Science Advances
AAAS

advances.sciencemag.org/cgi/content/full/4/12/eaau4212/DC1

# Supplementary Materials for

## Effective approach to epidemic containment using link equations in complex networks

Joan T. Matamalas*, Alex Arenas*, Sergio Gómez*

*Corresponding author. Email: joantomas.matamalas@urv.cat (J.T.M.); alexandre.arenas@urv.cat (A.A.); sergio.gomez@urv.cat (S.G.)

Published 5 December 2018, *Sci. Adv.* **4**, eaau4212 (2018)
DOI: 10.1126/sciadv.aau4212

**The PDF file includes:**

Fig. S11. Comparison of the number of connected components after total containment between the link epidemic importance and eigenscore strategies, calculated for the real networks and parameters as in Fig. 5.
Fig. S12. Graphical representation of the determination of the epidemic threshold.
Fig. S13. Computational time invested for each method to perform a single ranking and removal for BA networks ranging from 100 to 400,000 nodes, averaged over 36 repetitions.
Table S1. Structural characteristics of the 27 real networks obtained from the Network Repository (http://networkrepository.com) and used in Fig. 6 and fig. S11.

**Other Supplementary Material for this manuscript includes the following:**

(available at advances.sciencemag.org/cgi/content/full/4/12/eaau4212/DC1)

Data file S1 (.graphml format). Air transportation network data.

## Section S1. Link epidemic importance and connected components

One important property of our proposed definition of *link epidemic importance* is that it tends to maintain the connectivity of the network when the selected link is removed. Here we show that, when the network is formed by two subnetworks $A$ and $B$, which are connected by just one link $(i_A, j_B)$, the link epidemic importance of that link is lower than that of another link internal to $A$ or $B$. This means our containment strategy of removing the edge with largest link epidemic importance will not break the network in two disconnected components $A$ and $B$, unlike the betweenness approach, for which the link $(i_A, j_B)$ plays the role of a bridge and thus it has maximal edge betweenness.

Let $\langle k\rangle_A$ and $\langle k\rangle_B$ be the average degrees of subnetworks $A$ and $B$, respectively, with $\langle k\rangle_A > \langle k\rangle_B$. Let us also call $\rho^A$ and $\rho^B$ their respective incidence of the epidemics. The link epidemic importance of a link between nodes $i$ and $j$ has been defined as

$$I_{ij} = \bar{n}_{ij} + \bar{n}_{ji}\,, \tag{S.1}$$

where

$$\bar{n}_{ij} = \beta P(\sigma_j = S, \sigma_i = I) \sum_{r=1}^{N} A_{jr} \beta P(\sigma_r = S | \sigma_j = I)\,. \tag{S.2}$$

Supposing independence of the states of the nodes, $\bar{n}_{ij}$ can be approximated as

$$\bar{n}_{ij} \approx \beta P(\sigma_j = S) P(\sigma_i = I) \sum_{r=1}^{N} A_{jr} \beta P(\sigma_r = S)\,. \tag{S.3}$$

In a homogeneous mean field approximation we may substitute $k_j = \langle k\rangle$, $P(\sigma_i = I) \approx \rho$, and $P(\sigma_j = S) \approx 1 - \rho$, which lead to the following expressions for the importance of a link:

$$I^A \approx 2\beta^2 \rho^A (1-\rho^A)^2 \langle k\rangle_A\,, \tag{S.4}$$

$$I^B \approx 2\beta^2 \rho^B (1-\rho^B)^2 \langle k\rangle_B\,, \tag{S.5}$$

$$I^{AB} \approx \beta^2 \left[\rho^A (1-\rho^B)^2 \langle k\rangle_B + \rho^B (1-\rho^A)^2 \langle k\rangle_A\right]\,. \tag{S.6}$$

Here, $I^A$ and $I^B$ denote the link epidemic importance of links inside $A$ and $B$, respectively, and $I^{AB}$ the link epidemic importance of the link connecting subnetworks $A$ and $B$.

We need an expression relating the average degree $\langle k \rangle$ and the incidence of the epidemics $\rho$ to be able to calculate the approximate values of the link epidemic importances in Eqs. (S.4) to (S.6). It can be obtained using the nonperturbative heterogeneous mean field (npHMF) equations in [Gómez et al., Phys. Rev. E, 84 (2011) 036105]. In particular, the npHMF equations for the SIS model without one-step reinfections (WOR) read as

$$0 = -\mu\rho_k + (1-\rho_k)(1-q_k)\,, \tag{S.7}$$

$$q_k = \prod_{k'}(1-\beta\rho_{k'})^{C_{kk'}}\,, \tag{S.8}$$

where $\rho_k$ represents the fraction of infected nodes of degree $k$, $q_k$ the probability that nodes of degree $k$ are not infected by nodes of any other degree $k'$, and $C_{kk'} = kP(k'|k)$ the expected number of links from a node of degree $k$ to nodes of degree $k'$. In the nonperturbative homogeneous mean field (npHoMF) approximation, this reduces to

$$0 = -\mu\rho + (1-\rho)(1-q)\,, \tag{S.9}$$

$$q = (1-\beta\rho)^{\langle k \rangle}\,. \tag{S.10}$$

Thus, after some algebra we get

$$\langle k \rangle = \frac{\log\left(1-\mu\dfrac{\rho}{1-\rho}\right)}{\log(1-\beta\rho)}\,. \tag{S.11}$$

An immediate consequence of Eq. (S.11) is that $0 \leqslant \rho \leqslant 1/(1+\mu)$. We can see a plot of the npHoMF relationship between $\rho$ and $\langle k \rangle$ in the inset of Fig. S1. Note that $\rho$ is an increasing function of $\langle k \rangle$, thus the larger the average degree, the greater the incidence of the epidemics.

Now, we can substitute Eq. (S.11) for $\langle k \rangle_A$ and $\langle k \rangle_B$ into Eqs. (S.4) to (S.6) to obtain approximations of the three different link epidemic importances. The results are presented in Fig. S1. Fixing a certain average degree $\langle k \rangle_B$ for subnetwork $B$, we consider subnetworks $A$ with $\langle k \rangle_A \geqslant \langle k \rangle_B$. Since $\rho$ increases with $\langle k \rangle$, this is equivalent to fixing $\rho^B$ and consider subnetworks $A$ with $\rho^A \geqslant \rho^B$. We observe that, in all cases, the link epidemic importance $I^A$ of links in subnetwork $A$ is larger than $I^{AB}$ of the bridge link between subnetworks $A$ and $B$ ($I^A \geqslant I^{AB}$), thus confirming that our epidemic containment strategy driven by link epidemic importance does not disconnect the network. This result has been obtained under independence and homogeneous mean field approximations, and for a specific structural configuration of the network. In practice, we observe that the epidemic containment approach based on link epidemic importance is the one considered which better preserves the connectivity of the networks, as shown in Figs. S10 and S11.

## Section S2. Linearization of the ELE model

The determination of the epidemic threshold from the Epidemic Link Equations (ELE) requires the consideration of states in which the probabilities of having infected nodes are very small, i.e. $\Phi_{ij}, \Phi_{ji}, \Theta_{ij}^I \ll 1$. Therefore, we may suppose that $\Phi_{ij}, \Phi_{ji}, \Theta_{ij}^I \sim O(\varepsilon)$, with $\varepsilon \ll 1$, and in consequence $\Theta_{ij}^S \sim 1 - O(\varepsilon)$. Using these approximations the epidemic link equations become linear in these $O(\varepsilon)$ probabilities, since $O(\varepsilon^2)$ terms should be neglected.

We start with the linearization of the hostility:

$$\begin{aligned} h_{ij} &= \frac{\Phi_{ij}}{\Phi_{ij} + \Theta_{ij}^S} \\ &= \frac{\Phi_{ij}}{1 - (\Phi_{ji} + \Theta_{ij}^I)} \\ &= \Phi_{ij} \left(1 + \Phi_{ji} + \Theta_{ij}^I + O(\varepsilon^2)\right) \\ &= \Phi_{ij} + O(\varepsilon^2)\,, \end{aligned} \tag{S.12}$$

where we have used the normalization $\Phi_{ij} + \Phi_{ji} + \Theta_{ij}^I + \Theta_{ij}^S = 1$, and we realize that terms $\Phi_{ij}\Phi_{ji}$ and $\Phi_{ij}\Theta_{ij}^I$ are both $O(\varepsilon^2)$. Substituting hostility in the expression for $q_{ij}$ we get:

$$\begin{aligned} q_{ij} &= \prod_{\substack{r=1 \\ r\neq j}}^{N} (1 - \beta A_{ri} h_{ir}) \\ &= \prod_{\substack{r=1 \\ r\neq j}}^{N} \left(1 - \beta A_{ri}\Phi_{ir} + O(\varepsilon^2)\right) \\ &= 1 - \beta \sum_{\substack{r=1 \\ r\neq j}}^{N} A_{ri}\Phi_{ir} + O(\varepsilon^2) \\ &= 1 - \beta \sum_{r=1}^{N} A_{ri}(1 - \delta_{rj})\Phi_{ir} + O(\varepsilon^2)\,, \end{aligned} \tag{S.13}$$

where the Kronecker $\delta_{rj}$ has been introduced to make zero the $j$th term of the sum. Now we are in condition to find the linear approximations of the main ELE model equations. First, the equation for $\Theta_{ij}^I$ becomes:

$$\begin{aligned} \Theta_{ij}^I &= \Theta_{ij}^S\,(1 - q_{ij})\,(1 - q_{ji}) \\ &\quad + \Phi_{ij}\,(1 - (1 - \beta)q_{ij})\,(1 - \mu) \\ &\quad + \Phi_{ji}\,(1 - \mu)\,(1 - (1 - \beta)q_{ji}) \\ &\quad + \Theta_{ij}^I\,(1 - \mu)^2 \\ &= O(\varepsilon^2) \\ &\quad + \Phi_{ij}\,\beta(1 - \mu) + O(\varepsilon^2) \\ &\quad + \Phi_{ji}\,(1 - \mu)\beta + O(\varepsilon^2) \\ &\quad + \Theta_{ij}^I\,(1 - \mu)^2 \\ &= \beta(1 - \mu)\,\Phi_{ij} + \beta(1 - \mu)\,\Phi_{ji} + (1 - \mu)^2\,\Theta_{ij}^I + O(\varepsilon^2)\,. \end{aligned} \tag{S.14}$$

