Evolution of topology and information transfer in networks of agents
Start date of the project: January 2004

Duration: 50 months

Project Coordinator: Prof. Dr. math. Friedhelm Meyer auf der Heide
Heinz Nixdorf Institute, University of Paderborn, Germany

Due date of deliverable: December 2007
Actual submission date: December 2007
Dissemination level: PU – public

Work Package 2.2: Dynamics and stability of networks

Participants: Università Degli Studi di Roma “La Sapienza” (UDRLS), Italy
Université de Paris-Sud (UPS), France

Authors of deliverable: Alain Barrat (Alain.Barrat@th.u-psud.fr)
Guido Caldarelli (Guido.Caldarelli@roma1.infn.it)
Alessandro Panconesi (Panconesi@di.uniroma1.it)
1 Introduction

We begin by describing our contributions to the study of security. In this report we deal with the problem of massive attacks against sensor networks. A massive attack is when the adversary is able to compromise a large fraction of the communication links of the network, by means of capturing nodes and thereby gaining control of the cryptographic material they contain. Since sensor networks are typically unattended, collecting nodes is considered to be a realistic way to tamper with their security (see [1] among others). Even though it may partly fall under control of the adversary, the network as a whole can continue to operate in a useful manner if the inflicted damage is confined to a small part. This is indeed the case in many applications of sensor networks, such as monitoring of the environment and of the infrastructure. In these cases, it can be much more important to protect the global functionality of the network than few individual communication links.

If the adversary has enough resources to capture a large fraction of the nodes, then there is not much that can be done to salvage the functionality of the network. Therefore, a meaningful question is to ask whether a massive attack can be effected cheaply, that is, by capturing just a small fraction of the nodes.

This is the problem we investigate. We prove in a mathematically rigorous fashion that the well-known random key pre-distribution scheme [2] makes a sensor network secure against any massive attack whatsoever. We establish security against massive attacks regardless of the strategy used to collect nodes. For instance, the adversary could obtain information on the network by means of traffic analysis or by exploiting weaknesses of the key-discovery protocol, and use this information to select an appropriate strategy and a small set of nodes. We show that it is possible to fix the relevant parameters for the security of the network—pool and key-ring size—in such a way that, with overwhelming probability and regardless of the information available to the adversary, the network is at the same time connected and secure against massive attacks. In other words, only “expensive” massive attacks can succeed.

In the literature, security against massive attacks, and security properties in general, are often investigated only experimentally. One of our main contributions is to show that they can be certified a-priori in a precise and rigorous way. We remark that the guarantees we provide are out of reach for experimental approaches.

Besides massive attacks, we analyse another kind of attack and certify random pre-distribution of keys against it. Informally, the question is whether the adversary can split the network into two large chunks and compromise all communications between them just by acquiring a small set of nodes. We show that this is not possible, again regardless of the information available to the adversary. Besides their intrinsic interest, we feel that our results introduce an important element of novelty at the methodological level in the large body of research on random pre-distribution of keys and, more generally, on pairwise-communication security. Furthermore, our formal framework and methodological approach is general enough to be applicable to other contexts, such as peer-to-peer networks.

We now describe our contribution in the area of co-evolution of networks. One of the possible explanation for the ubiquitous presence of power-laws and scale-invariant properties in technological networks is that they naturally arise for some form of self-organization of the system. In other words, these systems have an inherent tendency to increase their internal order without any planned design. Their importance for our project relies in the fact that this property could be used to realize more robust and reliable networks of communication. Indeed, the self-organization is usually triggered by internal variation processes, rather than external intervention. Therefore, the particular kind of
environment is no more crucial, and the system can maintain its properties in a variety of different external conditions. At its core the self-organization can be described by a series feedbacks both positive and negative in such a way to allow that some rearrangements tend to reinforce themselves as in the auto-catalytic reactions. Despite the potential importance of this explanation and despite the possible applications of such ideas, no systematic study of these adaptive network topologies is present and there are still many open questions [6, 7, 8, 9, 10, 11, 12]. For our purpose a network is adaptive if both the process on the network and the network topology are dynamical structures interacting with each other with complex feedback effects. Such an interaction is indeed expected in real-world networks. For example, it is a common observation that social networks are dynamic objects. People break and establish connections on a regular basis based on the behavior of the other person. Besides, they adjust their actions and behavior influenced by their peers. In order to understand the results of this dynamical aspect of the network, one has to assess the possible outcomes of these feedback effects.

