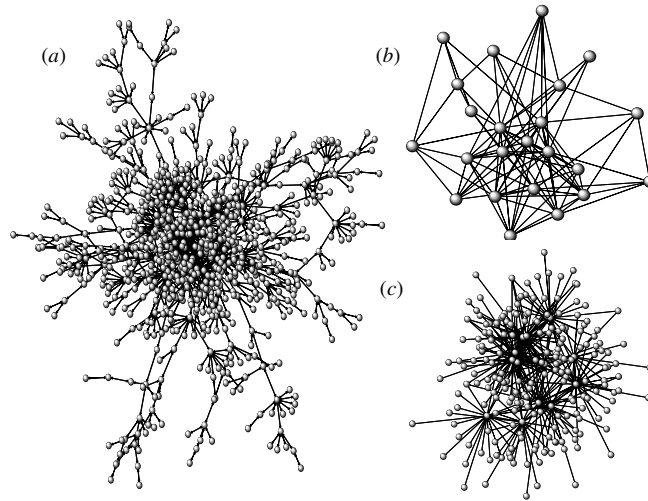
Figure 2. A complex network with broad distribution of links (*a*) and two nested subgraphs: (*b*) Its *K*-core ($K = 4$) and (*c*) the corresponding *K*-scaffold ($K = 20$).

3. Percolation of nested subgraphs

Previous to determining the specific statistical properties of the obtained subgraphs, we are interested in knowing whether there is a giant component in S_K, i.e., if the operation of nesting breaks (or not) the initial graph $\mathcal{G}$ into many small components. We consider first the general problem.

Let us define the generating functions for an arbitrary *K*-nested subgraph with the associated nesting function $\varphi_K(k)$ defined on $\mathcal{G}$ with arbitrary (but smooth) degree distribution $P(k)$:

$$F_0(z) = \sum_k^\infty P(k)\varphi_K(k)z^k, \tag{12}$$

$$F_1(z) = \frac{1}{\langle k \rangle} \sum_k^\infty k P(k)\varphi_K(k)z^{k-1}. \tag{13}$$

The averages—i.e. the values at $z = 1$ of equations (5) and (6)— are, respectively, $\mu \equiv F_0(1)$ and $\omega \equiv F_1(1)$. Here, μ is the fraction of nodes from $\mathcal{G}$ that belong to S_K. Similarly, ω is the relation among $\langle k \rangle$ and the average number of nodes from V reachable after computing the nested subgraph. The generating function for the size of components other than the giant component, which can be reached from a randomly chosen node is

$$H_1(z) = 1 - \omega + z F_1(H_1(z)) \tag{14}$$

and the generating function for the size of the component to which a randomly chosen node belongs is [16, 21]

$$H_0(z) = 1 - \mu + z F_0(H_1(z)), \tag{15}$$

thus, the average component size other than the giant component is

$$\langle s \rangle = H_0'(1) = \mu + F_0'(1)H_1'(1). \tag{16}$$

5

If we compute the derivative, it is straightforward to see that it leads to a singularity when $F_1'(1) = 1$. Thus, if

$$F_1'(1) = \frac{1}{\langle k \rangle} \sum_k k(k-1)\varphi_K(k)P(k) \tag{17}$$

to ensure the presence of a giant S_K, the following inequality has to hold:

$$\sum_k k(k-2)P(k) > \sum_k k(k-1)\widehat{\varphi}_K(k)P(k), \tag{18}$$

where $\widehat{\varphi}_K(k) = 1 - \varphi_K(k)$. This can be seen as the natural extension of the Molloy and Reed criterion [22] for any nested subgraph S_K, with the associated nesting function $\varphi_K(k)$. A more compact expression of such a criterion is