Note that, for the terms with a factor $\Phi_{ij}$, $\Phi_{ji}$ or $\Theta_{ij}^I$, which are $O(\varepsilon)$, we just need to keep $O(1)$ contributions in the rest of the term, thus we may use the approximations $q_{ij} = q_{ji} = 1 + O(\varepsilon)$.

The equation for $\Phi_{ij}$ reads:

$$\begin{aligned}
\Phi_{ij} &= \Theta^S_{ij}\, q_{ij}(1-q_{ji}) \\
&\quad + \Phi_{ij}\,((1-\beta)q_{ij})\,(1-\mu) \\
&\quad + \Phi_{ji}\,\mu\,(1-(1-\beta)q_{ji}) \\
&\quad + \Theta^I_{ij}\,\mu(1-\mu) \\
&= \beta\sum_{r=1}^{N} A_{rj}(1-\delta_{ri})\,\Phi_{jr} + O(\varepsilon^2) \\
&\quad + \Phi_{ij}\,(1-\beta)(1-\mu) + O(\varepsilon^2) \\
&\quad + \Phi_{ji}\,\mu\beta + O(\varepsilon^2) \\
&\quad + \Theta^I_{ij}\,\mu(1-\mu) \\
&= \beta\sum_{r=1}^{N} \left(A_{rj}(1-\delta_{ri}) + \mu\delta_{ri}\right)\,\Phi_{jr} \\
&\quad + \Phi_{ij}\,(1-\beta)(1-\mu) + \Theta^I_{ij}\,\mu(1-\mu) + O(\varepsilon^2) \\
&= \beta\sum_{r=1}^{N} (A_{rj} - (1-\mu)\delta_{ri})\,\Phi_{jr} + (1-\beta)(1-\mu)\,\Phi_{ij} + \mu(1-\mu)\,\Theta^I_{ij} + O(\varepsilon^2)\,. \qquad \text{(S.15)}
\end{aligned}$$

For the last step we have made use of $A_{rj}\delta_{ri} = A_{ij}\delta_{ri} = \delta_{ri}$, since these equations correspond to a link between nodes $i$ and $j$, thus we are implicitly assuming that $A_{ij} = 1$.

Summarizing, the linearized equations of the ELE model can be expressed as:

$$\Theta^I_{ij} = \beta(1-\mu)\,\Phi_{ij} + \beta(1-\mu)\,\Phi_{ji} + (1-\mu)^2\,\Theta^I_{ij}\,, \qquad \text{(S.16)}$$

$$\Phi_{ij} = \beta\sum_{r=1}^{N}(A_{rj} - (1-\mu)\delta_{ri})\,\Phi_{jr} + (1-\beta)(1-\mu)\,\Phi_{ij} + \mu(1-\mu)\,\Theta^I_{ij}\,. \qquad \text{(S.17)}$$

## Section S3. Epidemic threshold

We have shown in *Methods* that the epidemic threshold is obtained by finding the non-trivial solutions of the system of equations

$$\frac{\mu}{\beta}\epsilon_i = \sum_j B_{ji}\epsilon_j \,, \tag{S.18}$$

where the components of matrix $B$ read

$$B_{ij} = (1-\Upsilon)A_{ij} - \Upsilon k_i \delta_{ij} \,, \tag{S.19}$$

with the constant $\Upsilon$ being

$$\Upsilon = \frac{\beta(1-\mu)}{\mu(2-\mu)+2\beta(1-\mu)} \,. \tag{S.20}$$

The non-trivial solutions of Eq. (S.18) require $\frac{\mu}{\beta}$ to be an eigenvalue $\Lambda(B)$ of matrix $B$:

$$\Lambda(B) = \frac{\mu}{\beta} \tag{S.21}$$

Unfortunately, for any given fixed value of the recovery rate $\mu$, matrix $B$ also depends on the infection rate $\beta$ through $\Upsilon$, thus Eq. (S.21) becomes an implicit equation for $\beta$. Moreover, since we are interested in the onset of the epidemics, we must find the lowest value of the infection probability, $\beta_c$, which satisfies Eq. (S.21). At first sight, one would say that we must choose the maximum eigenvalue of $B$ to obtain the lowest value of $\beta$. However, the dependence of $B$ on $\beta$ may rise the question whether a different eigenvalue (e.g., the second largest eigenvalue) could solve the equation at a lower value of $\beta$, since each eigenvalue has a different functional form.

We show in Fig. S12 that, no matter the functional form of the eigenvalues of $B$ on $\beta$, if there exists a solution to Eq. (S.21), the curve $\frac{\mu}{\beta}$ always crosses the largest eigenvalue line most-to-the-left than for any other of the eigenvalues, due to the decreasing behavior of $\frac{\mu}{\beta}$. Thus, we can safely say express that

$$\beta_c = \frac{\mu}{\Lambda_{\max}(B)} \,, \tag{S.22}$$

which is the final implicit equation for the epidemic threshold $\beta_c$. In Fig. S12 we have made use of an Erdős-Rényi network with 100 nodes, average degree $\langle k \rangle = 6$, and setting $\mu = 0.5$, but the previous result is general no matter the network or the parameters.

## Section S4. Data description

Description of the 27 real networks used in Figs. 6 and S11, sorted by increasing number of nodes. They have been obtained from the Network Repository (`http://networkrepository.com`). Table S1 provides their main structural characteristics.

**ia-infect-dublin** Human contact network where nodes represent humans and edges represent proximity (i.e., contacts in the physical world), during the Infectious SocioPatterns event that took place at the Science Gallery in Dublin, Ireland.

**soc-wiki-Vote** Wikipedia voting data from the inception of Wikipedia till January 2008. Nodes represent Wikipedia users and a directed edge from node $i$ to node $j$ represents that user $i$ voted on user $j$.

**ca-CSphd** Genealogy network of PhD's in computer science.

**ia-fb-messages** The Facebook-like Social Network originate from an online community for students at University of California, Irvine. The dataset includes the users that sent or received at least one message.

**soc-hamsterster** Network of the friendship and family links between users of Hamsterster social network.

**socfb-USFCA72** A social friendship network extracted from Facebook consisting of people (nodes) with edges representing friendship ties.

**socfb-nips-ego** A social friendship network extracted from Facebook consisting of people (nodes) with edges representing friendship ties.

**socfb-Santa74** A social friendship network extracted from Facebook consisting of people (nodes) with edges representing friendship ties.

**ca-GrQc** Collaboration network of arXiv General Relativity. Nodes represent scientists, and links coauthorship.

**web-spam** Web Spam Challenge 2008 network.

**power-US-Grid** US Power grid graph.

**ca-Erdos992** Erdős collaboration network. Nodes represent scientists, and links coauthorship.

**soc-advogato** Advogato is a social community platform where users can explicitly express weighted trust relationships among themselves. The dataset contains a list of all of the user-to-user links.

**p2p-Gnutella08** Gnutella peer to peer network from August 8 2002.

**ia-reality** Reality mining network data consists of human mobile phone call events between a small set of core users at the MIT whom actually were assigned mobile phones for which all calls were collected. A node represents a person; an edge indicates a phone call or voicemail between two users.

**ca-HepTh** Collaboration network of arXiv High Energy Physics Theory. Nodes represent scientists, and links coauthorship.

**soc-anybeat** Anybeat is an online community, a public gathering place where you can interact with people from around your neighborhood or across the world.

**ca-AstroPh** Collaboration network of arXiv Astrophysics. Nodes represent scientists, and links coauthorship.

**ca-CondMat** Collaboration network of arXiv Condensed Matter. Nodes represent scientists, and links coauthorship.

**soc-gplus** Google+ social network.

**tech-as-caida2007** Internet network at the level of autonomous systems as of 2017. Nodes represent autonomous systems, and there exists a link between them if they have a business agreement for the routing of packets.

**ia-email-EU** The network was generated using email data from a large European research institution. Nodes are users and edges represent email exchanges between two users in both directions.

