Robustness of complex networks

UNIVERSITA' DI PARMA

Dottorato di Fisica

XXVII ciclo

Relatore di tesi:

Prof. Davide Cassi

Direttore di dottorato:

Prof. Cristiano Viappiani

Autore:

Bellingeri Michele

INDEX

1. Introduction

A complex network is a set of items, named vertices or nodes, with connections between them, called edges or links. Complex networks describe a wide range of systems in nature and society (Albert and Barabasi 2002). Networks (or graphs) can describe real world systems such as the Internet, the World Wide Web, social networks (of acquaintance, organizational among individuals, work relationship, sexual relationship, and so on), networks of business relations between companies, neural networks, metabolic networks, food webs or other ecological networks, distribution networks such as roads, airlines, blood vessels or postal delivery routes, networks of citations between papers, and others (Newman 2003).

Since the network structure is useful to shape a very large range of real world systems, it has been extensively studied in many field of science in the last twenty years. The study of complex networks analyze the properties of these systems, such as the degree distribution, the connectivity, the diameter, the clustering coefficient, and so on; these properties and the network structure are then compared with the aim to describe more general patterns and functioning of the real systems (Albert and Barabasi 2002; Newman 2003).

A fundamental issue concerning the functioning of a complex network is the robustness of the overall system to the failure of its constituent parts (Albert et al. 2000; Holme 2004). The integrity and the functioning of the network can be addressed by analyzing how the network structure changes as vertices or links are removed. To understand how the networks functioning changes with the removal of the vertices is the same to understand the degree of system robustness. Many works have considered how the structure of complex networks change as vertices are removed uniformly at random, in decreasing order of their degree, or in decreasing order of their betweenness centrality (Albert et al 2000; Albert and Barabasi 2002; Holme 2004). The robustness of the network is an interdisciplinary field ranging from social and Internet network, to biological networks as food

webs. For these reasons, the robustness of real-world complex networks, such as Internet, electrical power grids, airline routes, ecological and biological networks to "node failure" (i.e. node malfunctioning or removal) is a topic of fundamental importance for both theoretical and applied network science. Node failure can cause the fragmentation of the network, which has consequences in terms of system performance, properties, and architecture, such as transportation properties, information delivery efficiency and the reachability of network components (i.e. ability to go from node of the network to another). In ecology, food webs have been central to ecological research for decades and the study of the robustness of food webs to species loss is increasingly relevant for species and ecosystem conservation (Dunne et al. 2002; Allesina and Bodini 2004). The loss of a species in ecosystems (primary extinction) can cascade into further extinctions (secondary extinctions), as consumers' persistence depends on the persistence of their resources. Many theoretical and empirical studies have investigated how food web properties, such as modularity, degree-distribution (i.e. the probability distribution of the number of trophic connections of species), presence and distribution of keystone species may influence the pattern of secondary extinctions in ecosystems as well as food web robustness (Dunne et al. 2002; Jordan et al. 2003; Solé and Montoya 2001). In the vast majority of studies on extinction patterns in food webs, a species is assumed to go extinct after a primary extinction when is left without any resources to exploit (Allesina and Bodini 2004; Allesina and Pascual 2009; Dunne et al. 2002; Solé and Montoya 2001).

This thesis investigates the robustness of real and model networks. In the first chapter "*Efficiency of attack strategies on complex model and real-world networks*" we analyse the robustness of physics networks models and real world networks introducing new strategies to remove nodes. In the second chapter "*Optimization strategies with resource scarcity: from immunization of networks to the traveling salesman problem*", we analyse the immunization process in complex networks

introducing the problem of resource scarcity. In the third chapter "*Robustness of ecological networks*" we study the robustness of ecological networks focusing on empirical food webs with a new stochastic methods to select species to remove. Each chapter starts with a summary paragraphs that describes the research focus and the main results.

2. Efficiency of attack strategies on complex model and real-world networks

2.1 Summary

We investigated the efficiency of attack strategies to network nodes when targeting several complex model and real-world networks. We tested 5 attack strategies, 3 of which were introduced in this work for the first time, to attack 3 model networks (Erdos and Renyi, Barabasi and Albert preferential attachment network, and scale-free network configuration models) and 3 real networks (Gnutella peer-to-peer network, email network of the University of Rovira i Virgili, and immunoglobulin interaction network). Nodes were removed sequentially according to the importance criterion defined by the attack strategy, and we used the size of the largest connected component (*LCC*) as a measure of network damage. We found that the efficiency of attack strategies (fraction of nodes to be deleted for a given reduction of *LCC* size) depends on the topology of the network, although attacks based on either the number of connections of a node or betweenness centrality were often the most efficient strategies. Sequential deletion of nodes in decreasing order of betweenness centrality was the most efficient attack strategy when targeting real-world networks. The relative efficiency of attack strategies often changed during the sequential removal of nodes, especially for networks with power-law degree distribution (Bellingeri, Cassi and Vincenzi 2014).

2.2 Introduction

2.2.1 The resilience of complex networks

The resilience of real-world complex networks, such as Internet, electrical power grids, airline routes, ecological and biological networks (Callaway et al. 2000; Albert and Barabasi 2002; Holme et al. 2002; Bodini et al. 2009; Crucitti et al. 2004; Bellingeri et al. 2013) to "node failure" (i.e. node malfunctioning or removal) is a topic of fundamental importance for both theoretical and applied network science. Node failure can cause the fragmentation of the network, which has consequences in terms of system performance, properties, and architecture, such as transportation properties, information delivery efficiency and the reachability of network components (i.e. ability to go from node of the network to another) (Holme et al. 2002).

2.2.2 The attack strategies

Several studies (Holme et al. 2002; Crucitti et al. 2004; Bakke et al. 2006; Dong et al. 2013) have investigated the resilience of model networks using a number of "attack strategies", i.e. a sequence of node removal according to certain properties of the nodes (Albert and Barabasi 2002; Holme et al. 2002; Crucitti et al. 2004). A widely-applied attack strategy consists in first ranking the nodes with respect to an importance criterion (e.g. number of connections or some measure of centrality) and then remove the nodes sequentially from the most to the least important according to the chosen criterion until the network either becomes disconnected or loses some essential qualities (Albert et al 2000; Holme et al. 2000). However, little is known on how the efficiency of attack strategies (i.e. the fraction of nodes to be deleted for a given change in the network) varies when considering different real-world and model networks.

In this context, an underappreciated problem is how the relative efficiency of attack strategies may change during the attack to the network. For example, an attack strategy might be more efficient

when the targeted (i.e. under attack) network is still pristine, while other strategies may be more efficient when the network has already been fragmented and some of its properties have been compromised. Testing the efficiency of the different attack strategies when targeting different networks may also allow to identify the most important nodes for network functioning, and therefore which nodes should be primarily protected, as in the case of computer (Cohen et al. 2001) or ecological networks (Bellingeri et al. 2013; Solé and Montoya 2001; Curtsdotter et al. 2011; Ebenman 2011) or removed, as in the case of immunization/disease networks (Pastor-Satorras and Vespignani 2002).

2.2.3 Introducing new attack strategies on complex networks

In this work, we test the efficiency of both well-known attack strategies and new strategies introduced for the first time in this paper when targeting either model or real-world networks. We used the size of the largest connected component (*LCC*) (i.e. the largest number of nodes connected among them in the network, (Albert and Barabasi 2002)) as a measure of network damage. We found for model networks that the best strategy to reduce the size of the *LCC* depended on the topology of the network that was attacked. For real-world networks, the removal of nodes using betweenness centrality as importance criterion was consistently the most efficient attack strategy. For some networks, we found that an attack strategy can be more efficient than others up to a certain fraction of nodes removed, but other attack strategies can become more efficient after that fraction of nodes has been removed.

2.3.Methods

2.3.1 Attack strategies

We attacked the networks by sequentially removing nodes following some importance criteria. We compared the efficiency of a pool of attacks strategies, some of which have been already described in the literature while others are introduced in this work for the first time.

Most of the analyses on the robustness of network have investigated the effect of removing nodes according to their rank (i.e. number of links of the node) or some measures of centrality (Holme et al. 2000; Albert et al. 2000; Gallos et al. 2006). In this work, we introduce new attack strategies that focus entirely or in part on less local properties of a node, in particular its number of second neighbors, as explained in detail below.

2.3.2 The Largest Connected Component (*LCC*)

Several indexes and measures have been proposed in order to describe network damage. We use the size of the largest connected component (*LCC*), i.e. the size of the largest connected sub-graph in the network (Albert and Barabasi 2002; Holme et al. 2002), as a measure of network damage during the attack, where a faster decrease in the size of the *LCC* indicates a more efficient attack strategy. In order to compare attack strategies across networks, we normalized *LCC* size at any point during the attack with respect to the starting *LCC* size, i.e. the number of nodes in the *LCC* before the attack.

2.3.3 The recalculated attack strategies

For each attack strategy, we applied both the recalculated and non-recalculated method. With the recalculated method, the property of the node relevant for the attack strategy (e.g. number of links) was recalculated after each node removal. On the other hand, when applying the non-recalculated

method the property of the node was measured before the first node removal and was not updated during the sequential deletion of nodes. With *q* we indicate the fraction of nodes removed during the sequential removal of nodes. An attack strategy is less efficient than another when a higher *q* to reduce the *LCC* to zero (or any other size).

In this work, we used 2 attack strategies that have already been described in the literature. *First-degree neighbors (First)*: nodes are sequentially removed according to the number of first neighbors of each node (i.e. node rank). In the case of ties (i.e. nodes with the same rank), the sequence of removal of nodes is randomly chosen. *Nodes betweenness centrality (Bet)*: nodes are sequentially removed according to their betweenness centrality, which is the number of shortest paths from all vertices to all others that pass through that node (Holme et al. 2002; Barthelemy 2004).

2.3.4 The introduced attack strategies

We introduced in the present work the following new attack strategies. *Second-degree neighbors (Sec)*: nodes are sequentially removed according to the number of second neighbors of each node. Second neighbors of node *j* are nodes that have a node in common with - but are not directly connected to - node *j*. *First + Second neighbors (F+S)*: nodes are deleted according to the sum of first and second neighbors of each node. *Combined first and second degree* (*Comb*): nodes are removed according to their rank. In the case of ties, nodes are removed according to their second degree.

For all attack nodes were sequentially removed from most to least connected. In the case of *Bet*, nodes were sequentially removed from higher to lower betweenness centrality. For each network, we tested the relative efficiency of the five attack strategies in reducing the *LCC* to zero. In addition, we tested whether the relative efficiency of attack strategies changed along the removal

sequence, i.e. whether an attack strategies was less efficient than another at the beginning of the attack, but more efficient after a fraction *q* of nodes was removed.

2.3.5 Networks

We tested the new attack strategies on 3 types of model networks and 3 real world networks.

The 6 networks are undirected and unweighted graphs in which nodes are connected by links or edges, and rank *k* of a node is the number of links of that node. Each link may represent several real world interactions. For instance, in social networks links between nodes represent interactions between individuals or groups, such as co-authorship in scientific publications or friendship (Albert and Barabasi 2002). In cellular networks, nodes are chemicals species connected by chemical reactions (Ma and Zeng 2003), while in ecological networks links describe the trophic interactions between species or group of species, e.g. the energy and matter passing from prey to predator (Bellingeri et al. 2013; Ebenman 2011; Dunne et al. 2002).

2.3.6 Model networks

We tested the attack strategies on (*i*) Erdos and Renyi graphs (Erdos and Renyi 1969), (*ii*) Barabasi and Albert preferential attachment networks (Albert and Barabasi 2002), and (*iii*) scale-free network configuration models (Dorogotsev et al. 2008). For each model network, we tested the efficiency of attack strategies on networks of different size, as explained below. Since each model network is a random realization of the network-generating mechanism, we tested the attack strategies on 50 random realizations of each model network used the mean across replicates of the normalized *LCC* size at each fraction *q* of nodes removed as a measure of network damage. We observed a small variation of *LCC* size at each fraction *q* of nodes removed across different realizations of networks, thus the mean *LCC* size across replicates well represented the overall behavior of the attack strategy.

The Erdos and Renyi (*ER*) model generates a random graph with *N* nodes connected by *L* links, which are chosen randomly with an occupation probability *p* from $L_{max} = N(N-1)/2$ possible links, i.e. *p* is the proportion of realized links from L_{max}. The expected number of links is $<L> = (N^2p)/2$ and the expected rank of a node is $<k> = Np$. The random graph can be defined by the number of nodes *N* and the occupation probability *p*, i.e. *ER(N,p)* [21]. We analyzed *ER* graphs with different values of *N* and *p*, specifically: *ER*(*N* = 500, *p* = 0.008), *ER*(1000, 0.004), *ER*(10000, 0.0004).

The Barabasi and Albert preferential attachment network (*BA*) is created starting from few isolated nodes and by then growing the network by adding new nodes and links (Albert and Barabasi 2002). At each step in the creation of the network, one node and *m* outgoing links from the new node are added to the network. The probability θ that the new node will be connected to node *i* already in the network is function of the degree k_i of node *i*, such that $\theta(k_i) = k_i / \sum_{j=N}^{j=1} k_j$ (i.e. preferential attachment, since more connected nodes are more likely to be connected to the new node). The *BA* network is defined by parameters *N* and *m*. We built *BA* scale free networks with parameters *BA*(*N*=500, *m* = 2), *BA*(1000, 2), *BA*(10000, 2).

We created networks with power-law degree distribution using the configuration model for generalized random graphs (Albert and Barabasi 2002; Dorogotsev et al. 2008). This model is defined as follows. A discrete degree distribution $P(K = k) = k^{-\alpha}$ is defined, such that *P(k)* is the proportion of nodes in the network having degree *k*. The maximum node degree k_{max} is equal to *N*, where *N* is the number of nodes. The domain of the discrete function *P(k)* becomes (1, k_{max}). We generated the degree sequence of the nodes by randomly drawing *N* values $k_1,\ldots,k_n$ from the degree distribution. Then, for each node *i* we assigned a link with node *j* with probability $P(k_i)P(k_j.)$ A scale free configuration model network is defined by parameters *N*, α and $<k>$. We analyzed scale

free network with parameters *CM*(*N* = 500, α = 2.5, <*k*> = 3.8), *CM*(1 000,2.5,3.8), *CM*(10 000,2.5,3.9).

2.3.7 Real world networks

We tested the attack strategies on the following real-world networks: (*i*) The Gnutella P2P (peer-to-peer) network (*Gnutella*) (Ripeanu et al. 2004), (*ii*) the email network of the University Rovira i Virgili (URV) in Tarragona, Spain (*Email*) (Guimerà et al. 2003), and (*iii*) the immunoglobulin interaction network (*Immuno*) (Gfeller 2006). Nodes of *Gnutella* (*N*=8846, *L*=31839) represent hosts in the peer-to-peer network, while links represent connections between the hosts. *E-mail* (*N*=1134, *L*=10902) provides an example of the flow of information within a human organization. *Immuno* is the undirected and connected graph of interactions in the immunoglobulin protein (*N* = 1316, *L* = 6300) where nodes represent amino acids, and two amino acids are linked if they interact in the immunoglobulin protein.