$$\sum_k k^2 \varphi_K(k)P(k) - (1 + \omega)\langle k \rangle > 0. \tag{19}$$

4. Degree distribution of S_K

The following step is to compute the degree distribution of the nested subgraphs, $P_{S_K}(k)$. The key question is finding the average number of nodes a given node will reach, if it survived to the computation of S_K. Taking into account the set of all nodes of $\mathcal{G}$, the average connectivity will decrease a factor $\omega \equiv F_1(1) = 1/\langle k \rangle \times \sum_k k\varphi_K(k)P(k)$. Clearly, the probability for a *surviving* node with connectivity k in $\mathcal{G}$ to display connectivity $k' \leqslant k$ in S_K, $\mathbf{P}(k \to k')$ is

$$\mathbf{P}(k \to k') = \binom{k}{k'}\omega^{k'}(1-\omega)^{k-k'}. \tag{20}$$

And, in the absence of correlations, a node with connectivity k in $\mathcal{G}$ now will *survive* with a probability $\varphi_K(k)$ and it will be connected, on average, to ωk nodes. If we take into account all the possible contributions of the nodes of $\mathcal{G}$ to the abundance of nodes with certain degree k in S_K, we have

$$P_{S_K}(k) = \frac{1}{\mu} \sum_{i \geqslant k}^{\infty} \varphi_K(i)\binom{i}{k}\omega^k(1-\omega)^{i-k}P(i), \tag{21}$$

where $P_{S_K}(k)$ is the probability of finding a node of degree k after the computation of S_K. Note that the factor $\frac{1}{\mu}$ normalizes $P_{S_K}(k)$. Clearly, if we define $\delta(\omega, \lambda_{S_K})$ as

$$\delta(\omega, \lambda_{S_K}) \equiv \frac{1}{\mu} \sum_{i \geqslant k}^{\infty} (\lambda_{S_K} - \varphi_K(i))\binom{i}{k}\omega^k(1-\omega)^{i-k}P(i).$$
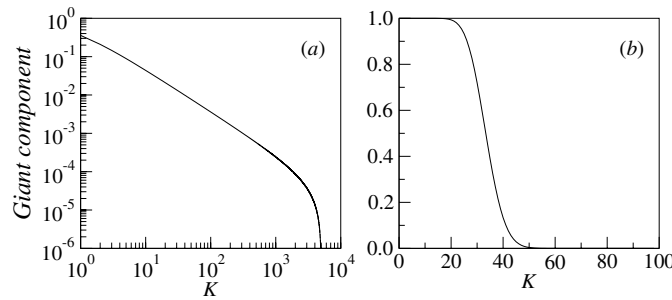
We can rewrite P_{S_K} as

$$P_{S_K}(k) = \frac{\lambda_{S_K}}{\mu} \sum_{i \geqslant k}^{\infty} \binom{i}{k}\omega^k(1-\omega)^{i-k}P(i) - \delta(\omega, \lambda_{S_K}). \tag{22}$$

But note that, due to relation (7), for large k's:

$$\frac{\lambda_{S_K}}{\mu} \sum_{i \geqslant k}^{\infty} \binom{i}{k}\omega^k(1-\omega)^{i-k}P(i) \gg \delta(\omega, \lambda_{S_K}). \tag{23}$$

Thus P_{S_K} is reduced to

$$P_{S_K}(k) \approx \frac{\lambda_{S_K}}{\mu} \sum_{i \geqslant k}^{\infty} \binom{i}{k}\omega^k(1-\omega)^{i-k}P(i). \tag{24}$$

Figure 3. The simplest family of nested subgraphs, obtained by removing all nodes whose connectivity is less than K: $\varphi_K(k) = \Theta(K, k)$, where $\Theta(K, k) = 1$ if $k \geqslant K$ and 0 otherwise. (*a*) Numerical computation of the size of the giant component $p_\infty = 1 - H_0(1) = \mu - F_0(u)$ where u is the first, non-trivial solution of $u = 1 - \omega + F_1(u)$, for $\varphi_K(k) = \Theta(K, k)$. This curve corresponds to a scale-free network with $\alpha \approx 2.15$. No specific scale is identified. The sharp decay for the large K values can be attributed to the finite size of the system. (In this simulation, we assumed $k_{\max} \approx 5000$.) (*b*) The same computation over an Erdös Rényi graph with $\langle k \rangle = 30$ displays a clear characteristic scale where the giant component is completely eliminated.