**ia-enron-large** Network of emails exchanged between senior managers of Enron Corporation, during the period which lead to its bankruptcy.

**soc-brightkite** Brightkite is a location-based social networking service provider where users shared their locations by checking-in. The dataset contains all links among users.

**soc-epinions** Who-trust-whom online social network of the general consumer review site Epinions.com.

**soc-slashdot** A technology-related news website known for its specific user community. The dataset contains friend/foe tags between the users of slashdot.

**soc-twitter-follows** Twitter follower network.

The prefix in the name of each network indicates the cathegory it belongs to, namely: (ia) interaction networks; (soc) social networks; (ca) collaboration networks; (socfb) Facebook networks; (web) we graphs; (power) power networks; (p2p) peer to peer networks; (tech) technological networks.

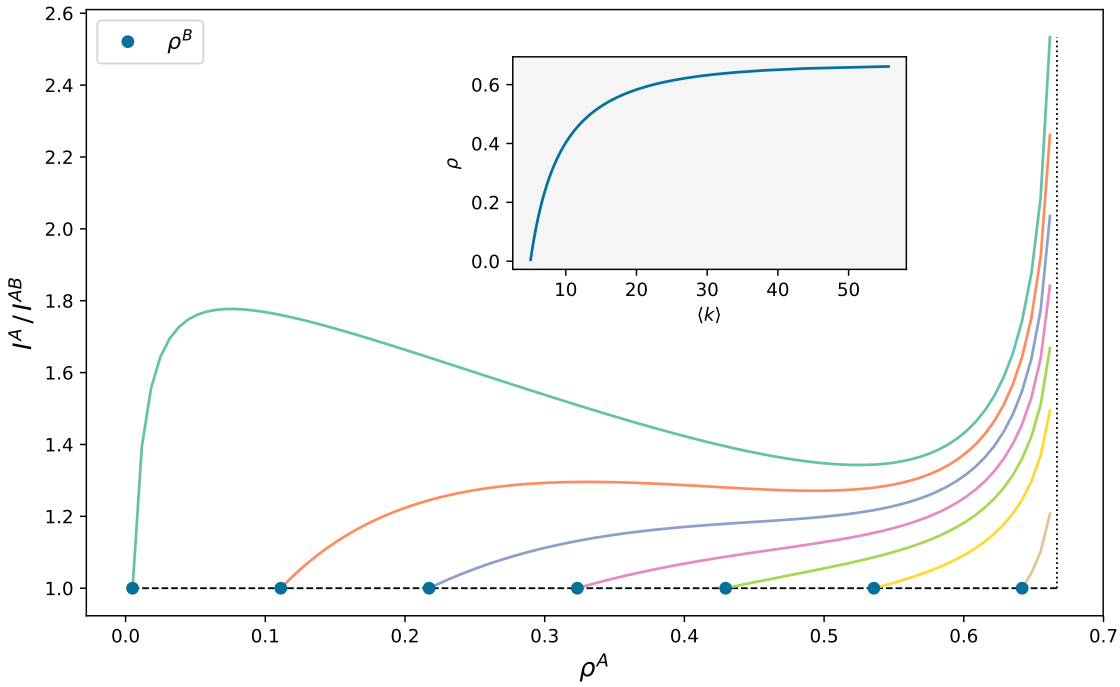


**Fig. S1.** Ratio between the link epidemic importance $I^A$ of a link in a subnetwork $A$ and the link epidemic importance $I^{AB}$ of a link that acts as the only bridge between subnetworks $A$ and $B$. First, we fix the average degree $\langle k \rangle_B$ of subnetwork $B$ (or equivalently, we fix its incidence $\rho^B$, the red circles), and then we consider subnetworks $A$ with average degree (and epidemic incidence) larger than that of $B$, i.e. $\langle k \rangle_A \geqslant \langle k \rangle_B$ (thus, $\rho^A \geqslant \rho^B$). We can see that, in all cases, $I^A \geqslant I^{AB}$, meaning that the ranking by link epidemic importance will not be leaded by the bridges. The vertical dotted line highlights the asymptote at $\rho = 1/(1+\mu)$. The inset shows the relationship between the incidence and the average degree. We have set the epidemic parameters to $\mu = 0.5$ and $\beta = 0.1$, and the calculations rely on a nonperturbative homogeneous mean field approximation (npHoMF). See section S1 for further details.

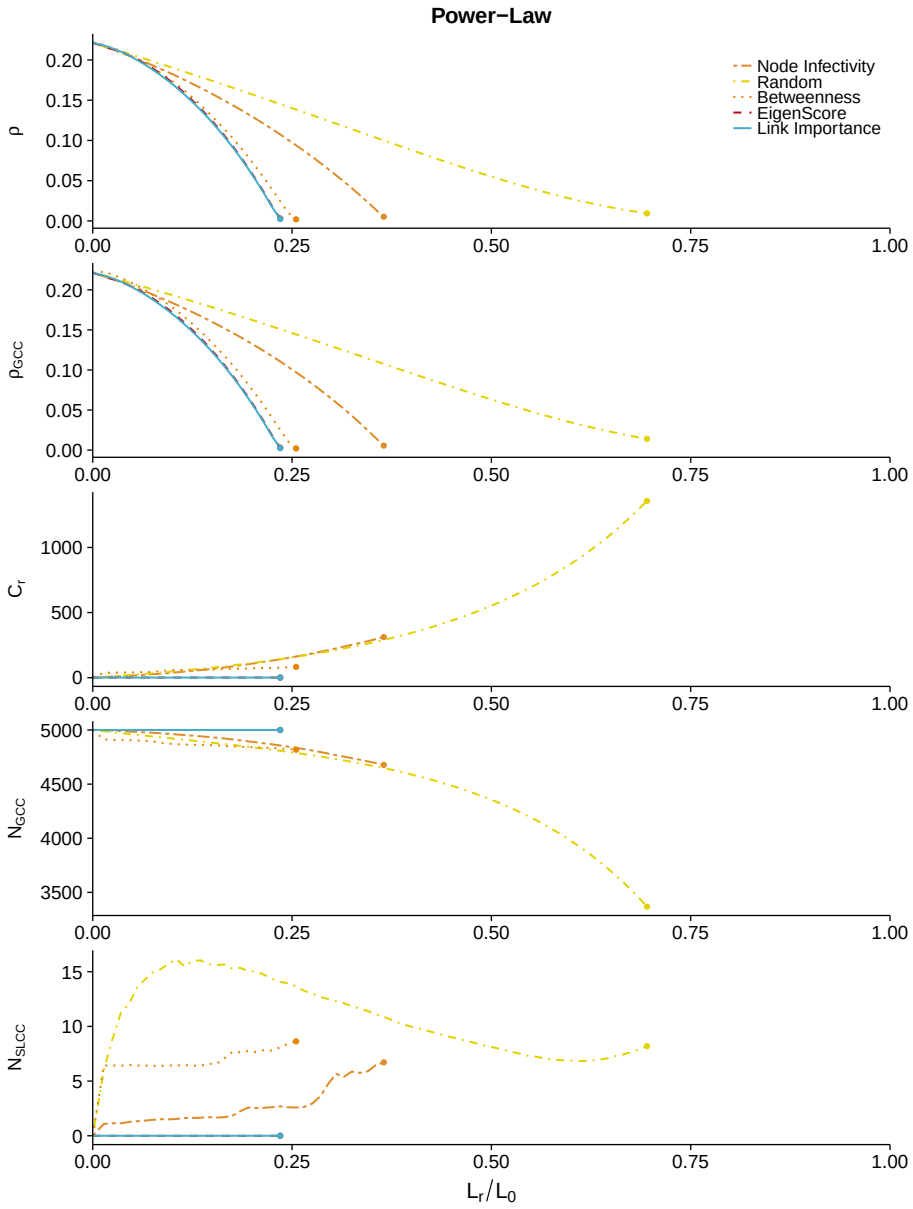


**Fig. S2**. Epidemic containment for a network with 5000 nodes, power-law degree distribution of exponent 3, and average degree $\langle k \rangle = 6$. Five containment strategies are compared: maximum probability of being infected (Node Infectivity), random link selection (Random), maximum edge betweenness (Betweenness), maximum eigenscore (EigenScore), and maximum link epidemic importance (Link Importance). In the horizontal axis we represent the fraction of removed links during the containment process ($L_r/L$). We show, from top to bottom: $\rho$, incidence of the epidemics on the network; $\rho_{\mathrm{GCC}}$, incidence of the epidemics on the Giant Connected Component; $C_r$, number of connected components; $N_{\mathrm{GCC}}$, size of the GCC; $N_{\mathrm{SLCC}}$, size of the Second Largest Connected Component. The dots mark the achievement of total containment.