2.4. Results

2.4.1 Non-Recalculated method

2.4.1.1 Model networks

ER: For all sizes of networks, the 5 attack strategies were equally efficient in reducing the size of the *LCC* up to *q* ~= 0.2. For *q* >0.2, *First* was the most efficient strategy to reduce the size of the *LCC* to 0.

CM: For *N* = 500, *Comb* was the most efficient strategy early in the removal sequence., while *First* became the most efficient strategy for *q* > 0.1. For *N* = 1 000, *Comb*, *Bet*, and *First* had the same efficiency. For *N* = 10000, *Comb*, *Bet*, and *First* were equally efficient up to *q* = 0.1, while for *q* > 0.1 *First* was the most efficient strategy.

BA: For N = 500, *First*, *Comb* and *Bet* were equally efficient in reducing the size of the *LCC*. For bigger networks, *First*, *Comb* and *Bet* were equally efficient up to q = 0.8 (N = 1 000) and q = 0.5 (N = 10 000). Then, *Bet* became more efficient than *First* and *Comb.*

For graphical representation of these results see Fig. 1.

2.4.1.2 Real-world networks

Email: *Bet* was the most efficient strategy to reduce *LCC* up to $q\sim= 0.3$. For greater fractions of nodes removed, *First* and *Comb* were slightly more efficient than *Bet*.

Immuno: *Bet* was distinctly more efficient than other strategies up to q = 0.55. For $q > 0.55$, all strategies were equally efficient.

Gnutella: *Bet* was the most efficient attack strategy.

For graphical representation of these results see Fig. 2.

2.4.2 Recalculated method

2.4.2.1 Model networks

ER: *First* and *Comb* were the most efficient strategies to reduce the *LCC* up to $q\sim=0.2$. For $q > 0.2$, *Bet* became more efficient than *First. Sec* was the least efficient strategy.

CM: *Comb* was the most efficient strategy up to $q\sim=0.1$. For $q > 0.1$, *Bet* was the most efficient strategy, while *Sec* was the least efficient strategies.

BA: *Comb* was the most efficient strategy up to $q\sim=0.1$. *First*, *F+S* and *Bet* attack induced a slightly slower decrease in *LCC* size. For $q > 0.1$, *Bet* became the most efficient strategy. *Sec* was the least efficient strategy. See Fig. 3.

2.4.2.2 Real-world networks

Email: All attack strategies were equally efficient up to $q = 0.12$. For $q > 0.12$, *Bet* was the most efficient attack strategy.

Immuno: *Bet* was largely the most efficient attack strategy.

Gnutella: All attack strategies were equally efficient up to $q = 0.1$. For $q > 0.1$, *Bet* was the most efficient attack strategy.

For graphical representation of these results see Fig. 4.

2.5 Discussion

We discuss the following main results of our work: (*i*) attacks were largely more efficient with the recalculated than with the non-recalculated method; (*ii*) the efficiency of attack strategies on model networks depended on network topology; (*iii*) the sequential removal of nodes according to their betweenness centrality was the most efficient attack to real-world networks; (*iv*) for some networks, the relative efficiency of attack strategies changed during the removal sequence.

2.5.1 The recalculated methods are more efficient

We found that the recalculated method provided more efficient attacks than the non-recalculated method, i.e. for a given fraction of nodes removed from the network, a larger reduction of *LCC* was obtained with the recalculated method. This result confirms the findings of other analyses on robustness of networks (Albert and Barabasi 2002; Holme et al. 2002), which found that updated information on the topology of the system after each removal allowed for more efficient attacks to networks.

However, non-recalculated attack strategies are implemented in various relevant settings and are equivalent in practice to the simultaneous removal of nodes, as it happens in the case of vaccination campaigns (i.e. the strategy is vaccinating at the same time nodes of the contact network with the highest probability of acquiring or transmitting the disease) or attacks to computer networks (Cohen et al. 2001).

2.5.2 The attack efficiency depends on networks topology

For model networks, the efficiency of the attack strategies depended on network topology. In the case of networks with power-law degree distribution, the efficiency of the attack strategies depended also on network size. Across all model networks and considering both the non-recalculated and recalculated methods, attack strategies based on either node betweenness centrality or node rank were the most efficient ones. However, the sequential deletion of nodes according to their betweenness centrality was consistently the most efficient attack strategy to real-world networks, with the only exception of the attack to the *Email* network with the non-recalculated method. While in some cases *Bet* was only slightly more efficient than other strategies in reducing the size of the largest connected component, in others *Bet* was largely the most efficient strategy. For example, in the immunoglobulin interaction network, deleting a very small fraction of nodes with high betweenness centrality reduced the size of the normalized *LCC* of more than 60% using either the recalculated and non-recalculated method, while - for the same fractions of nodes removed - other attack strategies caused only a 1-5% reduction in *LCC* size. Betweenness centrality describes how "central" a node is in the network by considering the fraction of shortest paths that pass through that node (Barthelemy 2004). Nodes with betweenness centrality greater than 0 play a major role in connecting areas of the network that would otherwise be either sparsely connected or disconnected (Newman 2003). Thus, betweenness centrality an important centrality measure for a social, technological, computer, and biological networks. The higher efficiency of the strategy based

on node betweenness centrality with respect to the attack based on node rank in real-world networks can be explained by the fact that in real-world networks some of the critical nodes (i.e. nodes whose persistence strongly contribute to maintaining network integrity) are either not highly linked, or that the highly-linked nodes are not located in the network core (Newman 2003).

2.5.3 The new introduced Combined strategy is more efficient in some cases

When applying the recalculated method, the newly-introduced *Combined* attack strategy was the most efficient strategy to decrease *LCC* size in the scale free network configuration model and in the Barabasi-Albert model up to $q = 0.1$. The *Combined* attack first select nodes according to their rank, then, in the case of ties (i.e. nodes with the same rank), it sequentially removes nodes according to their second degree. On the contrary, in the case of ties *First* randomly chooses the removal sequence for the nodes with the same rank. Thus, at the beginning of the attack to the network, when two or more major hubs have the same number of links to other nodes, removing first the hub with the greatest second degree causes a faster decrease in *LCC* size than to randomly select the removal sequence for those hubs.

Later in the attack sequence, the *Combined* strategy was less efficient than the *First* strategy to attack scale free networks; this might be due to the fact that after a certain fraction of hubs has been deleted, removing first (in the case of ties) the node(s) with the highest second degree(s) would remove more peripheral and less important nodes, i.e. nodes that are less likely to be part of the largest connected component.

2.5.4 The efficiency changed along the removal sequences

Further, the efficiency of attack strategies changed along the sequential removal of nodes. This was particularly clear for networks with power-law degree distribution. It follows that the percolation threshold, i.e. the fraction of nodes removed for which the size of the largest connected component

reaches zero, might be for some networks little correlated with the fraction of nodes to be removed in order to reduce the largest connected component to a size greater than 0. This result has important implications for applied network science and deserves further investigations. For example, in the case of immunization strategies, choosing the attack strategy according to the percolation threshold may be of little use when the goal is to reduce as much as possible the size of *LCC* with just a few targeted immunizations. Lastly, the use *LCC* as a measure of the efficiency of the network may not be appropriate for immune networks. Immune networks, such as neural or lymphocyte networks, reveal a specific and non-trivial architecture and they can display peculiar features when diluted. For this reason, differently from what happens in other kind of systems, when in immune networks the *LCC* decreases, the performance of the network can actually improve (Agliari et al. 2012; Agliari et al. 2013)

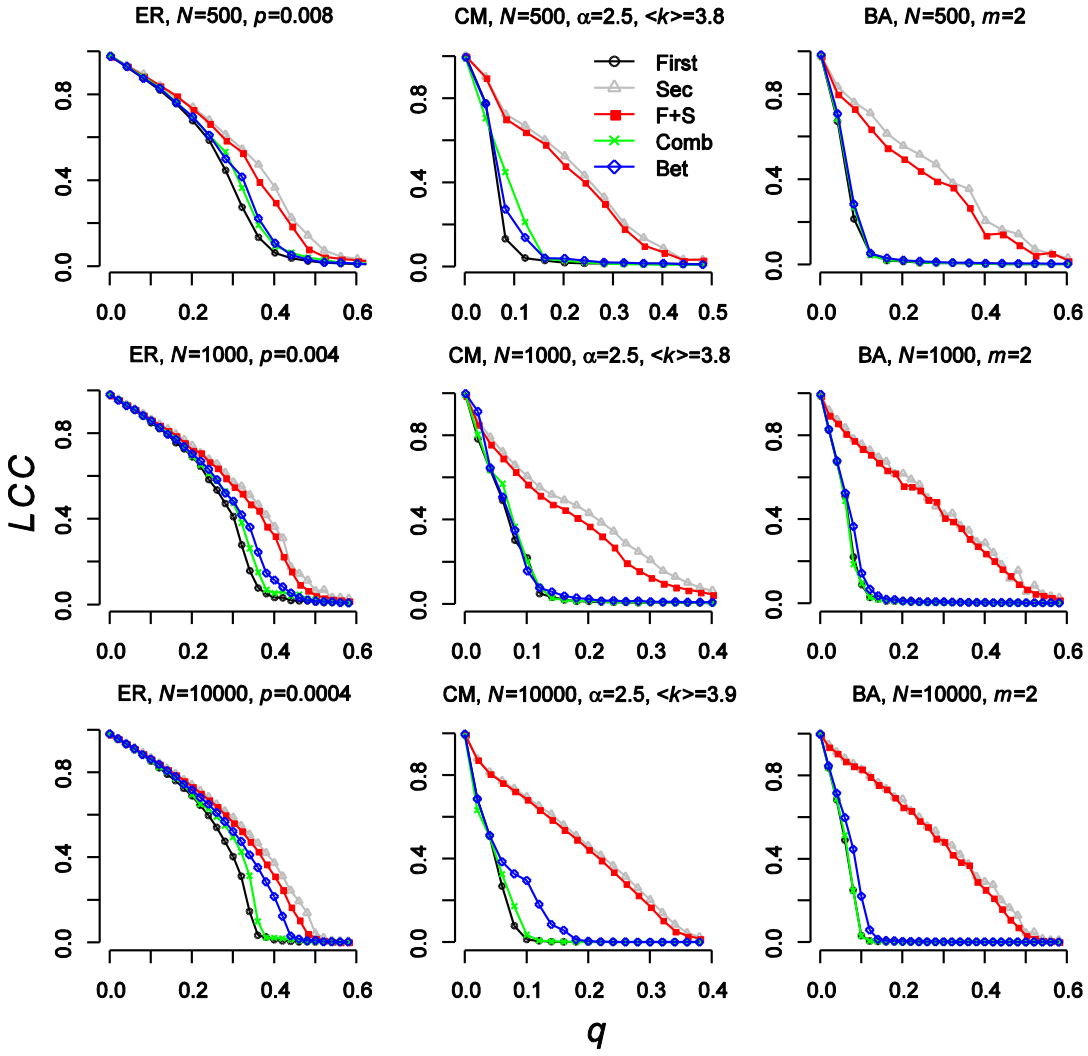

Figure 1. Size of normalized *LCC* and the fraction *q* of nodes removed for non-recalculated targeted attacks to model networks. Points are plotted every 20 nodes removed for networks with $N = 500$ and $N = 1000$, and every 200 nodes removed for $N = 10000$.

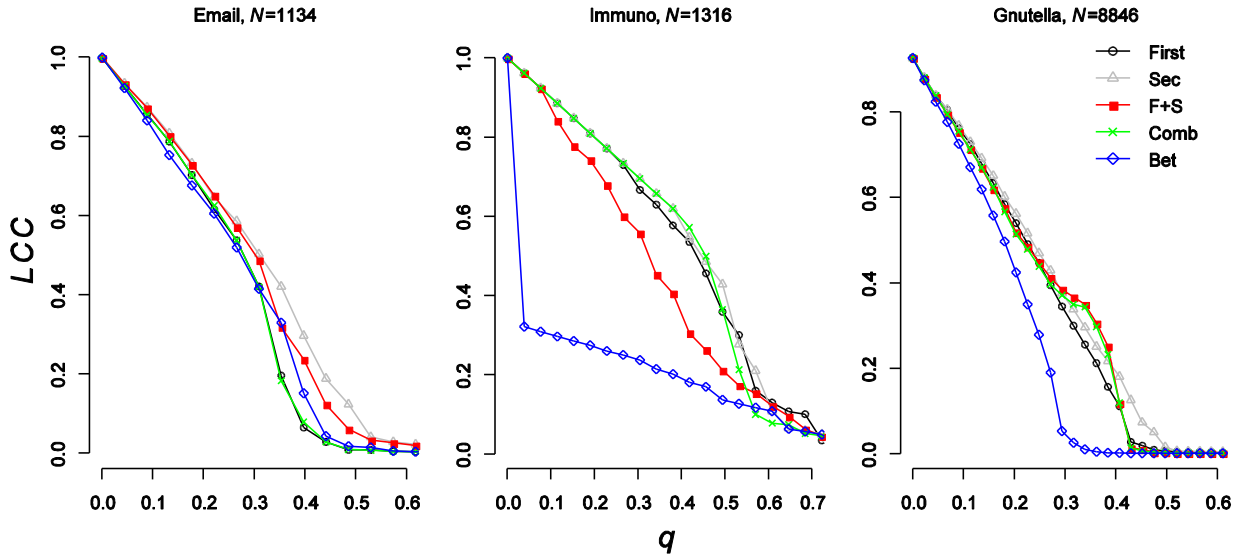

Figure 2. Size of normalized *LCC* and the fraction *q* of nodes removed for non-recalculated targeted attacks to real-world networks. Points are plotted every 50 nodes removed for *Email* and *Immuno* networks, and every 200 nodes removed for *Gnutella*.