In this report we present the first tentatives to characterize computationally and analytically the behaviour of simple statistical models with such characteristics. The idea is that from this simplified statistical model one could take inspiration in order to realize more complex devices based on the feedback between the vertices status and the set of connections. In all the known models either the network develops much faster than other vertices properties (no dynamics of the latter is considered), or the vertices properties undergo a dynamical process much faster than network evolution (the topology is kept fixed while studying the dynamics). This is because by separating the dynamical and topological time-scales, and by allowing the “fast variables” to evolve while the “slow variables” are fixed (quenched), makes the whole picture simpler and more tractable. However, this unavoidably implies that the slow variables must be treated as free parameters to be arbitrarily specified and, more importantly, that the feedback effects between topology and dynamics are neglected. From a different point of view we also study the effects of an adaptive topology on the formation of social consensus. To this purpose we have considered a particular model for opinion formation, namely the *Defiant model*, on adaptive networks[13]. In the Defiant model, opinions, the internal states of the agents, are continuous variables and neighbouring agents which have close enough opinions (as determined by a tolerance parameter) can reach a local consensus. In our model, the rate of evolution of the network’s topology is tunable and represents one of the parameters. We focus on simple evolution rules of the topology, that do not require prior knowledge of the state of agents to which new links are established. We study the role of the various parameters such as the tolerance of agents and the rate of topology evolution. The possibility of the interaction network to adapt to the changes in the opinion of the agents has important consequences on the system’s final state, the structure of this state, and the evolution mechanism.

The structure of this report is the following: the first part describes more in detail our results concerning security of ad-hoc networks. In the second part we firstly introduce the network model and we then present our analytical and numerical results for the various cases of study. Importantly, regardless the precise situation considered, we find that the networks and the functions defined on them self-organizes into a nontrivial state. A power-law decay of dynamical and topological quantities above a threshold emerges spontaneously, as well as a feedback between different dynamical regimes and the underlying network’s correlation and percolation properties.
2 The cost of a massive attack

We begin by reviewing briefly the random key pre-distribution model. We are given a network of \( n \) nodes and a pool of \( K \) cryptographic keys. Each node \( u \) is given a random subset \( K_u \) of \( k \) keys, its key ring. Two nodes in the network are connected by a link if and only if they are within transmission radius and share a key. The resulting graph is called a cryptograph. The basic question is how to set the values of \( K \) and \( k \) in order to have some desirable properties, such as various forms of connectivity and, simultaneously, security. It should be noted that security and connectivity are at odds. The former likes large key rings (relative to the pool size), since this increases the probability of having links. On the other hand, the latter likes them to be small, since in this fashion capturing one sensor is unlikely to corrupt many links. As shown in [3], if

\[
\frac{k^2}{K} \sim \frac{\log n}{n}
\]

then the network is connected with high probability, while it is likely to be disconnected below this threshold. Furthermore, the network is secure against a massive attack made by the so-called random adversary.

However we want to consider an extremely severe kind of massive attack [4]. The adversary selects a certain number of nodes with the aim of compromising links. Key rings of the selected nodes belong to the adversary. We assume pessimistically that a link \( uv \) is compromised as soon as the adversary owns a key in \( K_u \cap K_v \). The assumption is pessimistic because this key may or may not be used by \( u \) and \( v \). We also assume that the adversary has unlimited computing power. Crucially, we assume the following: the adversary knows the sets \( K_u \) for every node \( u \) in the network. Although its knowledge about the keys is complete, and its computing power is unlimited, the adversary has a limited budget. Denoting it with \( B \), and assuming that acquiring a node has unit cost, the question is: how many links can it compromise?

The point is that the adversary has the information and the computing power to cause maximum damage, within its limited budget. We call this the omniscient adversary. Our theorems show that in order to compromise a large (linear fraction) of the links, the omniscient adversary is forced with overwhelming probability to capture a large (linear fraction) of the nodes, i.e. the budget \( B \) must grow linearly with the number of nodes. Hence the term unassailable. This result holds assuming Condition 1. Therefore cryptographs are at the same time connected and secure against the omniscient adversary, with high probability. Our proofs show that, in essence, even if the adversary knows the composition of the key rings and operates its choices in the optimal (most disruptive) way, capturing a node essentially corrupts only the links incident on the captured node.