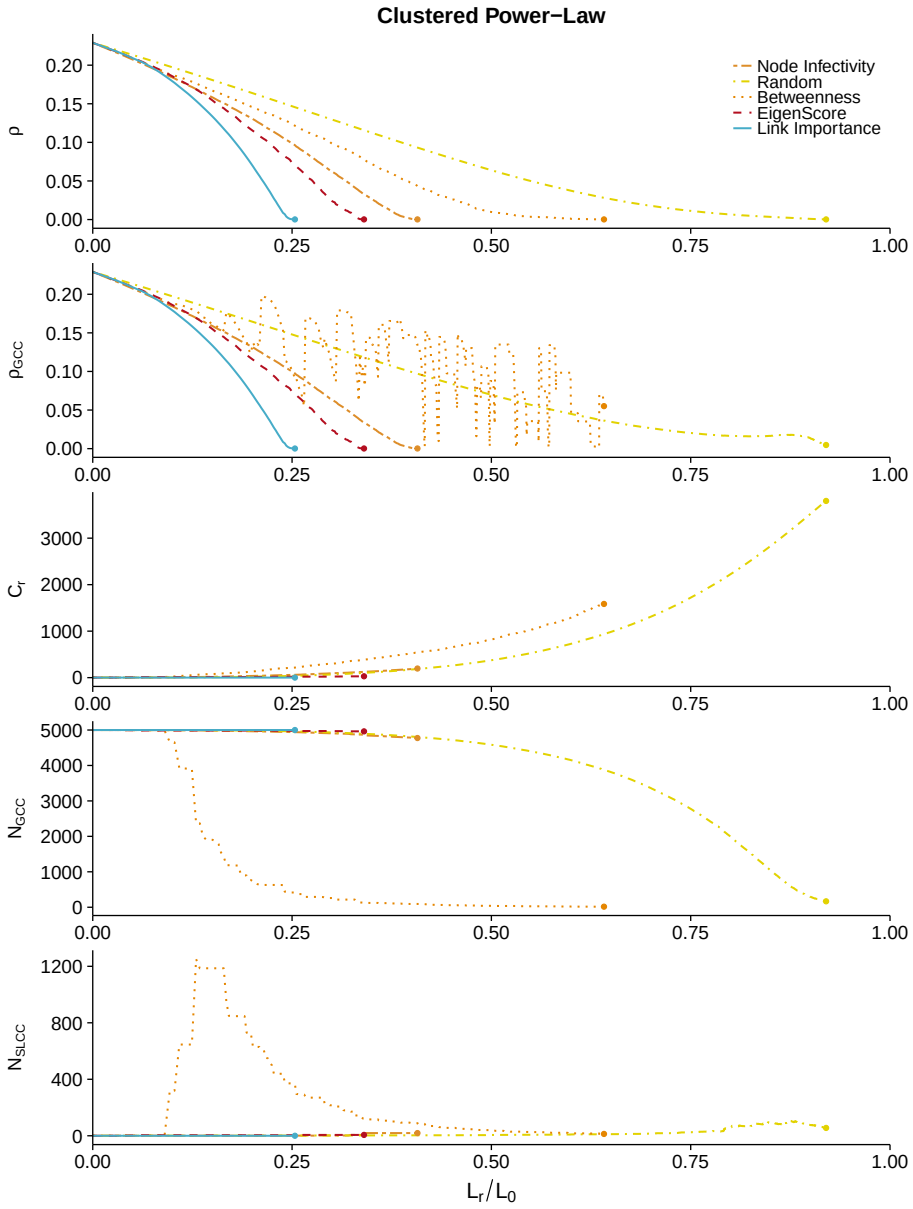


**Fig. S3.** Epidemic containment for a network with 5000 nodes, power-law degree distribution of exponent 3, high clustering coefficient, and average degree $\langle k \rangle = 6$ (see Methods for more details). Five containment strategies are compared: maximum probability of being infected (Node Infectivity), random link selection (Random), maximum edge betweenness (Betweenness), maximum eigenscore (EigenScore), and maximum link epidemic importance (Link Importance). In the horizontal axis we represent the fraction of removed links during the containment process ($L_r/L$). We show, from top to bottom: $\rho$, incidence of the epidemics on the network; $\rho_{\mathrm{GCC}}$, incidence of the epidemics on the Giant Connected Component; $C_r$, number of connected components; $N_{\mathrm{GCC}}$, size of the GCC; $N_{\mathrm{SLCC}}$, size of the Second Largest Connected Component. The dots mark the achievement of total containment.

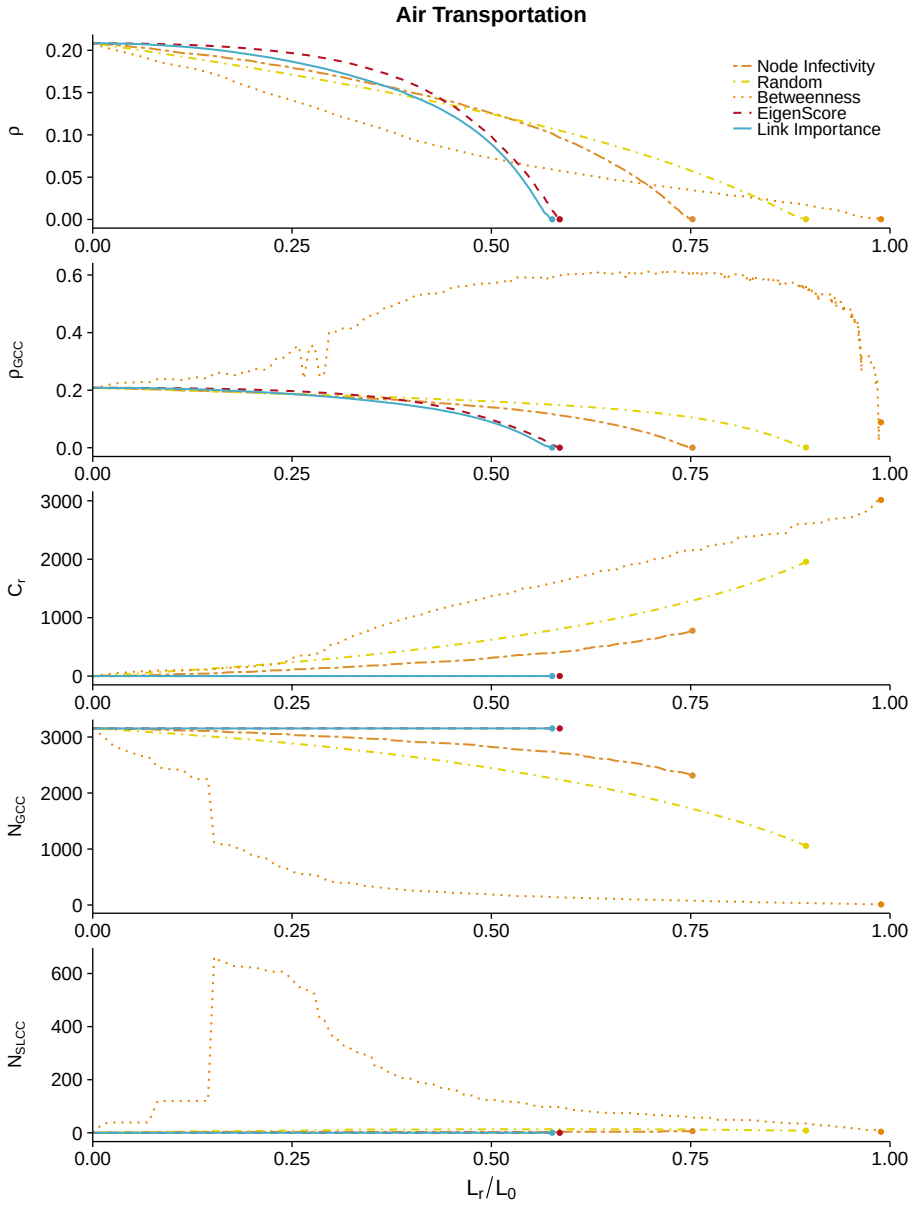


**Fig. 4**. Epidemic containment for the air transportation network (see Methods for more details). Five containment strategies are compared: maximum probability of being infected (Node Infectivity), random link selection (Random), maximum edge betweenness (Betweenness), maximum eigenscore (EigenScore), and maximum link epidemic importance (Link Importance). In the horizontal axis we represent the fraction of removed links during the containment process ($L_r/L$). We show, from top to bottom: $\rho$, incidence of the epidemics on the network; $\rho_{\mathrm{GCC}}$, incidence of the epidemics on the Giant Connected Component; $C_r$, number of connected components; $N_{\mathrm{GCC}}$, size of the GCC; $N_{\mathrm{SLCC}}$, size of the Second Largest Connected Component. The dots mark the achievement of total containment.