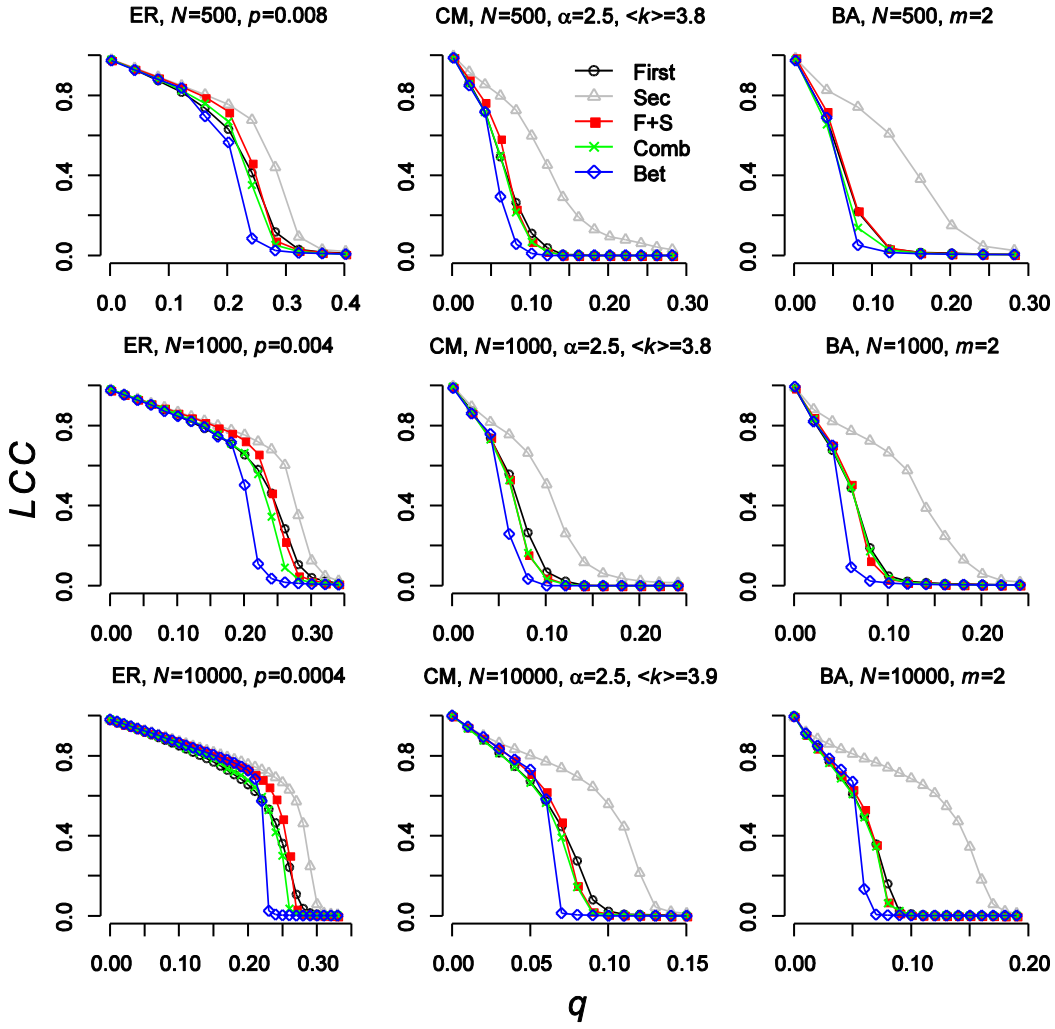

Figure 3. Size of normalized *LCC* and the fraction *q* of nodes removed for recalculated targeted attacks to model networks. Points are plotted every 20 nodes removed for networks with $N = 500$ and $N = 1\,000$, and every 200 nodes removed for $N = 10000$.

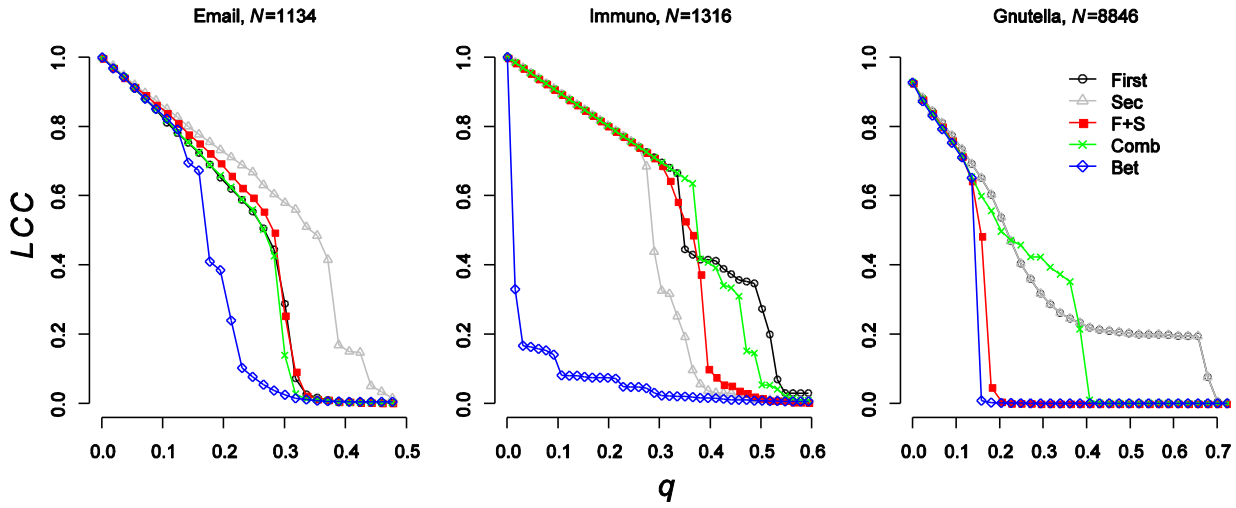

Figure 4. Size of normalized *LCC* and the fraction *q* of nodes removed for recalculated targeted attacks to real-world networks. Points are plotted every 50 nodes removed for *Email* and *Immuno* networks, and every 200 nodes removed for *Gnutella*.

3. Optimization strategies with resource scarcity: from immunization of networks to the traveling salesman problem

3.1 Summary

The development of efficient immunization strategies is a fundamental issue in the cross-field research between the physical network theory and the biological science. In these cases one typically applies targeted immunization strategies where individuals playing a special role in the network community are immunized first. Recently, many different immunization strategies have been developed. The efficiency of a strategy is usually evaluated in terms of percolation threshold, i.e. the number of vaccine doses producing the vanishes of the Largest Connected Cluster (*LCC*). The immunization strategies inducing the fastest vanishes in the *LCC* (e.g. the minimum percolation threshold) represents the optimal way to immunize the network. In this work we simulate several immunization processes on different kinds of model and real world networks. Here we show that the efficacy of the removal strategies can change during the immunization process. This means that in the first fraction of immunized individuals the strategy A can perform better than the strategy B. In correspondence of a certain fraction of vaccination q^* we assist to the efficacy transition and the strategy B becomes more efficient than A. We call the q^* the 'transition threshold'. This means that, if the number of doses is limited, the best strategy is not necessarily the one leading to the smallest percolation threshold. This outcome should warn about the adoption of global measures in order to evaluate the best immunization strategy. In addition, we evidence analogous phenomena for the traveling salesman problem (TSP). This new perspective has important implications for health policies, such as the tumor angiogenesis research, for immunology processes of short generation time virus (e.g., HIV), or in general, to evaluate the best solution of optimization problems in the presence of some constraints (Bellingeri, Agliari and Cassi 2015, under review).

3.2 Introduction

3.2.1 Immunization strategies in complex networks

Complex networks describe a wide range of systems in nature and society, such as Internet, electrical power grids, ecological, biological and social networks (Callaway et al. 2000; Albert and Barabasi 2002; Bellingeri et al. 2013; Agliari et al. 2011; Agliari et al. 2012). In this field, the immunization of large populations or apparati through vaccination or intrusion detection protocols is an extremely important issue with obvious implications for the public health and security (Schneider et al. 2011; Anderson and May 1992; Gallos et al. 2005; Hadidjojo and Cheong 2011). With this goal, many immunization strategies have been developed in the last years (Hadidjojo and Cheong 2011; Pastor-Satorras and Vespignani 2002; Serrano et al. 2009; Chen et al. 2008; Schneider et al. 2012; Iyer et al. 2013). Considering the topology of the network, the question how to immunize a given network with a minimum number of operations (e.g., vaccine doses) is mathematically equivalent to understand how to fragment the Largest Connected Cluster (*LCC*) with a minimum number of node removals (Hadidjojo and Cheong 2011; Chen et al. 2008), (see Figure 5 for an example of the immunization method).

In recent papers, the efficacy of the removal strategy is usually evaluated in terms of the percolation threshold q_c, i.e. the fraction of nodes removed for which the *LCC* reaches the value of zero (Chen et al. 2008; Hasegawa and Naoki 2011; Huang et al. 2011; Samuelsson and Socolar 2006; Zeng and Liu 2012). In all these cases one supposes to rely on potentially unlimited resources (e.g., unlimited vaccine doses, fuel, time, budget).

3.2.2 Immunization strategies with limited resources

Here, we start from an assigned amount of resources and we evaluate the best strategy as a function of this constraint. Using different strategies for selecting nodes, we present simulated immunization processes on different real world networks, i.e. the immunoglobulin interaction network (Gfeller

2007) and the Email network of the University Rovira i Virgili (Guimerà et al. 2003), and model networks, that is the "Hebbian networks" (Agliari and Barra 2012) and the Erdös-Rényi random graphs (Erdös and Rényi 1960). See the Methods section for a detailed explanation of the used immunization strategies.

3.3 Methods

3.3.1 Immunization strategies

We simulated the network immunization by sequentially vaccinating (removing) nodes (individuals) following some importance criteria. We compared a pool of immunization strategies. Most of the analyses on immunization processes on networks have investigated the effect of vaccinating nodes according to their rank (i.e. number of first or second neighbors) or some measures of centrality (Dorogotsev et al. 2008; Albert and Barabasi 2002). For each immunization strategy, we applied both the recalculated and non recalculated methods. In the recalculated method, the property of the node relevant for the immunization strategy (e.g. number of neighbors or centrality) was recalculated after each node immunization. In the non recalculated method, the property of the node was computed before the first node removal and was not updated during the process.

Several indexes and measures have been introduced to describe network immunization. In this research, we use the size of the largest connected component (*LCC*), i.e. the size of the largest connected sub-graph in the network (Chen et al. 2008; Zeng and Lyu 2012), as a measure of nodes vaccinated during the immunization process. A faster decrease in the size of the *LCC* indicates a more efficient immunization strategy.

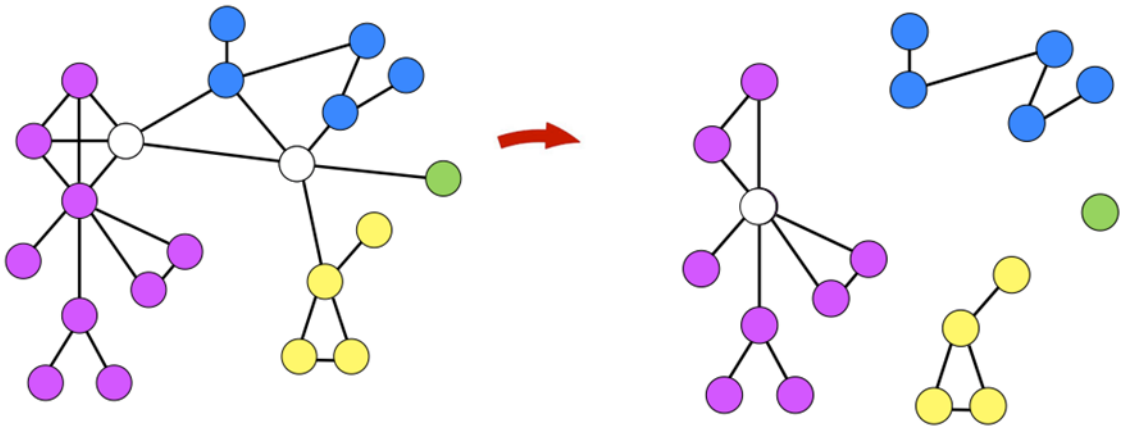

Figure 5: **The immunization procedure**. Starting from the original network (on the left) two nodes (in white color) are selected and consequently removed. The resulting network (on the right) turns out to be disconnected into four components. At this stage, the largest component can be further processed and one node (in white color) can be removed in such a way that this component will be fragmented in four small components. In this example targeted nodes are detected according to the betweenness strategy.

In literature, an immunization strategy is less efficient than another when a higher the fraction of nodes has to be vaccinated to reduce the *LCC* to quasi zero. With *q* we indicate the fraction of nodes removed during the sequential immunization of nodes. In our research we consider the efficacy of immunization strategy during the whole process. After the simulation, we contrasted the different strategies as a function of the immunized nodes *q*. *First-degree neighbors*: nodes are sequentially immunized according to the number of first neighbors of each node. *Second-degree neighbors*: nodes are sequentially immunized according to the number of second neighbors of each node. Second neighbors of a node are nodes that have a node in common with, but are not directly connected to, the node. *First + Second neighbors*: nodes are vaccinated according to the sum of first and second neighbors of each node. In the case of ties (i.e. nodes with the same rank), the sequence of nodes vaccinations was randomly chosen. *Nodes betweenness centrality*: nodes are sequentially vaccinated according to their betweenness centrality. The betweenness centrality of a node represents the number of shortest paths from all vertices to all others that pass through that node. For all strategies nodes were sequentially immunized from most to least ranked.

3.3.2 Networks

The networks we used are undirected and unweighted graphs in which nodes are connected by links or edges, and rank *k* of a node is the number of links of that node. Each link may represent several real world interactions. In social networks links between nodes represent interactions between individuals or groups, such as co authorship in scientific publications or friendship (Callaway et al. 2000; Albert and Barabasi 2002). In cellular networks, nodes are chemicals species connected by chemical reactions (Zeng and Lyu 2012), while in ecological networks links describe the trophic interactions between species or group of species, e.g. the energy and matter passing from prey to predator (Bellingeri et al. 2013). Moreover, in immunological networks, nodes represent lymphocytes or antibodies and a link connecting a couple of nodes mirrors a large enough affinity between the related receptors (Agliari et al. 2012).

3.3.3 Model networks

We tested the attack strategies on Erdös-Rényi random graphs and on the Hebbian network.

The Erdös-Rényi (ER) model generates a random graph with *N* nodes connected by *L* links, which are chosen randomly with an occupation probability *p* from *L*max = *N(N-1)/*2 possible links, i.e. *p* is the proportion of realized links from *L*max. The expected number of links is $<L> = (N^2p)/2$ and the expected rank of a node is $<k> = Np$. The random graph can be defined by the number of nodes *N* and the probability *p*, i.e. ER(N,p) (Erdös and Rényi 1960).