It might be objected that this adversary is unrealistically strong. This is however entirely desirable! If we are able to certify the security of the network against such a strong adversary it follows that it will be secure against more realistic attacks. For instance the result holds for an adversary that captures nodes blindly, at random. In fact, it implies that to compromise linearly many links, linearly many nodes are needed, no matter how the adversary selects the nodes. Also, assuming such a strong adversary has interesting practical consequences. In the literature, much attention has been given to the problem of finding protocols for key discovery between neighbours in such a way that no information is released to the adversary. One example is challenge response [2, 5], that is energy consuming (\( k \) messages and \( k^2 \) decryptions per peer to discover shared keys) but keeps key distribution as secret as possible. Our results on the omniscient adversary show that, if we are interested in protecting the network against massive attacks, we
do not (asymptotically) gain any advantage by using challenge response with respect to any light-weight, non-cryptographic protocol that does not keep key distribution secret. Broadcasting lists of key indices in the clear is good as well (within constant factors).

Interestingly, our results show that it is possible to build a secure sensor network that is very strong as a whole, while every individual link that composes the same network is extremely weak—it is not hard to see that the omniscient adversary can corrupt arbitrarily chosen individual links extremely easily, it is enough to select and corrupt any node in the network that contains the same key used to secure the target link.

We prove that kryptographs enjoy another strong security property. Suppose now that the aim of the adversary is to split the network into two large sets and to compromise all links between them, i.e. the adversary wants to partition the network. We show that, with overwhelming probability, the omniscient adversary cannot do this. This result is a consequence of the following structural property of kryptographs: with high probability and simultaneously for all large (linear size) sets of vertices, the number of edges between a set and its complement is \( \Theta(n \log n) \). This fact implies good fault-tolerant properties: to partition a network into two large sets, a huge (linear fraction) of the links must go down.

Finally, we extend our results to giant components. For many applications it suffices to have a giant connected component that covers the area within which the network is deployed. We show that if

\[
\frac{k^2}{K} \sim \frac{1}{n}
\]

then, with high probability, the network has a giant connected component, say, a connected component containing 99% of the vertices. Note that Condition 2 can be satisfied with much smaller key rings than Condition 1. This results in memory saving (smaller key ring) and computation saving (faster key discovery), an important consideration in resource starving environments typical of sensor networks. But more importantly, networks generated with Condition 2 are much sparser, not only globally but also locally at the neighbourhood level, than those set up under Condition 1. This is very beneficial because it translates in a reduced amount of interference, limiting the number of packet collisions and corresponding retransmissions. Nonetheless, we prove that the giant connected components are unassailable and unsplittable against the omniscient adversary as well, just like the whole kryptograph.

The properties we have been describing can be proven in the general case, but full proofs are given for the full-visibility case only. We made this choice because in this fashion the basic argument with the underlying reasons emerge more clearly, without cluttering technicalities. The general case is dealt with in the experiments. They complement and sharpen the predictions of our theorems rather than being the sole evidence of security.

To summarize, we believe that we have introduced an interesting approach from the methodological point of view that allows the rigorous investigation of security properties, not only for sensor networks. In particular, we have also shown that random key pre-distribution ensures surprisingly strong security properties against massive attacks.

3 Fitness or hidden variable models

One of the many definitions of complexity focus on the property that an ensemble of otherwise "simple" agents interact to form a common system where new global (and unexpected) properties emerge. It is then natural to think to networks as a prototype mathematical description for such systems. In this respect a model like that of the random graphs con-
sider all the vertices with the same properties. Conversely in many real situations, the vertices can represent agents or species or in general objects with a specific individuality. By assigning to every vertex \( i \) a quantity \( x_i \) that represents the specific properties of the agent we can therefore obtain a more careful representation of the system considered. This approach has proven to be rather successful in the past, and many efforts have been devoted in order to study the properties of networks obtained in such a way and if possible to apply the results of the model to real situations. We report below two activities made in this deliverable.

### 3.1 Self organization in a scale-free state

In the first model of network evolution the vertices of the network are characterised by a variable evolving through an extremal dynamics process. The network topology is in turn shaped by the variable itself. More specifically, to each vertex \( i \) a fitness \( x_i \) is assigned; the probability that a link exists between two vertices \( i, j \) is a function \( f(x_i, x_j) \) of their fitnesses[14]. Then, in the evolution, the vertex with minimum fitness and its neighbours are updated by extracting new fitnesses.