Let us rewrite equation (24) in order to extract analytical results. If the first generating function of the degree distribution of $\mathcal{G}$, without taking into account the nesting operation, is

$$G_0(z) = \sum_k^\infty P(k) z^k \tag{25}$$

it is straightforward that

$$\frac{\mathrm{d}^k}{\mathrm{d}z^k} G_0(z) = \sum_{i \geqslant k} \frac{i!}{(i-k)!} P(k) z^{i-k}. \tag{26}$$

Thus, we can rewrite the degree distribution (24) in terms of the derivatives of $G_0(z)$:

$$P_{S_K}(k) \approx \frac{\lambda_{S_K}}{\mu} \frac{\omega^k}{k!} \left. \frac{\mathrm{d}^k}{\mathrm{d}z^k} G_0(z) \right|_{z=1-\omega}. \tag{27}$$

In the following, we will apply our results to standard topologies of network theory: the Erdös Rényi graphs and the power-law graphs.

5. Erdös Rényi graphs

In the Erdös Rényi (E-R) graph,

$$P(k) = \frac{\langle k \rangle^k \, \mathrm{e}^{\langle k \rangle}}{k!} \tag{28}$$

and $\langle k^2 \rangle = \langle k \rangle^2$. To study specific percolation properties, we need to know the specific shape of $\varphi_K(k)$. In (figure 3(*b*)) we approached numerically the size of the giant component in an E-R graph where a successive nesting operation is performed. A clear threshold is observed, displaying a critical point where the giant connected component is completely eliminated. The special case of $\varphi_K(k) = \nu$ recovers the well-known percolation condition for E-R graphs under random damage, $\langle k \rangle > (1 + \nu)/\nu$. The predictions for the degree distribution are more

general and accurate. Indeed, the expression for $G_0(z)$ in E-R graphs is $G_0^{\mathrm{ER}}(z) = \mathrm{e}^{\langle k \rangle (z-1)}$. Thus, if, as we defined above, $\mu \equiv F_0(1)$:

$$P_{S_K}^{\mathrm{ER}}(k) \approx \frac{\lambda_{S_K}}{\mu} \frac{\langle \omega k \rangle^k \, \mathrm{e}^{\langle \omega k \rangle}}{k!}. \tag{29}$$

This implies that, for large k's, the nesting operation over an E-R graph results in an E-R graph but with a factor ω correcting the mean, whose value goes from $\langle k \rangle \to \omega \langle k \rangle$.

6. Scale-free nets

Let us assume a scale-free network with

$$P(k) \propto k^{-\alpha}, \tag{30}$$

with the scaling exponent $2 < \alpha < 3$. We will show that, at the thermodynamic limit, *any* family of subgraphs has infinite subgraphs. This has separately been shown for the *K*-core [4, 9] and the *K*-scaffold [6]. One of the main characteristics of such nets is that $\langle k^2 \rangle \to \infty$, and that $\langle k \rangle$ does not diverge with network size.

What we should prove is that, under these conditions, relation (19) holds *for all K*'s. In other words, there is no characteristic scale for the substructure generated by $\varphi_K(k)$. Indeed, our subgraphs need to fulfil the inequality

$$\sum_k k^2 \varphi_K(k) P(k) - (1 + \omega)\langle k \rangle > 0. \tag{31}$$

But we cannot work directly with an arbitrary nesting function φ_K. Thus, to prove the above claim, we build a minor nesting function $\gamma_K(k)$ of our $\varphi_K(k)$, as defined in (11), assuming k' as the smallest k such that $\varphi_K(k) > 0$. Thus, if $\omega_\gamma \equiv F_1^\gamma(1)$ has the form

$$\omega_\gamma = \epsilon \left(1 - \sum_{k < k'} \frac{k P(k)}{\langle k \rangle} \right) \equiv \epsilon', \tag{32}$$

the corresponding percolation condition for $S_{K\gamma}$ is

$$\epsilon \sum_{k \geqslant k'} k^2 P(k) - (1 + \epsilon')\langle k \rangle > 0.$$