The Hebbian network model generates a random graph with *N* nodes, each associated to a set of *P* binary attributes, namely $\{\xi i^1, \xi i^2, ..., \xi i^P\}$ for the *i*-th node, in such a way that the link connecting two nodes, say *i* and *j*, has a weight given by the Hebbian kernel $Jij = \sum_{\mu=1}^{P} \xi i^\mu \xi j^\mu$. The *N*x*P* entries $\{\xi i^\mu\}_{i,\mu}$ are identically and independently extracted from a uniform distribution $P(\xi i^\mu=0)=1-$

$P(\xi i^{\mu}=1)=(1-a)/2$, where $a \in [-1,1]$. Thus, by tuning *a*, different topological regimes can be recovered (Agliari and Barra 2011). Moreover, the emerging topology displays a number of peculiar features such as the small-world property and it recovers the Granovetter theory.

3.3.4 Real world networks

We tested the attack strategies on the email network of the University Rovira i Virgili (URV) in Tarragona, Spain (*E-mail*) (Guimerà et al. 2003), and the immunoglobulin interaction network (*Immuno*) (Gfeller 2007). As explained in the previous chapter, the *E-mail* (N=1134, L=10902) provides a representative example of the flow of information within a human organization and the *Immuno* is the undirected and connected graph of interactions in the immunoglobulin protein (N = 1316, L = 6300) where nodes represent amino acids and two amino acids are linked if they interact in the immunoglobulin protein.

3.3.5 The travelling salesman model (TSP)

We considered the travelling salesman problem (TSP), which consists in finding, for a given list of cities and distances between each pair of cities, the shortest possible route that visits each city exactly once and returns to the origin city. Otherwise state, defining a complete graph $G=(V,E)$ whose nodes (making up the set V) represents cities and whose links (making up the set of ordered pairs E) display a weight measuring the distance between the cities connected pairwise, the TSP asks for the cycle of minimum cost visiting all of the vertices of G exactly once. This problem is known to be NP-hard as, being n the number of cities, if we try to determine the solution of this problem systematically, we would end up with $(n-1)!$ possible solutions. Clearly, we cannot examine all possible solutions for minimum cost and several algorithms for a numerical solution of the problem have been built up. In general, one could follow empirical or practical algorithms according to the particular problem considered. For instance, one could follow a ``myopic route"

flying, step by step, to the closest city. This can be formalized as follows: starting from *i* we move to *j* such that $min\{k \in V\}d(i,k)=j$; in the following step we move to *z* such that $min\{ k \in V \setminus j\}d(i,k)=z$ and so on.

3.4 Results

3.4.1 The crossover threshold

We find that the rank of efficacy of the removal strategies can change depending on the amount of available resources. For example, in the first steps of an immunization process a strategy A can perform better than a strategy B. In correspondence of a certain fraction of removals q^* we assist to the efficacy transition and the strategy B becomes more efficient than A. We call the q^* the 'crossover threshold' (see Fig. 6). In this case, the percolation threshold (i.e. the fraction of nodes to be removed for the size of the largest connected component to vanish), commonly used as a measure of the strategy efficacy (Chen et al. 2008; Schneider et al. 2012; Iyer et al. 2013; Hasegawa and Naoki 2011; Huang et al. 2011; Samuelsson and Socolar 2006; Zeng and Lyu 2012), may be misleading, since the goal is to reduce as much as possible the size of *LCC* with just the few "shots" available. What matters here is being able to fast reduce the size of the *LCC* in the early stage of the immunization process and this, for many networks, may be little correlated with the fraction of nodes to be removed to underpercolate the network. In fact, the latter considers only the ending point of the immunization process, i.e. the total number of individuals that we have to immunize in order to completely destroy the *LCC* and consequently to reduce to zero the probability of new infection events.

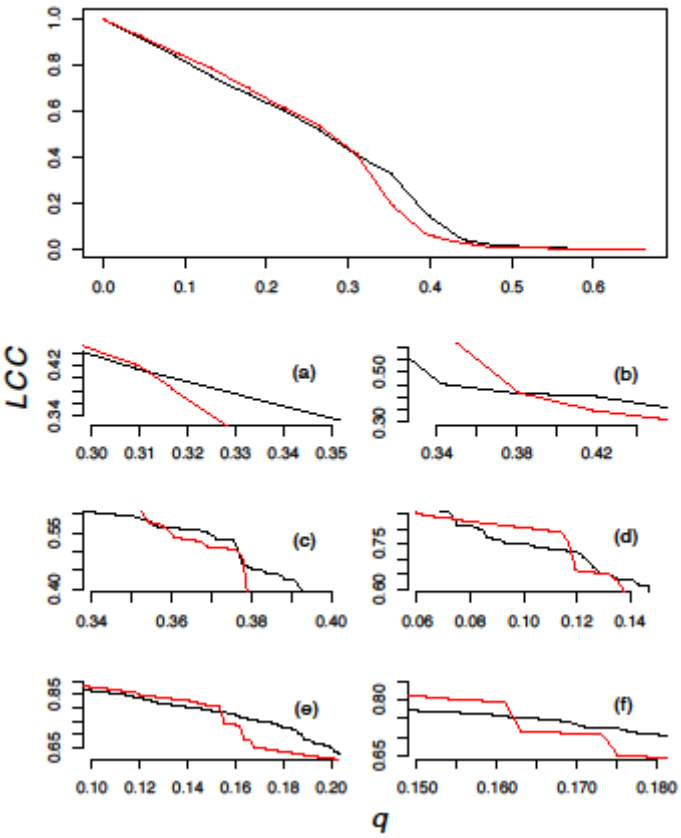

Figure 6: **The crossover threshold in the immunization process outcomes**. The *LCC* versus the fraction of immunized nodes *q* for different strategies and for different networks. The first main panel on the top shows the whole immunization process for the Email real world network; in the little panels (a-e) are highlighted the crossover threshold for different real world networks. (a) black line is the system response using non-recalculated betweenness strategy where nodes are removed according to their betweenness centrality, which is the number of shortest paths from all vertices to all others that pass through that node, and red line the removal based on the non-recalculated first degree where nodes are sequentially removed according to the number of first neighbors for the *Email* real world network; (b) black line indicates the recalculated first degree strategy and the red line the recalculated degree strategy for the *Immuno* real world network; (c) the black line is the recalculated first degree strategy and the red line represents the recalculated first+second degree strategy for the Hebbian model with parameters N=500, L=300, a=-0.983; (d) we focus on the cross between the black lines representing the random removal strategy and the red line represents the recalculated first+second degree strategy response for the Hebbian model with parameters N=500, L=500, a=-0.99: crossover threshold between recalculated first degree strategy (black line) and recalculated betweenness strategy (red line) for a Erdös-Rényi random graphs (e) N=500, p=0.008 and (f) N=500, p=0.007.

3.5 Discussion

3.5.1 The efficacy of the strategies depends on the number of vaccine doses

Having enough available doses to immunize the entire population is, of course, the best scenario, yet hardly applicable in real world conditions. In real world scenarios, the shortage of the resources affects all the decision policies. For example, when a political subject has to decide which strategy to adopt in order to immunize in the best efficient way a population, it has also to consider the limited number of vaccine units available (due to e.g., constraints in budget, qualified staff, number of doses (Dupuy and Freidel 1990; Schwalbe et al. 2010)). Further, the vaccination procedures take time, either for planning the operations, or to carry out the practical immunization of the individuals or the production mechanisms of the immunization agents may be inadequate or slow to supply the actual request. For this reason, the distributed immunization doses may cover only a small fraction of the entire population.

3.5.2 Resource shortage and policies in different type of model networks

The pattern shown here can be extended to many different contexts and with different purposes. In fact, if the immunization process is looked at as an attack process, where selected nodes are impaired, then effective immunization processes also turn out to be effective attack processes able to disconnect a network with a limited number of interventions. Thus, being aware of the existence of effective attack strategies for a given topology may allow a proper protection of the network by, e.g., reinforcing the most susceptible nodes. Moreover, in the case of denial-of-service (DoS) attack to a network, one could have to face the limited size of the botnet (Yu 2014). On the other hand, there may exist situations where the destruction or at least a rearrangement of the network is sought for.

For instance, in tumor angiogenesis research, a proper network attack is nowadays strongly investigated as a practice to inhibit tumor growth (Welter and Rieger 2013; Jain et al. 2007). Here the constraints could arise due to the individual resilience to the therapy toxicity.

Immunology offers other examples of systems where the amount and the effectiveness of resources play a fundamental role. In this case the threat is played by a virus load and the organization to be preserved is given by the lymphocyte network: Once a virus has invaded a lymphocyte cell, it starts to manipulate the host cell in such a way that the viral genome is multiplied and new virus particles are assembled and then released from the host cell. From this perspective it is the network of lymphocyte cells to be attacked: infected cells produce new virus, and new free virus infects healthy cells to produce new infected cells. Depending on the average life-time of virions and of infected cells, and according to the time scale of virus mutation, the outcome can be dramatically different (Nowak and May 2000).

In the case of virus mutating in fast time-scales, that is displaying a short generation time (this is the case of e.g., HIV), being aware of the number of cells that can be effectively targeted for further lysis in a given amount of time has strong relevance for the kind of strategy to implement (Nowak and May 2000).

3.5.3 The critical threshold pattern in other fields: the traveling salesman problem

In the end we discuss the crucial role of scarcity of resources in optimization problems in different fields. We consider the traveling salesman problem (TSP), which consists in finding, for a given list of cities and distances between each pair of cities, the shortest possible route that visits each city exactly once and returns to the origin city.

Otherwise stated, defining a complete graph *G* whose nodes represents cities and whose links display a weight measuring the distance between the cities connected pairwise, the TSP asks for the

cycle of minimum cost visiting all of the vertices of *G* exactly once. This problem is known to be NP-hard and it finds applications in many research areas such as operations research, theoretical computer science, and logistic (Applegate et al. 2006).

We simulated the salesman who visits the European capitals (Fig. 7): the best route obtained is compared with a "myopic route" which, at each step selects the closest city; for both strategies we account the number of visited cities as a function of the overall distance covered (i.e. time or consumed fuel), (see Fig. 7).

Interestingly, although the latter route takes more time to accomplish the complete task, its early rate is significantly larger. Indeed, we can still outline a crossover threshold q^* which roughly corresponds to 75% of the overall investment. Moreover, at half of the investment the myopic strategy has visited almost twice cities. It should also be noted that this strategy also requires a negligible computing time for its evaluation with respect to the optimized way.

This means that the efficacy of a strategy (number of visited towns) depends on the available resources (fuel). This has important consequences especially when some constraints on fuel, time or, more generally, "coverage" are included. In fact, an overall good strategy may result to perform rather poorly at short coverage.

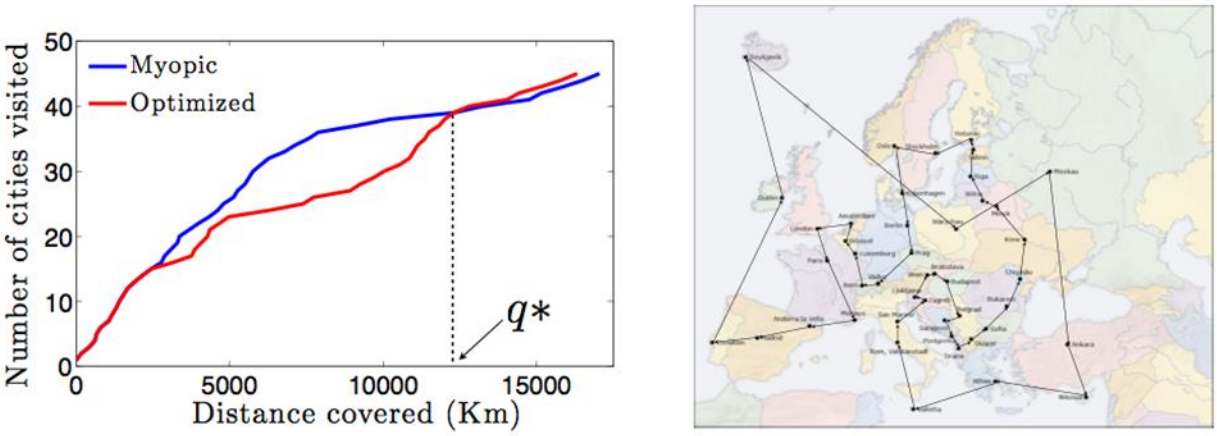

Figure 7: **The travelling salesman problem**. The set *V* of European cities and the set E of their pairwise distances are considered. We found a solution for the TSP on *G(V,E)* according to different strategies (left panel): myopic (blue), and best (black). The starting point is fixed on Rome and the path corresponding to myopic strategy is depicted (right panel). Notice that if, due to constraints, the overall distance covered should be limited to, say, 20 or to 70, respectively, the strategy ranking would be different.

4. Robustness of ecological networks

4.1 Summary

The robustness of ecological networks, i.e. the food webs, is a fundamental topic in ecology. The robustness of food web is generally analysed removing species in the systems and counting how many species go secondarily extinct. The removal can be at random or selecting most connected species (attack). Random removal and the attack from most- to least-connected node (i.e. species) are the two limit criteria for sequential extinction of species in food webs, but a continuum of possibilities exists between them.

We use simulations to test the robustness of 14 empirical food webs to species loss by varying a parameter *I* (intentionality) that defines the removal probability (extinction risk) of species with high number of trophic connections. The removal probability of highly-connected species increases with *I*. We found that food web robustness decreases slowly when the extinction risk of highly-connected species increases (we call this region *random removal regime*), until a threshold value of *I* is reached. For greater values of the threshold, we found a dramatic reduction in robustness with increasing intentionality in almost all the food webs (*intentional attack regime*).

Link-dense networks were more robust to an increase of *I*. Larger food webs (i.e. higher species richness) were more sensitive (i.e. robustness decreased faster) to the increase of extinction risk of highly-connected species. The existence of a clear transition in system behaviour has relevant consequences for the interpretation of extinction patterns in ecosystems and prioritizing species for conservation planning (Bellingeri, Cassi and Vincenzi 2013).

4.2 Introduction

4.2.1 Ecological networks: the food webs

Food webs are ecological networks that describe the feeding (trophic) relationship among diverse species in communities or ecosystems (Camerano 1880; Cohen et al. 1990; Dunne 2006; Bellingeri and Bodini 2013). In other terms, a food web is the natural interconnection of food chain and generally a graphical representation of what-eats-whom in an ecological community (Allesina and Bodini 2004). In the graph-network representation of the food web nodes indicates species and a link (or edges) connecting two nodes is the trophic relationship between the two species. The food webs are directed networks and the direction of the link indicates the transferring of energy and matter from the prey (or resource) and the predator (or consumer). See Figure 8 for a very simple example of a food web and Figure 9 for a detailed graphical depiction of the intertidal food web of Sanak Island, Alaska with 171 species.