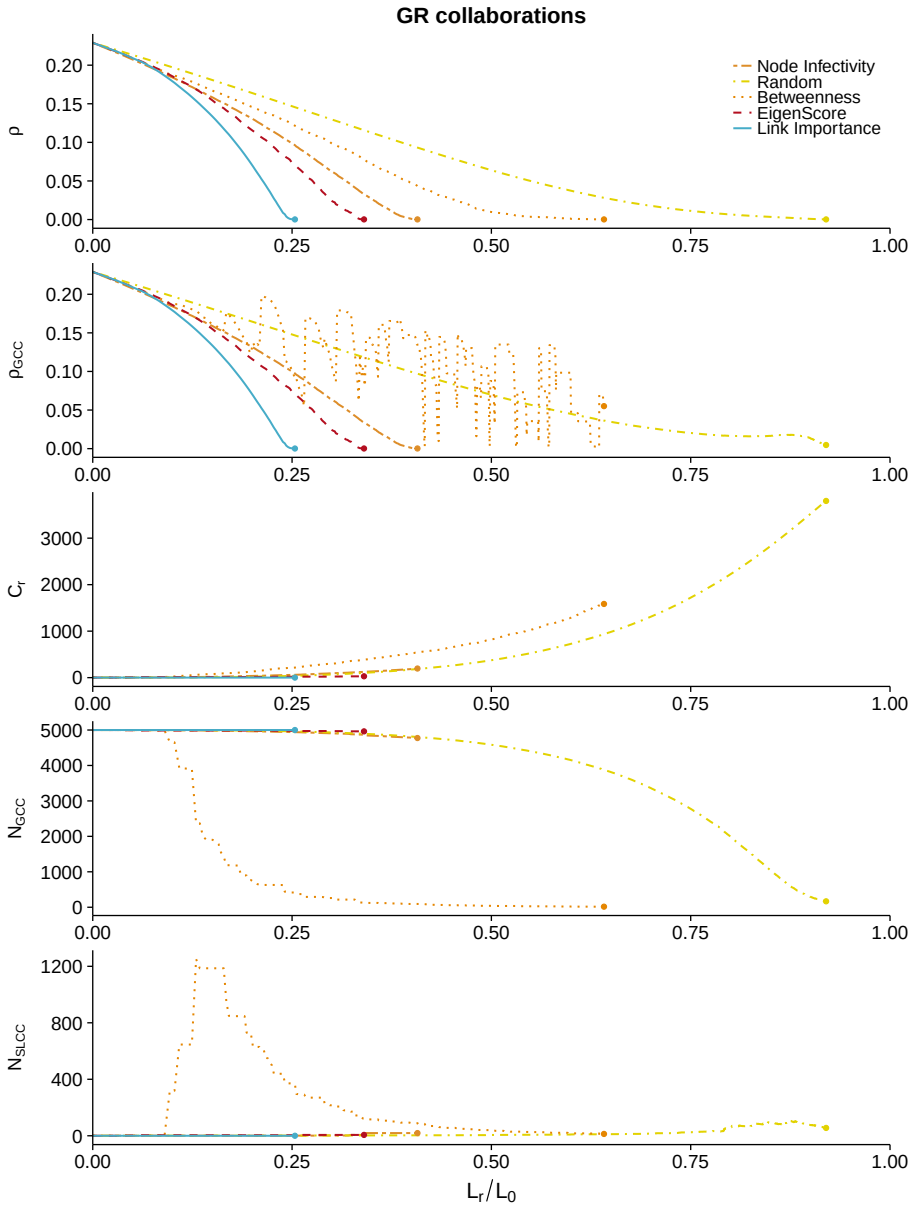


**Fig. 5**. Epidemic containment for the general relativity collaborations network (see Methods for more details). Five containment strategies are compared: maximum probability of being infected (Node Infectivity), random link selection (Random), maximum edge betweenness (Betweenness), maximum eigenscore (EigenScore), and maximum link epidemic importance (Link Importance). In the horizontal axis we represent the fraction of removed links during the containment process ($L_r/L$). We show, from top to bottom: $\rho$, incidence of the epidemics on the network; $\rho_{\mathrm{GCC}}$, incidence of the epidemics on the Giant Connected Component; $C_r$, number of connected components; $N_{\mathrm{GCC}}$, size of the GCC; $N_{\mathrm{SLCC}}$, size of the Second Largest Connected Component. The dots mark the achievement of total containment.

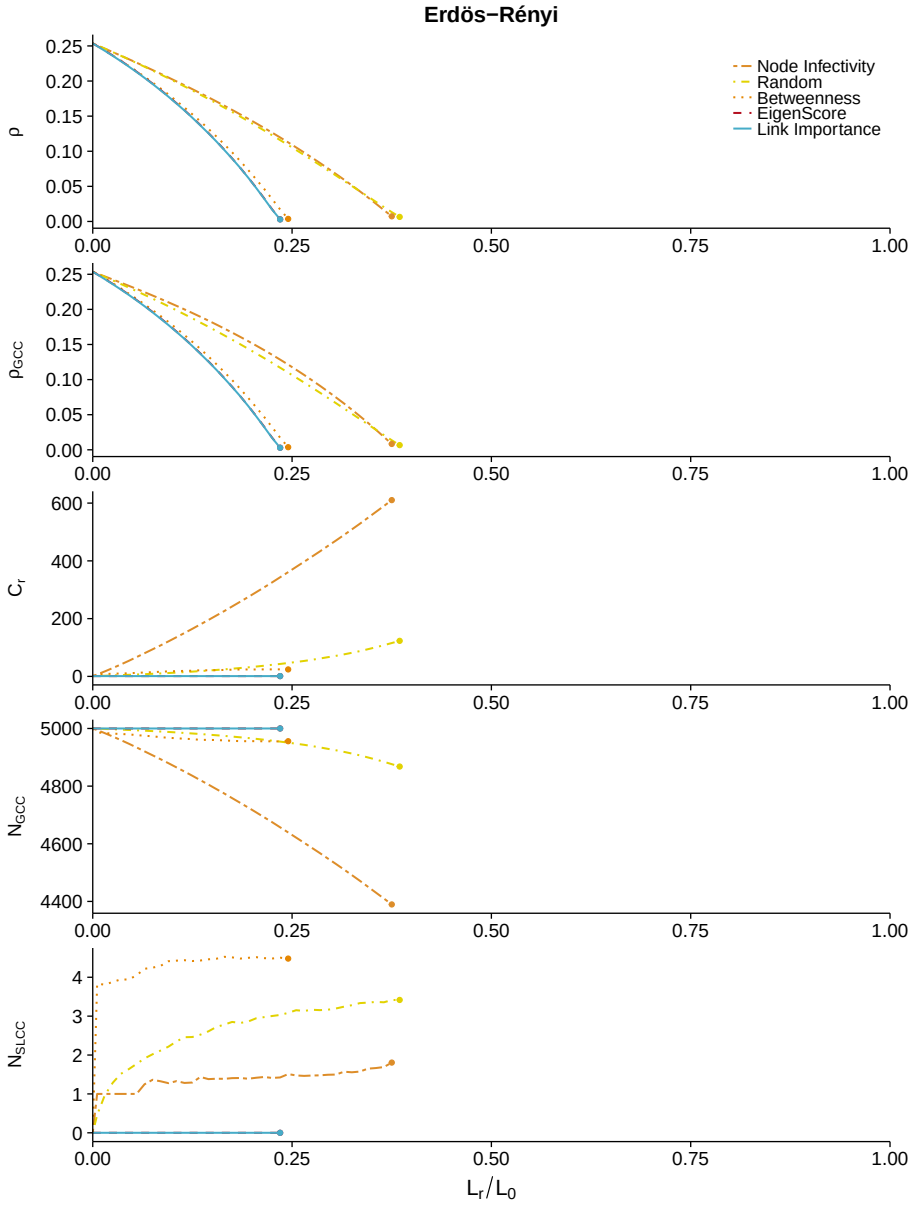


**Fig. S6**. Epidemic containment for an ER network with 5000 nodes and average degree $\langle k \rangle = 6$. Five containment strategies are compared: maximum probability of being infected (Node Infectivity), random link selection (Random), maximum edge betweenness (Betweenness), maximum eigenscore (EigenScore), and maximum link epidemic importance (Link Importance). In the horizontal axis we represent the fraction of removed links during the containment process ($L_r/L$). We show, from top to bottom: $\rho$, incidence of the epidemics on the network; $\rho_{\mathrm{GCC}}$, incidence of the epidemics on the Giant Connected Component; $C_r$, number of connected components; $N_{\mathrm{GCC}}$, size of the GCC; $N_{\mathrm{SLCC}}$, size of the Second Largest Connected Component. The dots mark the achievement of total containment. We observe that the containment is similar to the one in the power-law network in Fig. S2 due to the absence of transitivity and modular structure.