But since $\langle k^2 \rangle$ diverges, we will have $\epsilon \sum_{k \geqslant k'} k^2 P(k) \to \infty$ and condition (19) always holds, provided that $\langle k \rangle$ is finite. This implies that percolation of any nested subgraph of an arbitrarily large scale-free network is guaranteed, as far as $S_{K_\gamma} \subseteq S_K$. Numerical simulations (see (figure 3(a))) of the size of the giant component display no critical scale for the emergence (elimination) of the giant connected component.

The above mathematical machinery will lead us to demonstrate that our families of nested subgraphs exhibit invariance in degree distribution. If we put the distribution $P(k) = \mathcal{C}^{-1} k^{-\alpha}$, ($\mathcal{C} = \zeta(\alpha)$), equation (27) becomes

$$P_{S_K}(k) \approx \lambda_{S_K} \frac{\omega^k}{k!} \left. \frac{\mathrm{d}^k}{\mathrm{d}z^k} G_0^{\mathrm{SF}}(z) \right|_{z=1-\omega}. \tag{33}$$

Thus the problem lies in finding the *k*th derivative of $G_0^{\mathrm{SF}}(z)$. The computation is slightly more complex than the E-R graphs, and involves some approaches. First, we compute the generating function for a scale-free net $P(k) = \mathcal{C}^{-1} k^{-\alpha}$ whose exponent lies between 2 and 3, $G_0^{\mathrm{SF}}(z)$:

$$G_0^{\mathrm{SF}}(z) = \mathcal{C}^{-1} \mathbf{Li}_\alpha(z)$$

$$= \mathcal{C}^{-1} \frac{z}{\Gamma(\alpha)} \int_0^\infty \mathrm{d}t \, \frac{t^{\alpha-1}}{\mathrm{e}^t - z},$$

where $\mathbf{Li}_\alpha(z) = \sum_k^\infty \frac{z^k}{k^\alpha}$ is the polylogarithm function and, to obtain the last step, we used its integral form. But, actually, we are interested in the derivatives of $G_0^{\mathrm{SF}}(z)$. If we assume $z \to 1^-$ the kth derivative of $G_0^{\mathrm{SF}}(z)$ can be approached by

$$\frac{\mathrm{d}^k}{\mathrm{d}z^k} G_0^{\mathrm{SF}}(z) \approx \mathcal{C}^{-1} \frac{k!}{\Gamma(\alpha)} \int_0^\infty \mathrm{d}t \, \frac{t^{\alpha-1}}{(\mathrm{e}^t - z)^{k+1}}$$

$$\approx \mathcal{C}^{-1} \frac{k!}{\Gamma(\alpha)} \int_0^\infty \mathrm{d}t \, \frac{t^{\alpha-1}}{(t + \tau)^{k+1}}$$

$$= \mathcal{C}^{-1} \frac{k! \tau^{\alpha-1-k}}{\Gamma(\alpha)} \int_0^\infty \mathrm{d}y \, \frac{y^{\alpha-1}}{(y + 1)^{k+1}},$$

where, in the first approach, we used the fact that, if $z \to 1$, we are near a singularity when $t \to 0$. Thus, the dominant terms of the sum will be those close to $t = 0$. This enables us to rewrite $\mathrm{e}^t \approx 1 + t + \mathcal{O}(t^2)$. In the last step, we made the coordinate change $\tau = 1 - z$ and, then, $t = y\tau$. If we evaluate such an expression at $z = 1 - \omega$, with ω small enough:

$$\left. \frac{\mathrm{d}^k}{\mathrm{d}z^k} G_0^{\mathrm{SF}}(z) \right|_{z=1-\omega} \approx \mathcal{C}^{-1} \frac{k! \omega^{\alpha-1-k}}{\Gamma(\alpha)} J_{k+1,\alpha+1},$$

where $J_{k+1,\alpha+1}$ is defined as

$$J_{k+1,\alpha+1} \equiv \int_0^\infty \mathrm{d}y \, \frac{y^{\alpha-1}}{(y + 1)^{k+1}} = \frac{\Gamma(\alpha)\Gamma(k - \alpha + 1)}{k!(k - \alpha + 2)}.$$