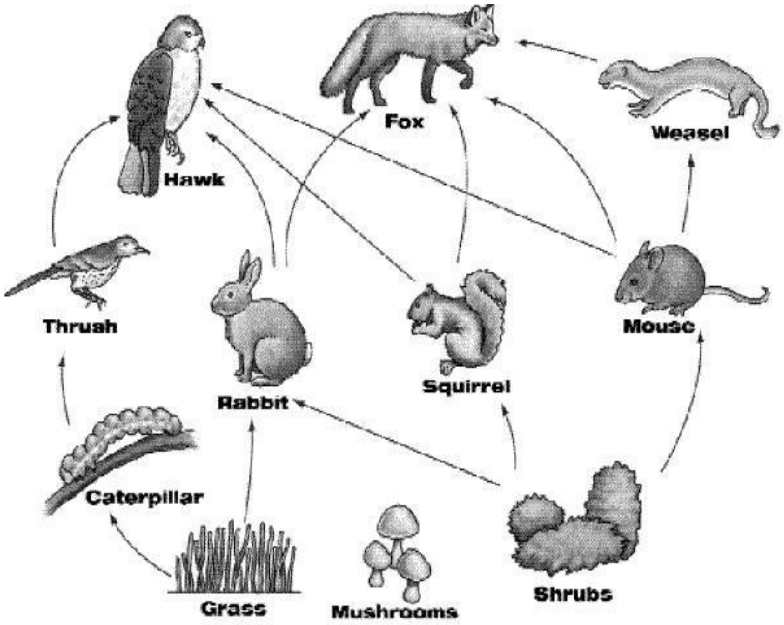

Figure 8: simple example of a terrestrial food web (http://eatingrecipe.com/).

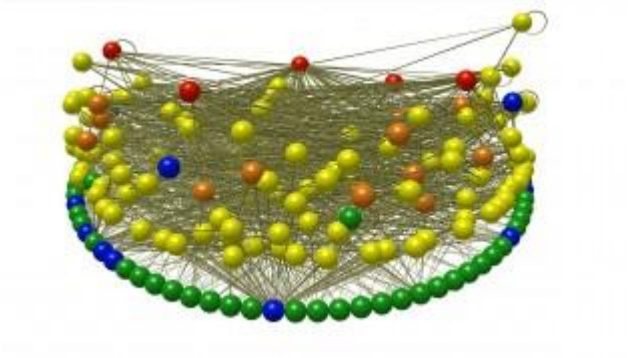

Figure 9: Depiction of the intertidal food web of Sanak Island, Alaska, spheres represent species or groups of species, and the links between them show feeding relationships. The Sanak Island food web has 171 species (http://www.newswise.com/articles/research-examines-ancient-humans-as-major-predators-in-marine-food-webs-suggesting-lessons-for-sustainability).

4.2.2 The robustness of food webs

Food webs have been central to ecological research for decades (Cattin et al. 2004; Jordan et al. 2003; May 1972; Mc Cann 2000; Montoya and Solé 2003), and the study of the robustness of food webs to species loss is increasingly relevant for species and ecosystem conservation (Montoya et al. 2006; Raffaelli 2004; Zavaleta 2004).

The loss of a species in ecosystems (primary extinction) can cascade into further extinctions (secondary extinctions), as consumers' persistence depends on the persistence of their resources. Many theoretical and empirical studies have investigated how food web properties, such as modularity, degree-distribution (i.e. the probability distribution of the number of trophic connections of species), presence and distribution of keystone species may influence the pattern of secondary extinctions in ecosystems as well as food web robustness (Allesina and Pascual 2009; Bascompte et al. 2005; Dunne et al. 2002; Jordan et al. 2003; Solé and Montoya 2001). In the vast majority of studies on extinction patterns in food webs, a species is assumed to go extinct after a primary extinction when is left without any resources to exploit (Allesina and Bodini 2004; Allesina and Pascual 2009; Dunne et al. 2002; Solé and Montoya 2001). This is clearly the best-case

scenario (Allesina and Pascual 2009; Dunne 2006), as the occurrence of other common effects, such as size-dependent-dynamics, top-down cascades or energetic thresholds, would result in additional losses (Bellingeri and Bodini 2013; Curtsdotter et al. 2011; Dunne 2006).

4.2.3 Simulations studies of cascading extinctions: attack and error removal

Simulation studies have shown that the extinction of highly-connected species is likely to generate a greater number of secondary extinctions than when species are randomly removed from the food web (Allesina and Bodini 2004; Dunne et al. 2002; Dunne and Williams 2009; Solé and Montoya 2001). Notions of error and attack sensitivity were first introduced in the physical literature and then successfully applied to the study of food webs (Albert and Barabasi 2002; Dunne et al. 2002; Solé and Montoya 2001; Strogatz 2001). A network is error resistant (or resistant to failure) when it is unlikely to be damaged by random removal of nodes. On the other hand, a network is sensitive to attack when it can be either highly damaged or destroyed by a targeted attack, such as the selective removal of highly-connected nodes (Albert and Barabasi 2002; Dunne et al. 2002).

4.2.4 New models of cascading extinctions: removal with probability

The sequential removal from most- to least-connected species (intentional attack) and random extinction of species (random removal) are two limit criteria for determining primary extinctions in food webs (least- to most-connected should be the other limit criterion, but it is rarely used in practice), and both approaches have been widely used to study patterns of secondary extinctions in ecosystems as well as to measure food web robustness. However, it is possible to introduce other removal criteria along the continuum from the random removal of species to the intentional attack. Across ecosystems, certain species - not necessarily the most connected - can be more prone to extinction, either because preferentially targeted by natural or human agents (e.g. pollution, species invasion, overexploitation, weather extremes) or for internal dynamics or properties of the

biological community (e.g. size-dependent dynamics). Other factors can decrease the species risk of extinction, e.g. the ability of consumers to use or prey on other resources in the case of resource loss (i.e. "rewiring of the food web"), or the human conservation efforts. In this context, a valuable approach to primary species extinction in food webs is to introduce non-uniform and non-deterministic criteria for species extinction. The introduction of probabilistic approaches to species extinction may offer more realistic predictions of both primary and secondary extinction dynamics in food webs as well as insights on possible transitions in system behaviour (e.g. from robustness to fragility). Further, a probabilistic approach can help understand how changes in the primary extinction risk of species affect secondary extinctions in ecosystem.

In a recent work, Gallos et al. (2006) studied the robustness of scale-free networks, i.e. networks whose degree-distribution follows a power law. They used the probability $W(k) \sim k^{\alpha}$ for a node of degree k (i.e. number of links of the node) to become inactive, where for: (i) $\alpha = 0$ the removal is random; (ii) $\alpha < 0$ low-degree nodes are more vulnerable; (iii) $\alpha > 0$ high-degree nodes are more likely to be removed than low-degree nodes. Gallos et al. (2006) showed that a little increase of α strongly reduces the percolation threshold p_c. In other words, with a moderate increase of the probability of removing highly-connected nodes, the scale-free network is quickly destroyed following the inactivation of a small number of nodes.

So far, how network robustness changes when increasing the probability of removing highly-connected nodes has not been studied either in model or empirical food webs. Here, we analyze the robustness of 14 empirical food webs to node loss by introducing a parameter I (intentionality) that defines the probability of removing highly-connected species. When I increases, so does the extinction risk of highly-connected species.

4.3 Material and methods

4.3.1 The empirical food webs data set

A food web can be described as a directed network with *S* species (nodes) and *L* trophic interactions among them (links), describing who eats whom (Dunne 2006; Montoya et al. 2006). In this work, we used empirical food webs that represent a wide range of species numbers, link densities, taxa, habitat types (terrestrial, aquatic an transition ecosystems). In Table 1, we report the basic properties of each food web, such as number of species (*S*), average number of links per species (*L/S*), and connectance ($C=L/S^2$). Since S^2 is the maximum possible number of trophic interactions in a *S* x *S* matrix, food web connectance describes the realized fraction of trophic interactions in the food web.

Table 1: Main features of food webs used in this study. *L*, total number of links in the food web; *S*, number of species; *C*, food web connectance (L/S^2). Refs: literature reference for the food web. Keys: short id of food web.

Food web	***S***	$C=L/S^2$	***L/S***	**Refs**	**Key**
Bridge Brook Lake	25	0.171	4.28	Havens, 2002	*Br*
Coachella Valley	29	0.312	9.03	Polis, 1991	*Co*
Cheasepeake Bay	31	0.071	2.19	Baird and Ulanowicz, 1989	*Ch*
St Martin Island	42	0.116	4.88	Goldwasser and Roughgarden, 1998	*SM*
St Marks Seagrass	48	0.096	4.60	Christian and Luczkovich, 1999	*SMk*
Grassland	61	0.026	1.59	Martinez et al. 1999	*Gr*
Ythan Estuary 91	83	0.057	4.76	Hall and Raffaelli, 1991	*Y91*
Scotch Broom	85	0.031	2.62	Memmot et al., 2000	*Sc*
Stony Stream	109	0.07	2.19	Townsend et al., 1996	*St*
Little Rock Lake	92	0.118	10.84	Martinez, 1999	*Li*
Canton Creek	102	0.067	6.83	Townsend et al., 1996	*Ca*
Ythan Estuary 96	124	0.038	4.76	Huxham et al., 1996;	*Y96*
El Verde Rainforest	155	0.063	9.74	Waide and Reagan, 1996	*El*
Mirror Lake	172	0.146	25.13	Dunne et al., 2002	*Mi*

4.3.2 Robustness

Food web robustness is usually tested with simulations in which a single species is removed at each step (i.e. primary extinction), and the number of secondary extinctions (i.e. extinctions following the primary extinction) is recorded (Allesina and Pascual, 2009; Dunne et al. 2002; Dunne and Williams 2009; Solé and Montoya 2001). Species going primarily extinct may be selected according to a particular criterion (i.e. random removal, decreasing or increasing number of connections etc..), and primary extinctions are repeated until all the species have gone extinct. With a topological approach (i.e. based on presence/absence or links, with no information on interaction

strength), a network node goes extinct when it loses all incoming links. In food webs, that means a species goes extinct when it is left without any exploitable resources.

Here, we test the robustness of 14 empirical food webs (Table 1) by introducing a novel criterion for primary extinctions. We assume that consumers cannot switch from one type of prey to another (i.e. no food web "rewiring"). Several measures of food web robustness have been proposed, such as secondary extinction area (Allesina and Pascual 2009), error and attack sensitivity (Allesina and Bodini 2004; Allesina et al. 2006), R_{25} (Srinivasan et al. 2007). In this work, we use 'structural robustness' (R), that is the proportion of primary extinctions leading to a particular proportion of total extinctions (Curtsdotter et al. 2011; Dunne et al. 2002; Dunne and Williams 2009; Dunne, 2006):

$$R_{\alpha} = \frac{E}{S} \qquad (1)$$

where E is the number of primary extinctions that produces a percentage α of total extinctions (primary + secondary) out of the total number of species S in the food web. We used two measures of R: (i) the proportion of primary extinctions triggering the loss of half of the species (R_{50}) (Curtsdotter et al. 2001; Dunne et al. 2002; Dunne and Williams 2009; Dunne, 2006) , and (ii) the proportion of primary extinctions causing food web collapse (i.e. extinction of all species, R_{100}) (Dunne 2006; Ebenman 2011). The maximum possible value of robustness when using R_{50} is 0.5 (i.e. half of the species must be removed to trigger the loss of half of the species in the food web), while the minimum is $1/S$ (i.e. the extinction of one species leads to the extinction of half of the species in the food web). For R_{100} , maximum and minimum values of robustness are 1 and $1/S$, respectively.

4.3.3 Attack strategies

We introduce a novel attack strategy where the extinction (removal) probability of species is determined by a probability mass function. The total number of trophic interactions *k* of a species in a food web (i.e. degree of the node/species) is the sum of the number of the ingoing links (resources or prey) and the number of the outgoings links (consumers or predators).

We used two different probability mass functions to define the removal probability of species in a food web, namely the exponential and the power-law probability mass functions.

In the first case, the probability of removing a species with *k* trophic interactions with intentionality *I*, $P_E(K = k|I)$, is defined by the family of exponential probability mass functions:

$$P_E(K = k \mid I) = \frac{(1-I)^{(k_{\max}-k)} N_k}{\sum_{i=k_{\min}}^{k_{\max}} (1-I)^{(k_{\max}-i)} N_i} \qquad 0 \le I < 1 \tag{2}$$

where k_{max} is the maximum number of trophic interactions for a species, k_{min} the minimum number, N_k the number of species with degree *k*. The subscript *E* in P_E specifies the exponential probability mass function. From now on, we simply use $P_E(k|I)$ in order to simplify notation. When $I \rightarrow 1$, we tend to sequentially remove the most connected species (intentional attack), where:

$$P_E(k \mid 1) \equiv \lim_{I \to 1} P(k \mid I) = \delta k, k_{\max} \tag{3}$$

When $I = 0$, species are randomly removed, i.e. all nodes share the same probability of being removed:

$$P_E(k \mid 0) = N_k / N_{tot} \tag{4}$$

where N_{tot} indicates the total number of nodes in the network. In the second case, the removal probability $P_P(k|I)$ of nodes is defined by the power-law probability mass function:

$$P_P(k \mid I) = \frac{k_i{}^I}{\sum_{i=1}^{N_{tot}} k_i{}^I} \qquad 0 \le I < \infty \tag{5}$$

where k_i indicates the degree of node *i*, the exponent *I* is the intentionality parameter (corresponding to parameter α in Gallos et al. (2006)) and the subscript *P* in P_P indicates the power law. With the power-law formulation, the probability of removing highly-connected species increases with *I*, where for $I = 0$ species are randomly removed, and with $I \to \infty$ nodes are removed from most- to least-connected. We chose the power-law probability mass function in order to compare robustness of food webs to that of scale-free networks in Gallos et al. (2006).

In addition, as a third scenario we removed species from the most- to the least-connected (i.e. intentional attack). The degree *k* is recalculated with each new primary extinction. In the case of ties, i.e. nodes with the same degree, we randomly ordered those nodes.

Since the result of simulations using Eq. (2), Eq. (5) and with the intentional attack are stochastic realizations, for each value of *I* and each food web we carried out 1000 simulations. We used the mean across replicates as our measure of robustness for both R_{50} ($\bar{R}_{50}$) and R_{100} ($\bar{R}_{100}$). We could not directly compare the results obtained with the two family of functions as we had to use different sets of values of *I* for the power-law and exponential probability mass functions. For the exponential removal probability function (Eq. (2)), we used *I* values obtained by bisections from $I \sim 1$ to $I=0$ (I = 0, 0.00098, 0.00196, 0.00390625, 0.0078125, 0.015625, 0.03125, 0.0625, 0.125, 0.25, 0.5, 0.9999). We used the following bisections in order to analyse in greater details the increase of the removal probability in the highest degrees region.

For the power-law removal probability function (Eq. (5)), we used the same set of *I* values used by Gallos et al. (2006) (I = 0, 0.25, 0.5, 1, 2, 4) in order to directly compare the response of scale-free networks presented in Gallos et al. (2006) to that of food webs.