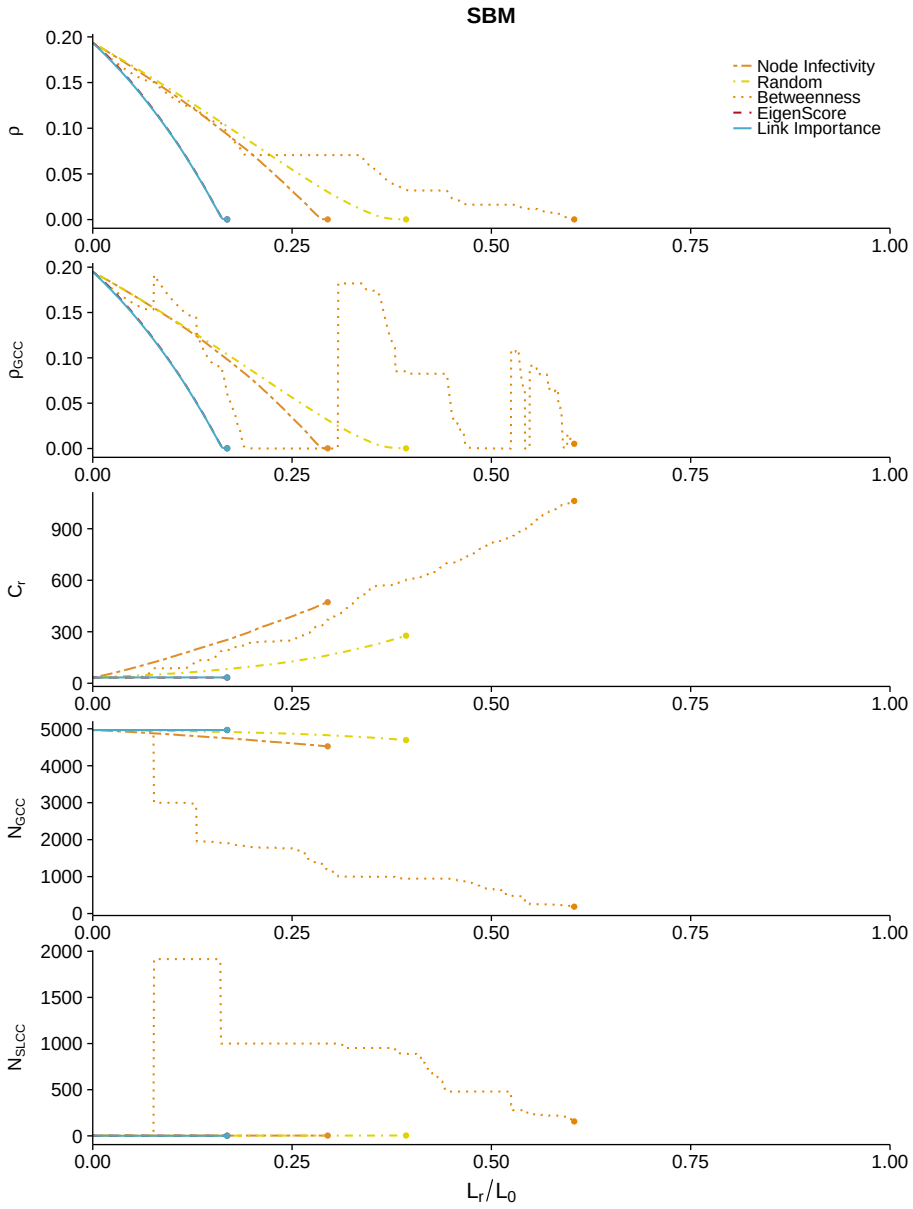


**Fig. S7**. Epidemic containment for a network with 5000 nodes generated with a stochastic block model, with four blocks of 250 nodes, two blocks of 1000 nodes, and one block of 2000 nodes, average degree 5, and mixing probability of 0.3. Five containment strategies are compared: maximum probability of being infected (Node Infectivity), random link selection (Random), maximum edge betweenness (Betweenness), maximum eigenscore (EigenScore), and maximum link epidemic importance (Link Importance). In the horizontal axis we represent the fraction of removed links during the containment process ($L_r/L$). We show, from top to bottom: $\rho$, incidence of the epidemics on the network; $\rho_{\mathrm{GCC}}$, incidence of the epidemics on the Giant Connected Component; $C_r$, number of connected components; $N_{\mathrm{GCC}}$, size of the GCC; $N_{\mathrm{SLCC}}$, size of the Second Largest Connected Component. The dots mark the achievement of total containment. We observe that our containment strategy based on link epidemic importance outperforms all other methods except eigenscore, with similar results, and the large fragmentation induced by the other strategies.

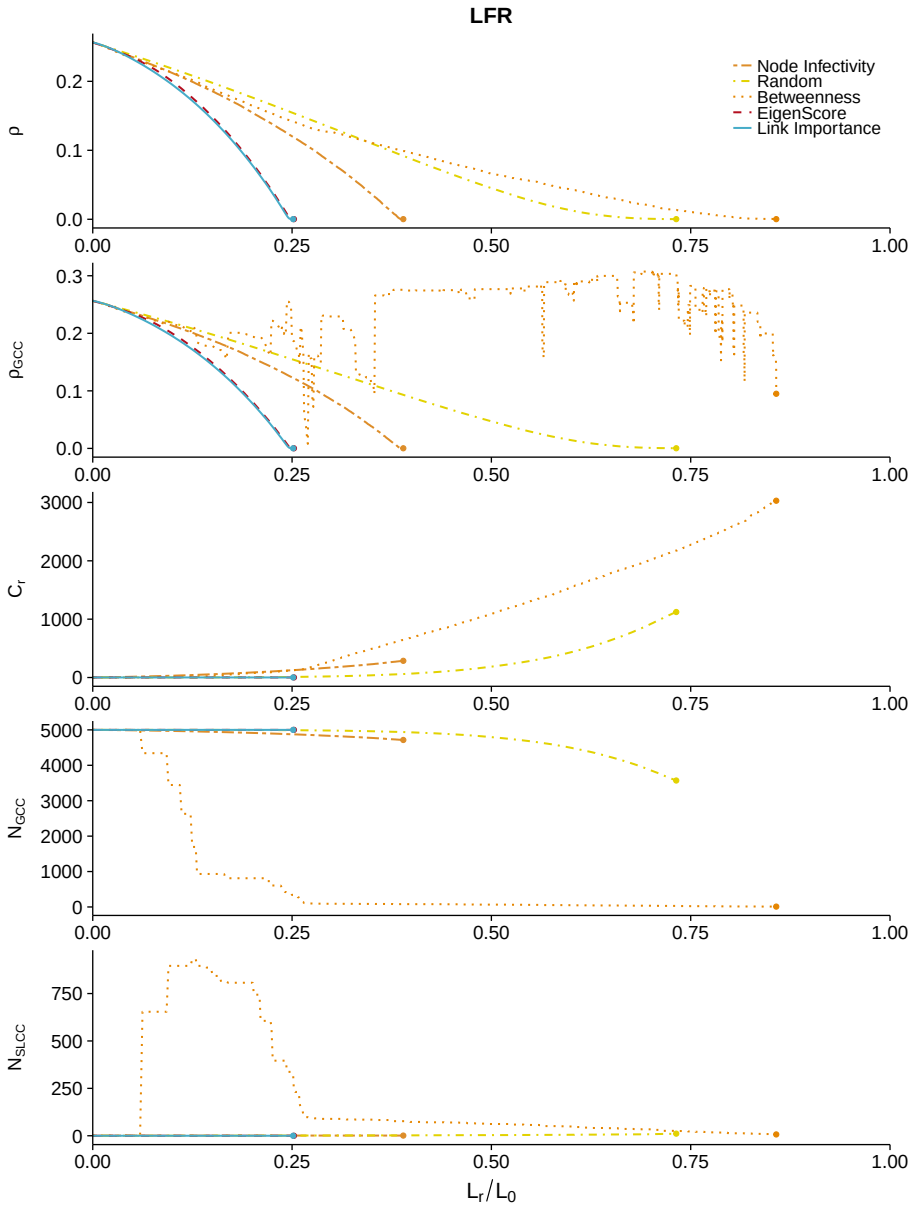


**Fig. S8**. Epidemic containment for a network with 5000 nodes generated using the LFR algorithm [Lancichinetti et al., Phys. Rev. E, 78 (2008) 046110], with average degree of 6, exponent of 3, and mixing probability 0.1. Five containment strategies are compared: maximum probability of being infected (Node Infectivity), random link selection (Random), maximum edge betweenness (Betweenness), maximum eigenscore (EigenScore), and maximum link epidemic importance (Link Importance). In the horizontal axis we represent the fraction of removed links during the containment process ($L_r/L$). We show, from top to bottom: $\rho$, incidence of the epidemics on the network; $\rho_{\mathrm{GCC}}$, incidence of the epidemics on the Giant Connected Component; $C_r$, number of connected components; $N_{\mathrm{GCC}}$, size of the GCC; $N_{\mathrm{SLCC}}$, size of the Second Largest Connected Component. The dots mark the achievement of total containment. We observe that our containment strategy based on link epidemic importance outperforms all other methods except eigenscore, with similar results, and the large fragmentation induced by the other approaches.