If we check the behavior of $J_{k+1,\alpha+1}$ for large k's, we see that

$$J_{k+1,\alpha+1} \approx \frac{\Gamma(\alpha)}{k^\alpha}. \tag{34}$$

Thus, if we introduce the above results into the definition of P_{S_K}:

$$P_{S_K}(k) \approx \frac{\lambda_{S_K}}{\mu} \frac{\omega^k}{k!} \left. \frac{\mathrm{d}^k}{\mathrm{d}z^k} G_0^{\mathrm{SF}}(z) \right|_{z=1-\omega}$$

$$= \mathcal{C}^{-1} \frac{\lambda_{S_K}}{\mu} \omega^{\alpha-1} k^{-\alpha}, \tag{35}$$

which can be rewritten in the standard form when describing self-similar objects:

$$P_{S_K}(k) \approx \rho^{-\alpha} P(k) = P(\rho k), \tag{36}$$

where ρ is a constant that, interestingly, depends on both the scaling exponent α and the nature of the nesting, namely:

$$\rho = \left(\frac{\mu}{\lambda_{S_K} \omega^{(\alpha-1)}} \right)^{\frac{1}{\alpha}}. \tag{37}$$

Note that the invariance in the degree distribution is also valid for scale-free nets with $\alpha > 3$. However, in this case, the percolation inequality (19) is not generally satisfied.

7. Discussion

Many interacting systems found in nature display a scale-free topology, $P(k) \propto k^{-\alpha}$, with $2 < \alpha < 3$. In this paper, we have shown that the assumptions of the configuration model are enough to explain many of the scaling and self-similar properties of the observed nested subgraph nets. The resulting prediction (36) reveals that, under no correlations, we should expect invariance in degree distributions of nested subgraphs to occur. This is what we observe
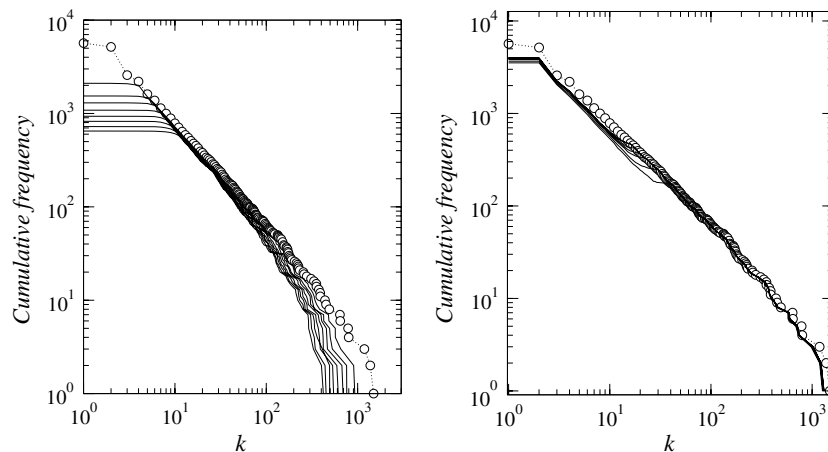
Figure 4. Analyzing the web obtained from the O Wilde's novel *The portrait of Dorian Gray*. The network was built up by tracing an arc between two adjacent words, if they appear one after the other within the same sentence. The obtained graph has $N = 5696$ nodes and displays a scale-free distribution $P(k) \propto k^{-\alpha}$ (gray circles), with an exponential cut-off at high connectivities ($k > 1000$). In this graph, $\alpha \approx 2.15$ and $\langle k \rangle \approx 8.814$. We plot the cumulative frequency for the K-cores, $4 \geqslant K \geqslant 11$ (left). Despite the strong connectivity requirements imposed for the K-core, the distribution behaves as a statistical invariant. The same is observed with successive *naked* K-scaffold subgraphs, $K = 14, 16, 18, 20, 22, 30, 40$ (right). The naked K-scaffold subgraph is obtained from the K-scaffold but deleting all the nodes with $k < K$ that are connected only to one node with $k' \geqslant K$.