4.3.4 The breakpoint threshold pattern

A visual inspection of plots of robustness *R* ($\bar{R}_{50}$ and $\bar{R}_{100}$) *vs. I* when using the exponential function in Eq. (2) showed that *R* was fairly constant for increasing *I* up to a threshold value after which it

sharply declined with further increases of *I*. To fit these trajectories, we used two-phase regression models, that is regression models in which two straight lines are connected at a breakpoint I_t, in the form:

$$\begin{cases} R = \alpha_1 + \beta_1 I & for \;\; I < I_t \\ R = \alpha_2 + \beta_2 I & for \;\; I \geq I_t \end{cases}$$

with the restriction for continuity $\alpha_1 + b_1 I_t = \alpha_2 + \beta_2 I_t$. We fitted the two-phase regressions using the library *segmented* available for *R* (R Development Core Team 2011). Parameters estimation proceeds in two parts: a generalized linear model (GLM) is first fitted to the data, then a broken-line relationship (estimation of slopes and breakpoint) is added by re-fitting the model. We used Davies test for significant difference-in-slope (Davies 1987). We set statistical significance at the 0.05 level. We carried out all simulations and statistical analyses using *R* 2.14.0 (R Development Core Team 2011).

4.3.5 Robustness-complexity relationship

We used linear regressions on both linear and log-log scales to explore the relationship between I_t and two parameters describing food web complexity, namely species richness (*S*) and connectance ($C = L/S^2$) (Table 1). We used AIC to select the best model (we corrected the likelihood when the response variable was log-transformed).

4.4 Results.

4.4.1 Robustness with exponential probability of primary extinction

Using the exponential probability mass function, a value of intentionality close to 1 (i.e. close to the maximum value allowed by the probability mass function) was necessary across food webs to

obtain robustness values for both $\bar{R}_{50}$ and $\bar{R}_{100}$ comparable to those obtained with the intentional attack (Figs. 11 and 12).

For increasing *I*, we observed a slow-to-fast decrease in robustness after reaching a threshold value. The Davies test for difference-in-slope was significant for each food web and for both $\bar{R}_{50}$ and $\bar{R}_{100}$ (Table 2). For all the food webs, the slope of the regression line for values of intentionality $I > I_t$ was on average an order of magnitude greater than the slope of the regression line for $I < I_t$ (Table 2). Only in Mirror lake food web (for $\bar{R}_{50}$) a sharp decrease in robustness was not observed. Interestingly, some food webs showed an increase in robustness with increasing intentionality before the sharp decline in robustness for values of *I* greater than I_t.

4.4.2 Robustness with power law probability of primary extinction

For both measures of robustness, increasing intentionality generally decreased robustness, although for two food webs (Bridge for $\bar{R}_{50}$ and Coachella for $\bar{R}_{100}$) robustness tended to increase for values of *I* up to 2. Across food webs, robustness for $I = 1$ in Eq. (5) was substantially greater than robustness obtained with the intentional attack for both $\bar{R}_{50}$ and $\bar{R}_{100}$ (Fig. 13, Fig. 14). Only when setting $I = 4$ in Eq. (5), and only for some food webs, we obtained values of robustness similar to the one given by the intentional attack.

Table 2: Two-phase linear regression of the robustness measures ($\bar{R}_{50}$ and $\bar{R}_{100}$) on intentionality *I* for the exponential probability function. The breakpoint I_t indicates the value of the intentionality at which the transition of system response occurred (i.e. slow-to-fast decrease in robustness). b_1 and b_2 are the slopes of the straight lines on the left and on the right of I_t, respectively, while a_1 and a_2 are the respective intercepts. We present standard errors for all parameters estimates except for a_2, since it was calculated and not estimated in the two-phase regression. *p*-values of the Davies test for difference-in-slope are all smaller than 0.01 except for $\bar{R}_{50}$ for Mirror ($p = 0.078$).

Table 2, cont'd

Food Web	Breakpoint I_t	β_1	β_2	α_1	α_2
$\bar{R}_{50}$					
Bridge Brook Lake	0.9±0.01	0.0021±0.00034	-0.03±0.0009	0.463±0.011	0.8
Coachella Valley	0.163±0.002	0.00017±0.0003	-0.0244±0.002	0.4769±0.002	0.7055
Cheasepeake Bay	0.095±0.002	-0.0018±0.001	-0.04168±0.003	0.3551±0.006	0.6949
St Martin Island	0.194±0.03	-0.0023±0.0009	-0.0499±0.0049	0.4263±0.005	0.8807
St Marks Seagrass	0.1175±0.004	-0.000716±0.0008	-0.05483±0.003	0.4306±0.005	0.911
Grassland	0.17±0.05	-0.000301±0.0007	-0.0398±0.004	0.2377±0.004	0.606
Ythan Estuary 91	0.025±0.003	-0.00618±0.008	-0.0587±0.012	0.3812±0.003	0.7289
Scotch Broom	0.022±0.0002	-0.00553±0.006	-0.0527±0.008	0.3158±0.02	0.6153
Stony Stream	0.027±0.001	-0.0012±0.001	-0.0156±0.001	0.4918±0.003	0.5883
Little Rock Lake	0.011±0.002	0.00186±0.0043	-0.0236±0.005	0.4006±0.014	0.537
Canton Creek	0.0115±0.013	0.00014±0.001	-0.01081±0.001	0.4787±0.003	0.5386
Ythan Estuary 96	0.015±0.002	-0.0029±0.008	-0.0534±0.01	0.3717±0.03	0.6695
El Verde Rainforest	0.026±0.003	-0.00187±0.002	-0.03421±0.002	0.4394±0.01	0.6551
Mirror Lake	0.011±0.006	-0.00124±0.001	-0.00408±0.001	0.4665±0.004	0.4818
$\bar{R}_{100}$					
Bridge Brook Lake	0.34± 0.008	0.0004±0.0007	-0.054±0.00375	0.8213±0.004	1.345
Coachella Valley	0.18±0.001	0.0041±0.0013	-0.081±0.0074	0.7501±0.008	1.571
Cheasepeake Bay	0.072±0.0002	-0.0009±0.0012	-0.053±0.0023	0.660±0.0053	1.078
St Martin Island	0.17±0.0009	-0.0042±0.0012	-0.094±0.007	0.7496±0.007	1.578
St Marks Seagrass	0.096±0.001	-0.003977±0.002	-0.0816±0.0042	0.761±0.007	1.424
Grassland	0.11±0.009	-0.00113±0.001	-0.0483±0.003	0.4754±0.005	0.8879
Ythan Estuary 91	0.0122±0.03	0.0011±0.012	-0.0794±0.014	0.6058±0.039	1.054
Scotch Broom	0.0123±0.002	-0.0059±0.011	-0.0688±0.013	0.4500±0.037	0.8009
Stony Stream	0.027±0.01	-0.004848±0.0034	-0.0765±0.0048	0.888±0.013	1.37
Little Rock Lake	0.026±0.009	-0.00245±0.0021	-0.049±0.003	0.7283±0.008	1.042
Canton Creek	0.045±0.04	-0.0066±0.0038	-0.0986±0.008	0.8913±0.017	1.576
Ythan Estuary 96	0.0135±0.01	-0.0043±0.011	-0.082±0.013	0.6168±0.036	1.062
El Verde Rainforest	0.02±0.0003	-0.0059±0.004	-0.0598±0.0051	0.7963±0.014	1.134
Mirror Lake	0.038±0.004	-0.0038±0.0022	-0.0619±0.0041	0.9202±0.001	1.34

4.4.3 Relationship between breakpoint and complexity

Food web connectance *C* showed no statistically-significant relationship with breakpoint I_t for $\bar{R}_{50}$ with variables either on either linear or log-log scale, while we found a negative relationship

between C and I_t for $\bar{R}_{100}$ on both scales (p<0.01, Table 3). We found a negative linear relationship between species richness and I_t on both linear and log-log scales (p<0.01, Table 3). AIC strongly indicated the model on the linear scale as the best one. We found a negative linear relationship between S and the breakpoint of the two-phase regression I_t for $\bar{R}_{100}$ (p<0.01), and also in this case AIC was lower for the model on the linear scale (Table 3).

Table 3: Parameters of the regression models on the linear scale for the relationship between I_t and a) connectance, b) species richness.

	Breakpoint $\bar{R}_{50}$			**Breakpoint $\bar{R}_{100}$**		
	Slope	***Intercept***	***p***	***Slope***	***Intercept***	***p***
Connectance	1.1412±0.832	0.0150±0.102	0.20	0.7262±0.291	0.0113±0.035	<0.01
Species richness	-0.0027±0.001	0.352066±0.11	<0.05	-0.0013±0.0004	0.1953±0.04	<0.01

4.5 Discussion of the results

4.5.1 The sharp transition in the number of secondary extinction

Our study shows that when increasing the probability of deleting highly-connected species there is a sharp transition in system behaviour, from a region where food webs show high resistance to species loss ('*random removal regime*') to a region where robustness decreases rapidly and quickly reaches the robustness obtained with the sequential attack from most- to least-connected species ('*intentional attack regime*'). This pattern was particular clear when using the exponential probability removal function, although a fast decrease in robustness with increasing intentionality

was observed in the majority of food webs when using the power-law probability mass function, in particular for $\bar{R}_{100}$.

For $\bar{R}_{100}$, we found that connectance increased the value of intentionality at which the transition between the two regimes of robustness occurred (i.e. breakpoint of the two-phase regression) (Table 3). This result suggests that food webs with greater connectance are less affected than low-connectance food webs by an increase of the extinction risk of highly-connected species. This result is in agreement with previously investigations showing an increase of food web robustness with connectance (Dunne et al. 2002; Dunne et al. 2004; Dunne and Williams 2009). The regime transition at a larger value of intentionality that we found in our analyses for food webs with higher connectance may have two different explanations. First, it may be related to the buffer provided by an high number of trophic connections against further extinctions in the event of species loss (Dunne et al. 2002). Second, it may be explained by the degree-distribution of food webs, which typically changes from distributions similar to power-law to exponential or uniform with increasing connectance (Dunne 2006; Dunne et al. 2002; Montoya and Solé 2003). In fact, in food webs with highly-skewed degree-distribution, highly-connected species are more likely to function as "hubs" and their extinction may have dramatic effects on the stability of ecosystems, while in food webs with a more uniform degree-distribution the extinction of highly-connected species leads to a lower number of secondary extinction, thus preserving food web stability.

However, when using $\bar{R}_{50}$ as a measure of robustness, we did not observe a significant relationship between connectance and robustness, even after removing a clear outlier.

4.5.2 The complexity-stability relationship

On the contrary, for both robustness measures we observed a negative linear relationship between the breakpoint of the two-phase regression and species richness (*S*). In empirical food webs, no relationship is typically found between robustness and species richness (Dunne et al. 2002; Dunne

et al. 2004), whereas in model food webs species richness increases robustness (Dunne and Williams 2009). In our extinction scenarios, larger food webs seem to be more sensitive to the increase of intentionality (i.e. to the preferential targeting of highly-connected species). Also this pattern may be explained by variations in the shape of the degree-distribution, as its skewness tends to increase with species richness (Montoya and Solé 2003). In terms of conservation ecology, this result suggests that protecting highly-connected species may be more important in larger ecosystems. In fact, a smaller value of the breakpoint of the two-phase regression of robustness on intentionality (i.e., the transition from 'random removal' to 'intentional attack' regime) increases the probability of falling in the intentional attack regime in the case highly-connected are preferentially targeted. Since in larger food webs the transition in system behaviour occurred for lower values of intentionality, it follows that the preservation of highly-connected species may be particularly important for the stability of larger ecosystems.

In addition, this threshold effect strongly suggests that when modelling extinction dynamics to carefully assign or estimate a risk of primary extinction to species as a function of their number of trophic links. In fact, intentionality values slightly smaller or bigger than the breakpoint of the two-phase regression may lead to substantially different patterns of secondary extinction, as well as largely different estimates of food web robustness.

When using the power-law function to define the extinction probability of highly-connected species, patterns of robustness of food webs differed from those showed by Gallos et al. (2006) for scale-free networks. Scale-free networks are typically highly robust to random removal of nodes, but become fragile when highly-connected nodes are removed. In scale-free networks, the attack with $I = 1$ in Eq. (5) can reduce the percolation threshold in scale free network to $p_c = 0.25$, from $p_c = 1$ when nodes are randomly removed ($I = 0$) (Gallos et al. 2006). Contrary to what found for scale-free networks, a small increase in the extinction risk of highly-connected species does not strongly reduce the robustness of food webs. This is likely to be ascribed to the structural

differences between scale-free networks and food webs. First, the number of nodes N is much smaller in food webs (typically < 200, in our study $9 < N < 140$) than in scale-free networks (> 1000) (Camacho et al. 2002; Dunne 2006). Second, food webs exhibit smaller maximum node degree, and the degree-distribution in food webs is in general less skewed than a power-law (Camacho et al. 2002; Dunne 2006). Clearly, with respect to a scale-free network, a smaller maximum degree of the food web along with a less skewed degree-distribution would reduce the probability of attacking highly-connected nodes with an increase of intentionality. Since scale free networks did not show a threshold response to the increase in the intentionality parameter (Gallos et al. 2006), the emergence of the breakpoint not does not seem to be a general occurrence in all networks.

An interesting question is how the addition of ecological dynamics may modify the results presented here. In addition, the food webs we analysed in this work are binary, i.e. they describe only the presence of trophic interaction and do not describe the amount of energy passing from resource to consumer (i.e. interaction strength). Thus, it would be interesting to use our methodology with weighted food webs, that is food webs including information about the amount of the energy and matter passing along a trophic interaction (Bellingeri and Bodini 2013; Bodini et al. 2009; Thierry et al. 2011). Finally, the same approach we used in the present work could be applied to food webs where rewiring (i.e., modification of trophic interactions) may occur. Rewiring in the food web may simply occur when a predator consumes prey species not included in the trophic data. However, in modern data sets it is unlikely that potential resources resulting from switching prey go unregistered (Allesina and Pascual 2009). Alternatively, food web rewiring may occur when following the extinction of one of its competitors (competitive release), a consumer might expand its diet to include a prey that it had previously been not available (Staniczenko et al. 2010; Thierry et al. 2011). Finally, a consumer species forages on the subset of possible prey items that provides it with the highest net energy intake per unit effort (optimal foraging strategy).

Following the loss of its preferred prey item(s), a predator may expand its diet to include novel resources (Thierry et al. 2011). All the above processes can thus be potentially included in the analysis of food webs and exploring how food web robustness changes with increasing intentionality when modification of trophic interaction may occur is thus encouraged.