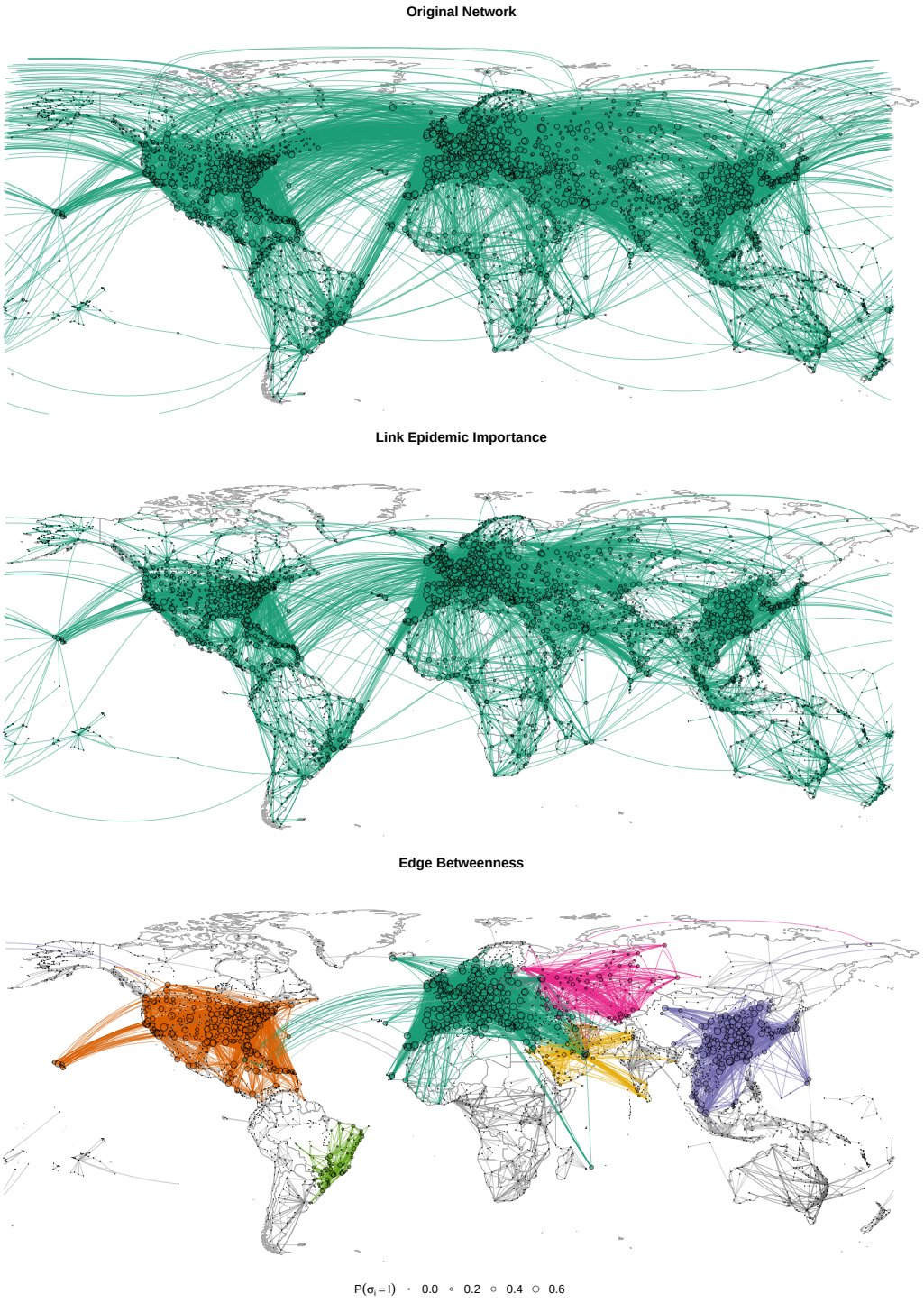


**Fig. S9**. Original air transportation network (top) and the results after a removal of 33.3% of the links using link epidemic importance (middle) and edge betweenness (bottom). Nodes and edges with the same color belong to the same connected component, with subcritical components in gray scale and using darker gray for larger components. The area of the nodes is proportional to their probability of being infected. We have set the epidemic parameters to $\mu = 0.5$ and $\beta = 0.06$.

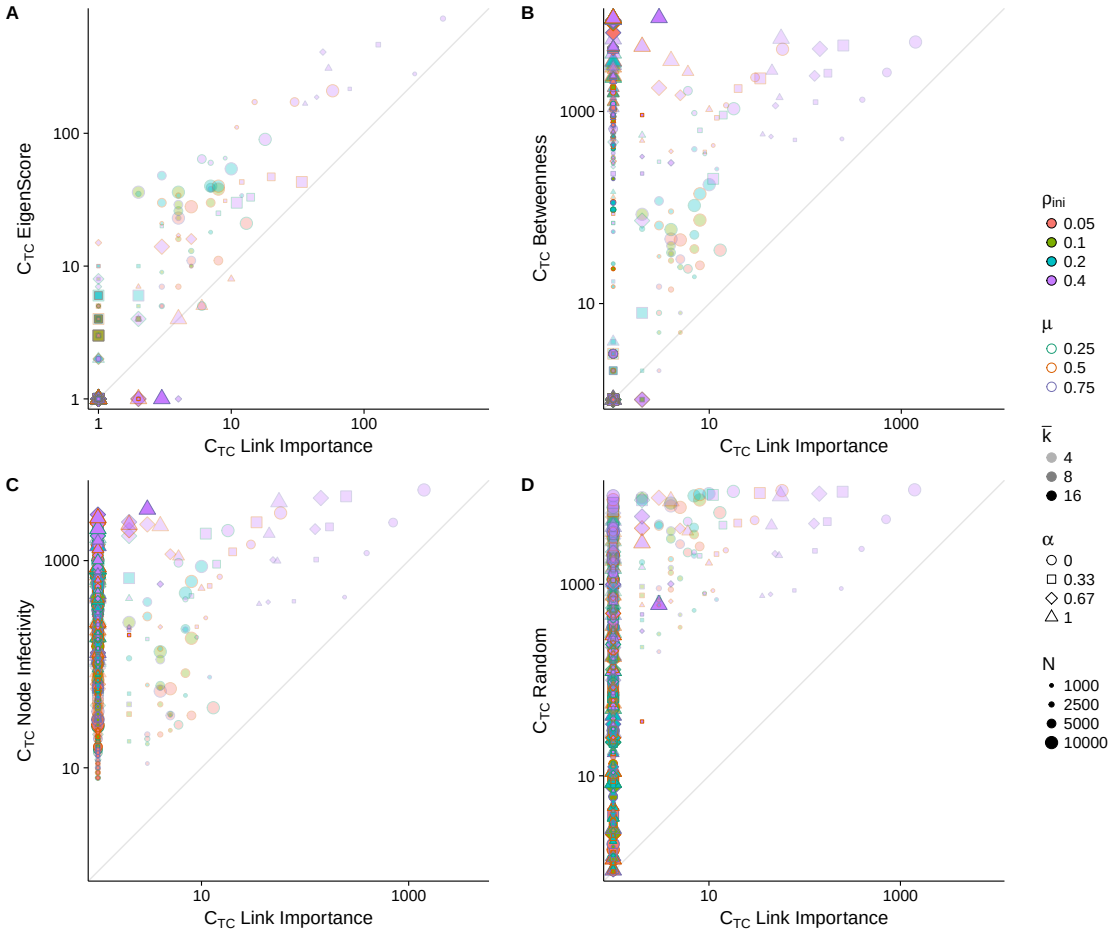


**Fig. S10**. Comparison of the number of connected components after total containment between the link epidemic importance strategy and the other four methods, calculated for the synthetic networks and parameters as in Fig. 4.

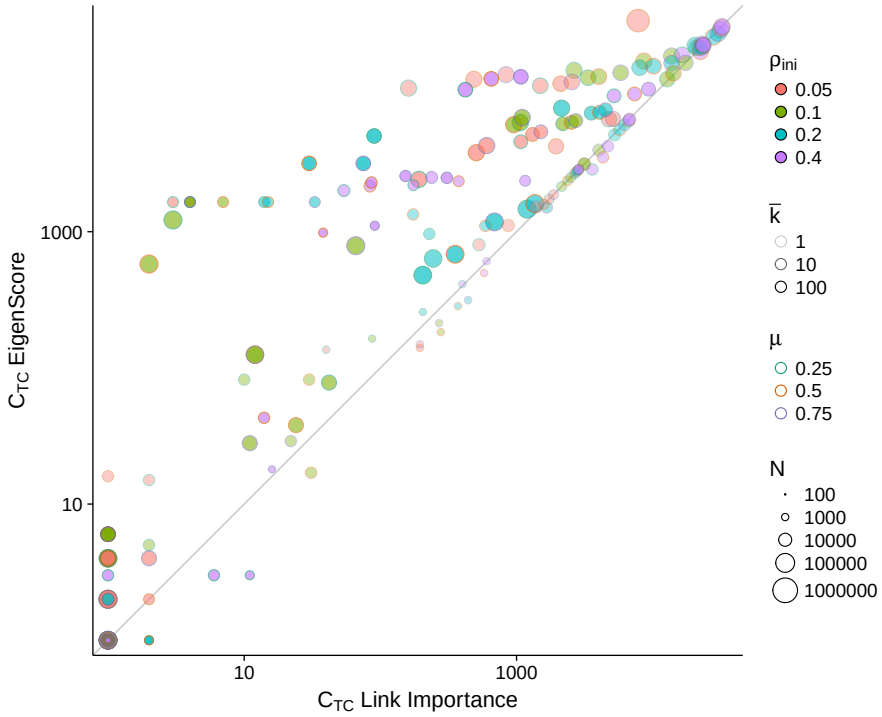


**Fig. S11**. Comparison of the number of connected components after total containment between the link epidemic importance and eigenscore strategies, calculated for the real networks and parameters as in Fig. 5.