in the analysis of real nets (see figure 4). Indeed, in the analysis of the degree frequency we see that, despite the finite size of our system, the degree frequency acts as an invariant, only modulated by a scaling factor. These results contrast with the previous work on sampled subnets obtained from scale-free graphs [23]. Although it is true that arbitrary subsets of nodes might not display invariance, our families of nested subgraphs are defined in such a way that our results are expected to hold. Further work should address the impact of the self-similarity in the functional aspects of the net as well as a broader study of nested subgraphs involving different types of real networks.

Acknowledgments

References

[1] Newman M E J 2004 *Eur. Phys. J.* B **38** 321
[2] Palla G, Derényi I, Farkas I and Vicsek T 2005 *Nature* **435** 814
[3] Guimerà R and Amaral L 2005 *J. Stat. Mech.* P02001
[4] Dorogovtsev S N, Goltsev A V and Mendes J F F 2006 *Phys. Rev. Lett.* **96** 040601
[5] Milo R, Shen-Orr S, Itzkovitz S, Kashtan N, Chklovskii D and Alon U 2002 *Science* **298** 824–7
[6] Corominas-Murtra B, Valverde S, Rodríguez-Caso C and Solé R V 2007 *Europhys. Lett.* **7** 18004

 [7] Bascompte J *et al* 2003 *Proc. Natl Acad. Sci. USA* **100** 9383
 [8] Goltsev A V, Dorogovtsev S N and Mendes J F F 2006 *Phys. Rev.* E **73** 056101 1–10
 [9] Fernholz D and Ramachandran V 2004 *Technical Report TR04-13* University of Texas at Austin. G/A
[10] Rodriguez-Caso C, Medina M A and Solé R V 2005 *FEBS J.* **272** 6423
[11] Song C, Havlin S and Makse H A 2005 *Nature* **433** 392
[12] Goh K-I, Salvi G, Kahng B and Kim D 2006 *Phys. Rev. Lett.* **96** 018701
[13] Alvarez-Hamelin J I, Dall'Asta L, Barrat A and Vespignani A 2005 *Preprint* arXiv.org:cs/0511007
[14] Guimerà R, Danon L, Díaz-Guilera A, Giralt F and Arenas A 2003 *Phys. Rev.* E **68** 065103
[15] Bollobás B 2001 *Random Graphs* 2nd edn (Cambridge: Cambridge University Press)
[16] Newman M E J, Strogatz S H and Watts D J 2001 *Phys. Rev.* E **64** 026118
[17] Gouyet J F 1995 *Physics of Fractal Structures* (Berlin: Springer)
[18] Dorogovtsev S, Goltsev A V and Mendes J F F 2007 *Preprint* arXiv:0705.0010
[19] Bekessy A, Bekessy P and Komlos J 1972 *Stud. Sci. Math. Hungar.* **7** 343
       Bender E A and Canfield E R 1978 *J. Comb. Theory* A **24** 296
       Bollobás B 1980 *Eur. J. Comb.* **1** 311
       Wormald N C 1981 *J. Comb. Theory* B **31** 156–68
[20] Pittel B, Spencer J and Wormald N 1996 *J. Comb. Theory* B **67** 111
[21] Callaway D S, Newman M E J, Strogatz S E and Watts D J 2000 *Phys. Rev. Lett.* **85** 5468
[22] Molloy M and Reed B 1995 *Random Struct. Algorithms* **6** 161–80
[23] Stumpf P H, Wiuf C and May R 2005 *Proc. Natl Acad. Sci.* **102** 4221