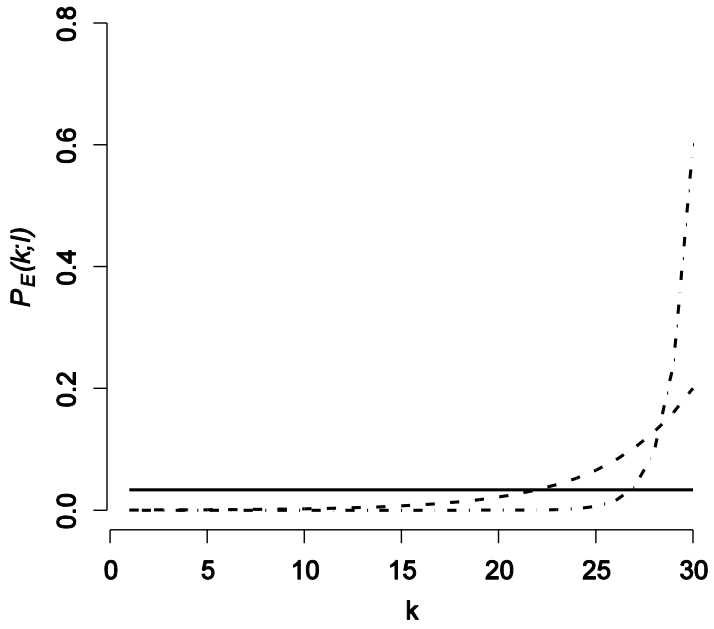

Figure 10: Removal probability $P_E(k|I)$ of a species with k trophic interactions for three value of intentionality I (solid line, $I = 0$; dashed line, $I = 0.1$; point-dashed line, $I = 0.2$) in Eq. (2.2) (exponential probability mass function) for St Marks food web. The solid horizontal line represents the random removal extinction scenario. The probability of removing highly-connected species increases with I.

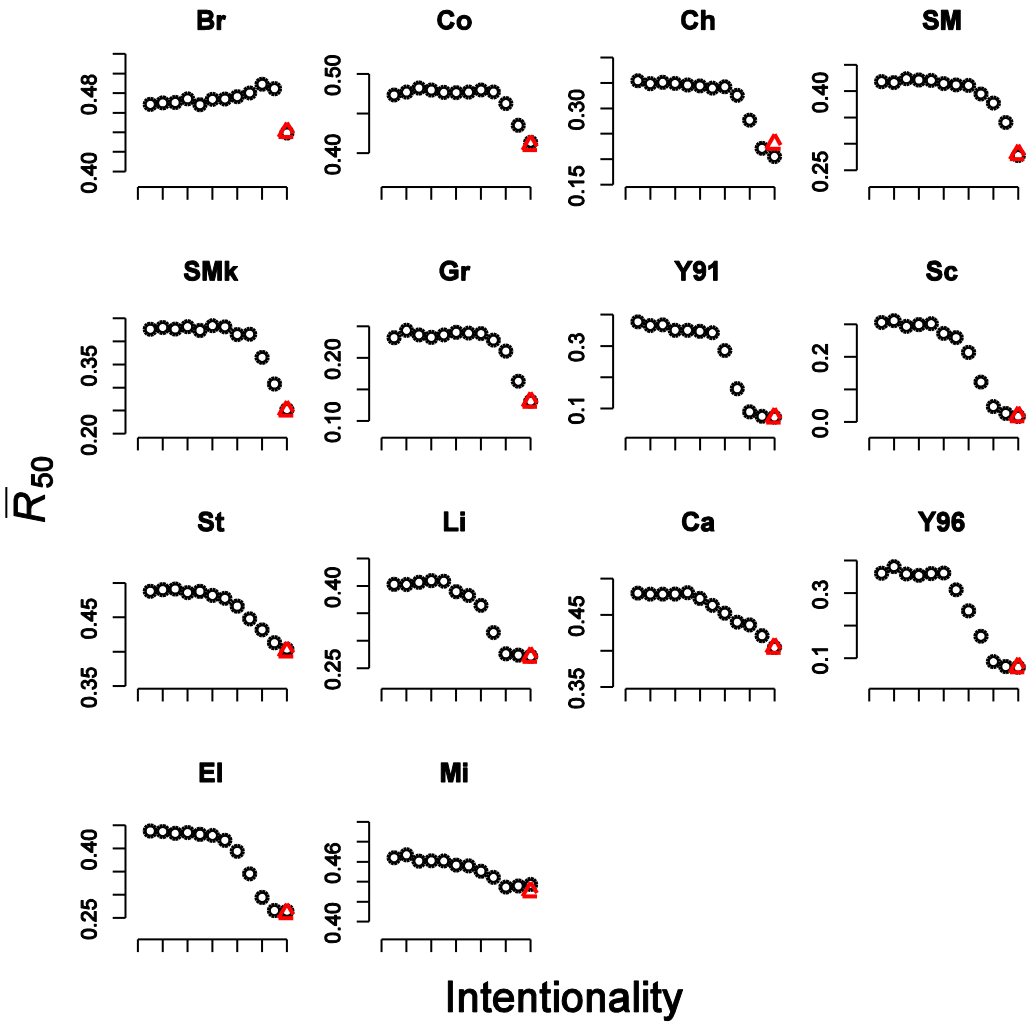

Figure 11: Robustness $\overline{R}_{50}$ as a function of intentionality I when using the exponential function in Eq. (2) for each of the 14 food webs we analysed. Note that the y axis is different for each food web in order to facilitate the visual analysis of patterns. ○ represent results using the exponential probability function; Δ (in red) indicates $\overline{R}_{50}$ for the intentional attack, from most- to least-connected species. Almost food webs show a slow decrease of $\overline{R}_{50}$ when I increases (Bridge, Coachella, Little Rock and Canton Creek food webs show a little increase) and then a sharp decrease after a critical value of I. The key on each panel identifies the food web as reported in Table 1. Left to right I = 0, 0.00098, 0.00196, 0.00390625, 0.0078125, 0.015625, 0.03125, 0.0625, 0.125, 0.25, 0.5, 0.9999.

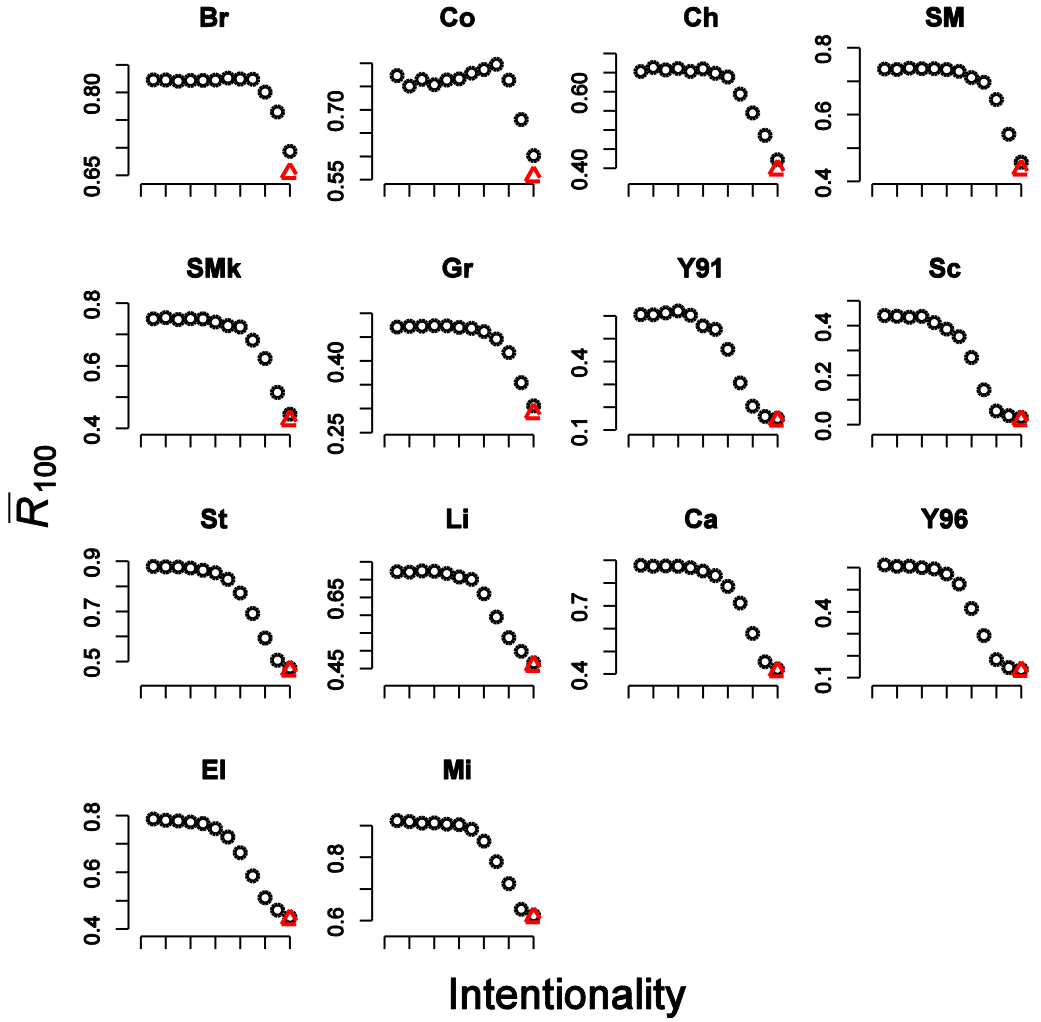

Figure 12: Robustness $\overline{R}_{100}$ as a function of intentionality I when using the exponential function in Eq. (2) for each of the 14 food webs we analysed. Note that the y axis is different for each food web in order to facilitate the visual analysis of patterns. Δ (in red) indicates $\overline{R}_{100}$ for the intentional attack, from most- to least-connected species. For the exponential function, almost all food webs show a slow decrease of $\overline{R}_{100}$ when the intentionality of the removal criterion is increased (Bridge, Coachella and YThan Estuary 91 food webs show a slight increase of robustness with increasing *I*) and then a sharp decrease after a critical value of intentionality is reached. The key on each panel identifies the food web as reported in Table 1. Left to right I = 0, 0.00098, 0.00196, 0.00390625, 0.0078125, 0.015625, 0.03125, 0.0625, 0.125, 0.25, 0.5, 0.9999.

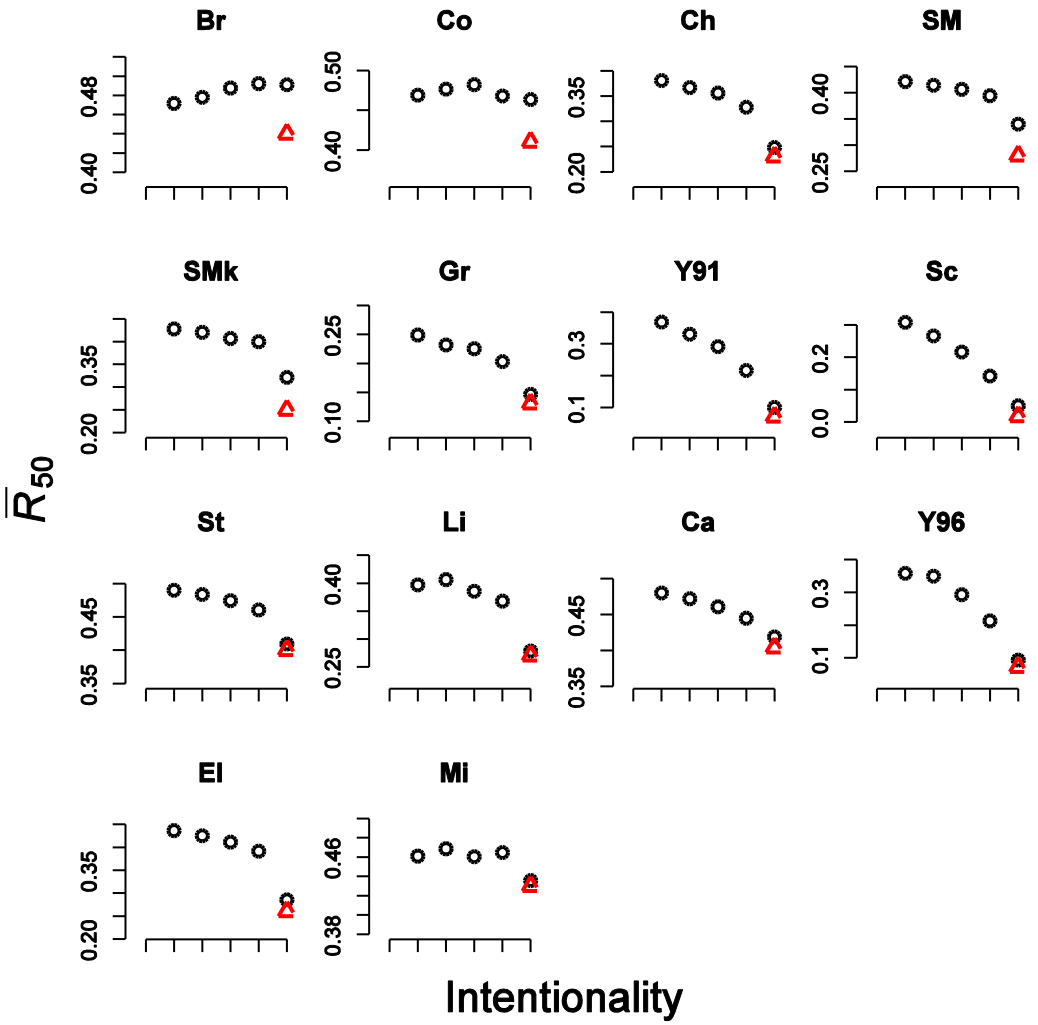

Figure 13: Robustness $\overline{R}_{50}$ as a function of intentionality *I* when using the power law function in Eq. (5) for each of the 14 food webs we analysed. Δ (in red) indicates $\overline{R}_{50}$ for the intentional attack, that is with sequential primary extinctions from the most- to the least-connected species. Left to right *I* = 0, 0.25, 0.5, 1, 2, 4.