The fitnesses are initially \( (t = 0) \) drawn from a probability distribution \( \rho(x, 0) \) uniform between 0 and 1. The edges between the vertices are formed with various fitness dependent probability to be defined later.

A sketch of the modification in on time step of evolution is shown in Fig. 1. The fitness distribution \( \rho(x, t) \) self-organizes spontaneously to a stationary probability density \( \rho(x) \), that is

\[
\lim_{t \to \infty} \rho(x, t) = \hat{\rho}(x)
\]  

The analytical solution of the model for an arbitrary linking function \( f(x, y) \) can be obtained by focusing on the master equation for the fitness distribution \( \hat{\rho}(x) \) at the stationary state. We find that the analytical expression for \( \hat{\rho}(x) \) is[10]

\[
\hat{\rho}(x) = \begin{cases} 
(\tau N)^{-1} & x < \tau \\
\frac{1}{N \int_0^\tau f(x, m)dm} & x > \tau 
\end{cases}
\]  

where \( \tau \) is a threshold value determined through the normalization condition \( \int_0^\tau \hat{\rho}(x)dx \),

![Figure 1: On the left a graph at time \( t \). The black vertex has the lowest fitness, the grey ones are its neighbours. At next time step \( t + 1 \), three new vertices (light grey) are introduced with their new fitnesses, determining a new set of connections.](image)
which reads
\[ \int_{\tau}^{1} \int_{0}^{1} dx \frac{f(x, m)dm}{\tau} = N - 1 \]  
\[ (5) \]

In the infinite size limit \( N \to \infty \), the distribution \( q(m) \) of the minimum fitness value \( m \equiv x_{\min} \) is uniform between 0 and \( \tau \), while all other values (except possibly a vanishing fraction) are above \( \tau \). In other words, \( q(m) = \Theta(\tau - m)/\tau \). This characterizes the stationary state completely. Once \( \hat{\rho}(x) \) is known, all the expected topological quantities can be determined as in the static fitness model [14, 15, 16]. Note that if \( \tau \) is nonzero the model preserves the step-like behaviour displayed on static networks [17, 18, 19, 20]. However, here we find the novel feature that \( \hat{\rho}(x) \) is in general not uniform for \( x > \tau \). Therefore the system spontaneously evolves from a random structure to a complex network with nontrivial dynamical and topological properties. Also, note that the uniform distribution of \( m \) below \( \tau \) generalizes the result obtained for the random-neighbour variant of the BS model [18]. Indeed, the latter is a particular case of our model that we recover below. The above analytical solution holds for any form of \( f(x, y) \).

3.2 The dynamics of agents opinions

The same concept of fitness, it is useful to describe a specific case of social network, that of the opinion formation. We assume that the previous fitness \( x_i \) are the opinions of agent \( i \) and we study their evolution in time, by assigning a rule of agreement between different agents. In the original definition of the Defiant model [13], \( N \) agents are endowed with a continuous opinion \( o \in [0 : 1] \) [13, 21, 22, 23, 24]. Starting from random values, the agents’ opinions evolve through binary interactions according to the following rules: at each time-step \( t \), two neighbouring agents are chosen at random. If their opinions are close enough, i.e., if \(|o(i, t) - o(j, t)| \leq d\), where \( d \) defines the tolerance range or threshold, they can communicate, and the interaction tends to bring them closer, according to the rule
\[
x_i(t + 1) = x_i(t) + \mu (x_j(t) - x_i(t)) \\
x_j(t + 1) = x_j(t) - \mu (x_j(t) - x_i(t))
\]
\[ (6) \]

where \( \mu \in [0, 1/2] \) is a parameter to ensure convergence in the iteration. For the sake of simplicity and to avoid too large number of parameters, we will consider the case of \( \mu = 1/2 \): \( i \) and \( j \) adopt the same intermediate opinion after communication [21]. The tolerance parameter plays a crucial role in the ability of the population of agents to reach a global consensus or not. It is indeed intuitively clear that, for large tolerance values, agents can easily communicate and converge to a global consensus. On the contrary, small values of \( d \) naturally lead to the final coexistence of several remaining opinions [13, 21].