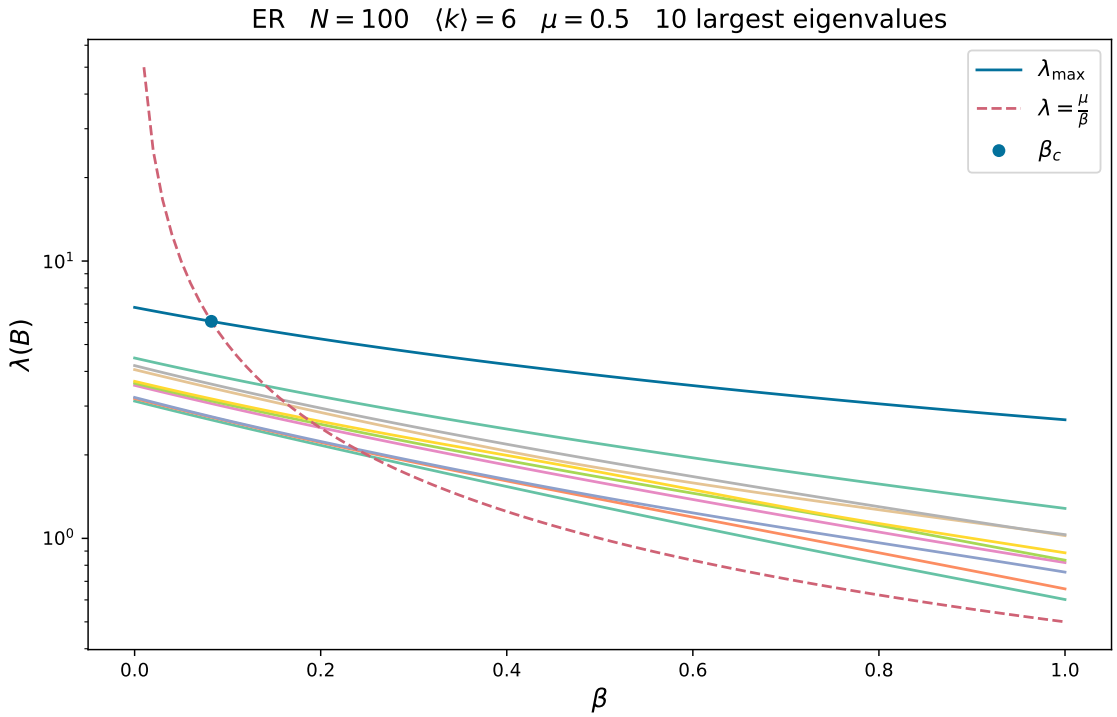


**Fig. S12**. Graphical representation of the determination of the epidemic threshold. The solid lines show the 10 largest eigenvalues of matrix $B$, which depends on the infection probability $\beta$ and the recovery probability $\mu$, for an Erdős-Rényi network of 100 nodes, average degree $\langle k \rangle = 6$, and setting $\mu = 0.5$. For the epidemic threshold, the eigenvalue must be equal to $\frac{\mu}{\beta_c}$, thus it must lay at the intersection between the eigenvalues curves and the dashed line $\frac{\mu}{\beta}$. From those intersections, the one providing the smallest value of $\beta$ is given by the largest eigenvalue of $B$, thus showing that $\beta_c = \frac{\mu}{\Lambda_{\max}(B)}$, which is an implicit equation for the epidemic threshold.

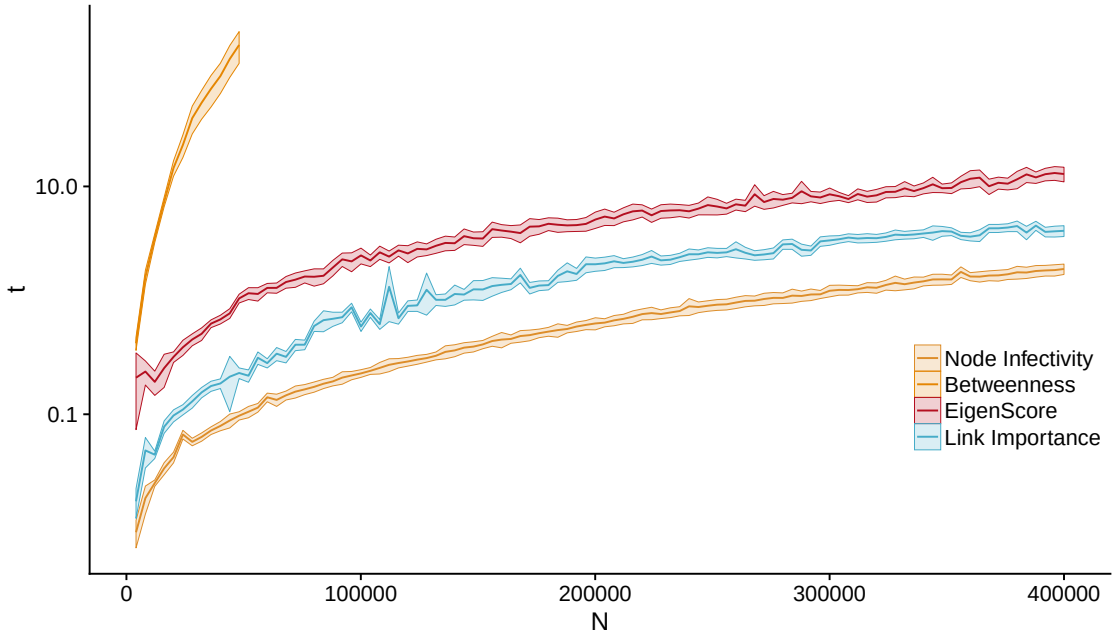


**Fig**. **S13**. Computational time invested for each method to perform a single ranking and removal for BA networks ranging from 100 to 400,000 nodes, averaged over 36 repetitions. The standard deviations are included as ribbons on the plot.

**Table S1**. Structural characteristics of the 27 real networks obtained from the Network Repository (`http://networkrepository.com`) and used in Fig. 6 and fig. S11. They correspond to the largest connected component of the networks with the same name in the repository. The structural descriptors shown are the number of nodes $N$, the number of links $L$, the average degree $\langle k \rangle$, the clustering coefficient $c$, and the assortativity $r$.

| Name | $N$ | $L$ | $\langle k \rangle$ | $c$ | $r$ |
|---|---|---|---|---|---|
| ia-infect-dublin | 410 | 2765 | 13.4878 | 0.1452 | 0.0000 |
| soc-wiki-Vote | 889 | 2914 | 6.5557 | 0.0901 | -0.0556 |
| ca-CSphd | 1025 | 1043 | 2.0351 | 0.0023 | -0.2532 |
| ia-fb-messages | 1266 | 6451 | 10.1912 | 0.6289 | 0.6392 |
| soc-hamsterster | 2000 | 16097 | 16.0970 | 0.0268 | -0.3816 |
| socfb-USFCA72 | 2672 | 65244 | 48.8353 | 0.0037 | -0.3886 |
| socfb-nips-ego | 2888 | 2981 | 2.0644 | 0.3178 | 0.2013 |
| socfb-Santa74 | 3578 | 151747 | 84.8222 | 0.2618 | 0.1253 |
| ca-GrQc | 4158 | 13422 | 6.4560 | 0.2294 | 0.0227 |
| web-spam | 4767 | 37375 | 15.6807 | 0.0006 | -0.8764 |
| power-US-Grid | 4941 | 6594 | 2.6691 | 0.0925 | -0.0952 |
| ca-Erdos992 | 4991 | 7428 | 2.9766 | 0.0439 | -0.0844 |
| soc-advogato | 5054 | 39374 | 15.5813 | 0.0207 | 0.0355 |
| p2p-Gnutella08 | 6299 | 20776 | 6.5966 | 0.0420 | -0.4531 |
| ia-reality | 6809 | 7680 | 2.2558 | 0.1032 | 0.0035 |
| ca-HepTh | 8638 | 24806 | 5.7435 | 0.1907 | 0.0917 |
| soc-anybeat | 12645 | 49132 | 7.7710 | 0.0024 | -0.6753 |
| ca-AstroPh | 17903 | 196972 | 22.0044 | 0.4357 | 0.2258 |
| ca-CondMat | 21363 | 91286 | 8.5462 | 0.1273 | -0.0288 |
| soc-gplus | 23576 | 39145 | 3.3207 | 0.0073 | -0.1946 |
| tech-as-caida2007 | 26475 | 53381 | 4.0326 | 0.0260 | -0.0651 |
| ia-email-EU | 32430 | 54397 | 3.3547 | 0.0004 | -0.6682 |
| ia-enron-large | 33696 | 180811 | 10.7319 | 0.2022 | 0.0706 |
| soc-brightkite | 56739 | 212945 | 7.5061 | 0.2811 | 0.2389 |
| soc-epinions | 61355 | 494372 | 16.1151 | 0.0217 | -0.1234 |
| soc-slashdot | 70068 | 358647 | 10.2371 | 0.0851 | -0.1165 |
| soc-twitter-follows | 404719 | 713319 | 3.5250 | 0.1105 | 0.0096 |