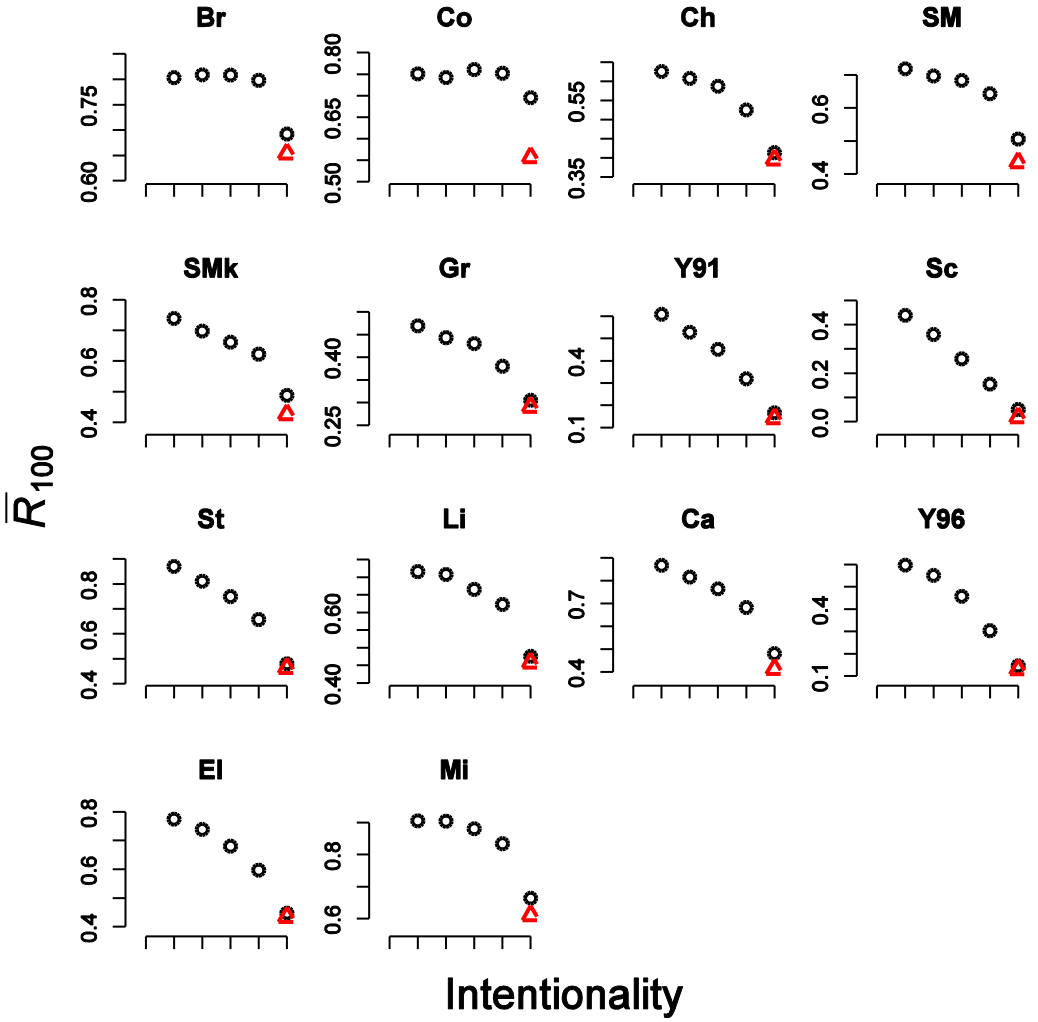

Figure 14: Robustness $\overline{R}_{100}$ as a function of intentionality *I* for the power law function in Eq. (5) for each of the 14 food webs we analysed. Δ (in red) indicates $\overline{R}_{100}$ for the intentional attack, that is with sequential primary extinctions from the most- to the least-connected species. Left to right *I* = 0, 0.25, 0.5, 1, 2, 4.

5. Acknowledgements

We thank Prof. Davide Cassi for the fundamental teaching in complex networks theory. Prof. Simoni Vincenzi, Dr. Elena Agliari and Prof. F. Scotognella for many discussions and fruitful exchanges. We thank Dr. Stefano Poletti.

6.References

Agliari, E., Barra, A., 2011. A Hebbian approach to complex-network generation. EPL (Europhysics Lett. 94, 10002. doi:10.1209/0295-5075/94/10002

Agliari, E., Barra, A., Galluzzi, A., Guerra, F., Moauro, F., 2012. Multitasking associative networks. Phys. Rev. Lett. 109, 268101.

Agliari, E., Cioli, C., Guadagnini, E., 2011. Percolation on correlated random networks. Phys. Rev. E 84, 031120.

Albert, R., Barabási, A., 2002. Statistical mechanics of complex networks. Rev. Mod. Phys. 74.

Allesina S., Bodini A., 2004. Who dominates whom in the ecosystem? Energy flow bottlenecks and cascading extinctions. J. Theor. Biol. 230, 351–358.

Allesina, S., Bodini, A., Bondavalli C., 2006. Secondary extinctions in ecological networks: bottlenecks unveiled. Ecol. Model. 194, 150–161

Allesina., S., Pascual M., 2009. Googling Food Webs: Can an Eigenvector Measure Species' Importance for Coextinctions? PLoS Comput Biol 5(9): e1000494.

Andserson, R.., May, R.M., 1992. Infectious Diseases of Humans: Dynamics and Control.

Applegate, D.L., Bixby, R.M., Chvátal, V., Cook, W.J., 2006. The Traveling Salesman Problem: A computational study.

Baird, D., Ulanowicz, R.E., 1989. The seasonal dynamics of the Chesapeake Bay ecosystem. Ecological Monograph 59, 329–364.

Bascompte, J., Melian, C.J., Sala, E., 2005. Interaction strength combinations and the overfishing. PNAS 102, 5443–5447.

Bellingeri, M., Cassi, D., Vincenzi, S., 2013. Increasing the extinction risk of highly connected species causes a sharp robust-to-fragile transition in empirical food webs. Ecol. Modell. 251, 1–8.

Bellingeri, M., Bodini, A., 2013. Threshold extinction in food webs. Theoretical Ecology. DOI: 10.1007/s12080-012-0166-0

Bellingeri, M., Cassi, D., Vincenzi, S., 2014. Efficiency of attack strategies on complex model and real-world networks. Physica A 414,174-180.

Bellingeri, M., Cassi, D., Agliari, E., 2015. Optimization strategies with resource scarcity: from immunization of networks to the traveling salesman problem. Physics Letter A, (Under Review).

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

Camacho, J., Guimerà, R., Nunes Amaral, L., 2002. Robust Patterns in Food Web Structure. Physical Review Letters 88, 8–11.

Cattin, M.-F., Bersier, L.-F., Banasek-Richter, C., Baltensperger, R., Gabriel, J.-P., 2004. Phylogenetic constraints and adaptation explain food-web structure. Nature 427, 835–9.

Chen, Y., Paul, G., Havlin, S., Liljeros, F., Stanley, H., 2008. Finding a Better Immunization Strategy. Phys. Rev. Lett. 101, 058701. doi:10.1103/PhysRevLett.101.058701

Christian, J.J., Luczkovich, R.., 1999. Organizing and understanding a winter's seagrass foodweb network through effective trophic levels. Ecological Modelling 117, 99–124.

Curtsdotter, A., Binzer, A., Brose, U., de Castro, F., Ebenman, B., Eklöf, A., Riede, J.O., Thierry, A., Rall, B.C., 2011. Robustness to secondary extinctions: Comparing trait-based sequential deletions in static and dynamic food webs. Basic and Applied Ecology 12, 571–580.

Davies, R., 1987. Hypothesis testing when a nuisance parameter is present only under the alternative - linear model case. Biometrika 33–43.

Dunne, J., Williams, R, Martinez, N., 2002a. Network structure and biodiversity loss in food webs: Robustness increases with connectance. Ecol Lett 5, 558–567.

Dunne, J.A., R.J., Williams, N.D., Martinez. 2002b. Food-web structure and network theory: the role of connectance and size. PNAS 99,12917-12922.

Dunne, J., Williams, R., Martinez, N., 2004. Network structure and robustness of marine food webs. Mar Ecol Prog Ser 273: 291–302.

Dunne, J.A., 2006. The network structure of food webs. 27-86 in Ecological Networks: Linking Structure to Dynamics in Food Webs. M. Pascual and J.A. Dunne, eds. Oxford University Press.

Dunne, J., Williams, R.J., 2009. Cascading extinctions and community collapse in model food webs. Philosophical transactions of the Royal Society of London. Series B, Biological sciences 364, 1711–23.

Dupuy, J.M., Freidel, L., 1990. Lag between discovery and production of new vaccines for the developing world. Lancet 336, 733–734.

Erdos, P., Renyi, A., 1960. On the evolution of random graphs. Publ. Math. Inst. Hung. Acad. Sci. 5, 17–60.

Ebenman, B., 2011. Response of ecosystems to realistic extinction sequences. The Journal of animal ecology 80, 307–9.

Gallos, L.K., Cohen, R., Argyrakis, P., Bunde, A., Havlin, S., 2006. Stability and topology of scale-free networks under attack and defense strategies. Phys. Rev. Lett. 94, 188701.

Goldwasser, L., Roughgarden, J., 1993. Construction and Analysis of a Large Caribbean Food Web. Ecology 74, 1216–1233.

Gfeller, R., 2007. Simplifying complex networks: from a clustering to a coarse graining strategyNo Title.

Guimerà, R., Danon, L., Díaz-Guilera, a., Giralt, F., Arenas, a., 2003. Self-similar community structure in a network of human interactions. Phys. Rev. E 68, 065103. doi:10.1103/PhysRevE.68.065103

Hadidjojo, J., Cheong, S.A., 2011. Equal graph partitioning on estimated infection network as an effective epidemic mitigation measure. PLoS One 6, e22124. doi:10.1371/journal.pone.0022124

Hall, A.S.J., Raffaelli, D., 1991. Food-Web Patterns : Lessons from a Species-Rich Web. Journal of Animal Ecology 60, 823–841.

Hasegawa, T., Naoki, M., 2011. of networks against propagating attacks under vaccination strategies. J. Stat. Mech. P09014.

Havens, K., 1992. Scale and Structure in Natural Food Webs. Science 257, 1107–1109.

Huang, X., Gao, J., Buldyrev, S. V., Havlin, S., Stanley, H.E., 2011. Robustness of interdependent networks under targeted attack. Phys. Rev. E 83, 065101. doi:10.1103/PhysRevE.83.065101

Huxham, M., Beany, S., Raffaelli, D., 1996- Do parasites reduce the chances of triangulation in a real food web? Oikos 76, 284–300.

Iyer, S., Killingback, T., Sundaram, B., Wang, Z., 2013. Attack robustness and centrality of complex networks. PLoS One 8, e59613. doi:10.1371/journal.pone.0059613

Jain, R.K., di Tomaso, E., Duda, D.G., Loeffler, J.S., Sorensen, a G., Batchelor, T.T., 2007. Angiogenesis in brain tumours. Nat. Rev. Neurosci. 8, 610–22. doi:10.1038/nrn2175

Jordán F., Scheuring I., Molnár I., 2003. Persistence and flow reliability in simple food webs. Ecol. Model. 161, 117–124.

Martinez, N., 1999. Artifacts or Attributes? Effects of Resolution on the Little Rock Lake Food Web. Ecological Monograph 61, 367–392.

Martinez, N., Hawkins, B.R.A., Dawah, H.A.A.L., Feifarek, P., 1999. Effects of sampling effort on characterization of food-web structure. Ecology 80, 1044–1055.

May, R., 1972. Will a large complex system be stable? Nature 238, 413–414.

McCann, K.S., 2000. The diversity-stability debate. Nature 405, 228–33.

Memmott, J., Martinez, N.D., Cohen, J.E., 2000. Predators, parasitoids and pathogens: species richness, trophic generality and body sizes in a natural food web. Journal of Animal Ecology 69, 1–15.

Montoya, J.M., Pimm, S.L., Solé, R.V., 2006. Ecological networks and their fragility. Nature 442, 259–64.

Montoya, J.M., Sole, R.V., 2003. Topological properties of food webs: from real data to community assembly models. Oikos 102, 614–622.

Nowak, M., May, R.M., 2000. Virus dynamics.

Pastor-Satorras, R., Vespignani, A., 2002. Immunization of complex networks. Phys. Rev. E 65, 036104. doi:10.1103/PhysRevE.65.036104

Polis, G., 1991. Complex trophic interactions in desert: An empirical critique of food web theory. American Naturalist 138, 123–155.

Raffaelli, D., 2004. Ecology. How extinction patterns affect ecosystems. Science (New York, N.Y.) 306, 1141–2.

R Development Core Team 2011. R: A language and environment for statistical computing, reference index version 2.15. R Foundation for Statistical Computing, Vienna, Austria.

Samuelsson, B., Socolar, J., 2006. Exhaustive percolation on random networks. Phys. Rev. E 74, 036113. doi:10.1103/PhysRevE.74.036113

Schneider, C.M., Mihaljev, T., Herrmann, H.J., 2012. Inverse targeting —An effective immunization strategy. EPL (Europhysics Lett. 98, 46002. doi:10.1209/0295-5075/98/46002

Schneider, C.M., Moreira, A.A., Andrade, S., Havlin, S., Herrmann, H.J., 2011. Onion-like network topology enhances robustness. J. Stat. Mech. Theory Exp. 1–4.

Schwalbe, N., El-Ziq, I., 2010. GAVI's Advance Market Commitment. Lancet 375, 638–639.

Serrano, M.Á., Boguñá, M., Vespignani, A., 2009. Extracting the multiscale backbone of complex. PNAS 106, 6483–6488.

Solé, R.V., Montoya, J.M., 2001. Complexity and fragility in ecological networks. Proceedings. Biological sciences / The Royal Society 268, 2039–45.

Srinivasan, U.T., Dunne, J. a, Harte, J., Martinez, N.D., 2007. Response of complex food webs to realistic extinction sequences. Ecology 88, 671–82.

Staniczenko, P. P. A., Lewis, O. T., Jones, N. S. & Reed-Tsochas, F., 2010. Structural dynamics and robustness of food webs. Ecology Letters, 13, 891–899.

Strogatz, S.H., 2001. Exploring complex networks. Nature 410, 268–76.

Thierry, A., Beckerman, A.P., Warren, P.H., Williams, R.J., Cole, A.J., Petchey, O.L., 2011. Adaptive foraging and the rewiring of size-structured food webs following extinctions. Basic and Applied Ecology 12, 562–570.

Townsend, C.R., Ross, M., Mcintosh, A.R., 1998. Disturbance, resource supply, and food-web architecture in streams. Ecology Letters 1, 200–209.

Waide, R., Reagan, D., 1996. The Food Web of a Tropical Rain Forest.

Welter, M., Rieger, H., 2013. Interstitial fluid flow and drug delivery in vascularized tumors: a computational model. PLoS One 8, e70395. doi:10.1371/journal.pone.0070395

Yu, S., 2014. Distributed Denial of Service Attack and Defense.

Zavaleta, E.S., 2004. Realistic Species Losses Disproportionately Reduce Grassland Resistance to Biological Invaders. Science 306, 1175–1177.

Zeng, A., Liu, W., 2012. Enhancing network robustness against malicious attacks. Phys. Rev. E 85, 066130. doi:10.1103/PhysRevE.85.066130