For large populations, it may be more realistic to consider that the interactions between agents define a network with a finite average connectivity: each agent has a limited number of neighbours and cannot communicate \textit{a priori} with all the other agents. A typical example of such an interaction network structure is given by an uncorrelated random graph in which agents have \( k \) acquaintances on average, i.e. the initial network corresponds to an Erdős-Rényi network with average degree \( k \). While such a topology lacks many interesting features displayed by real social networks, such as degree heterogeneity or community structures, it is nevertheless interesting to start from this simple case as a reference frame. As in the previous case, we focus on the fact that the network topology may evolve on the same timescale as the agents’ opinions. Agents can indeed break a connection or establish new ones, depending on the success of the corresponding relationship. The
rules defining the evolution of the network topology can be modeled in many different ways. A possibility is to consider that links decay at a constant rate, independently from the agents’ opinions [7]. In the case of opinion dynamics, we consider instead that only neighbouring agents with far apart opinions (i.e., \(|x_i(t) - x_j(t)| > d\)) may terminate their relationship. In order to keep the average number of interactions constant, a new link is then introduced between one of the agents having lost a connection and another agent, chosen at random. The new link may of course break again if the newly connected agents have too-far-apart opinions. The rewiring process thus occurs as a random search for agents with close-enough opinions.

Even a simple rewiring process such as the one depicted above leads to the introduction of two new parameters. The first one, \(w\), quantifies the relative frequencies of the two following processes: a local opinion convergence for agents whose opinions are within the tolerance range, and a rewiring process for agents whose opinions differ more. At each time step \(t\), a node \(i\) and one of its neighbours \(j\) are chosen at random. With probability \(w\), an attempt to break the connection between \(i\) and \(j\) is made: if \(|x_i(t) - x_j(t)| > d\), the link \((i, j)\) is removed and a new link is created. With probability \(1 - w\) on the other hand, the opinions evolve according to (6) if they are within the tolerance range. The second parameter concerns the creation of a new link whenever a link \((i, j)\) has been removed: a new node \(k\) is then chosen at random, and with probability \(p \in [0 : 1]\) a link \((i, k)\) is created, while with probability \(1 - p\) the new link instead connects \(j\) and \(k\). Since \(j\) is chosen as a neighbour of a randomly chosen node \(i\), it will have on average a larger degree. Larger \(p\) therefore favors the removal of links from larger degree nodes, while small \(p\) means that large degree nodes preferentially keep their links. We will also see how the parameter \(p\) affects the final structure of the agents groups.

The dynamics stops when no possible update is left. If \(w > 0\), this corresponds to a state in which no link connects nodes with different opinions. This can correspond either to a single connected network in which all agents share the same opinion, or to several disconnected clusters representing different opinions. For \(w = 0\) on the other hand, the final state is reached when neighbouring agents either share the same opinion or differ of more than the tolerance \(d\).

Before turning to a detailed analysis of the model, we illustrate in Fig. 2 the different behavior observed for static and adaptive networks. The figures show the evolution of the opinions of 250 out of \(N = 1000\) agents as a function of time, in each case for one single realization of the dynamics with \(d = 0.15\). The opinions are initially randomly distributed on the interval \([0, 1]\). When the interaction network is static, local convergence processes take place and lead to a large number of opinion clusters in the final state, with few macroscopic size opinion clusters and many small size groups: agents with similar opinions may be distant on the network and not be able to communicate. This is in contrast with the mean-field case in which all agents are linked together so that the final opinion clusters are less numerous and more separated in the opinion space. Figure 2, which corresponds to an adaptive network with \(w = 0.5\), is strikingly in contrast with the static case: no small groups are observed.

In particular, the whole cluster-size distribution gives a complete description of the system, where a cluster is defined as a connected group of agents sharing the same opinion. Interesting summaries are given by the number of clusters, the size of the largest and second-largest opinion-cluster which will tell us about the behavior of the clusters with macroscopic size (because of the possible large number of small clusters, the average size may be biased towards small values and is therefore of less interest).
4 Main results

4.1 Analytical computations of the steady state fitness probability distribution

We now consider the choice of this function. This function can be very general, anyway we believe that the more a vertex is connected to other vertices, the better for the general properties of the network. This picture is self-consistent provided that the larger \( x \) and \( y \), the larger \( f(x, y) \). Following the results of refs \([25, 26]\), the simplest unbiased \([26]\) choice for such a function is

\[
f(x, y) = \frac{zxy}{1 + zxy}
\]

where \( z \) is a positive parameter controlling the number of links. This choice generates a network with a nonrandom, fitness dependent expected degree sequence \([25, 26]\), which in this case is not known a priori and will be determined by the fitness distribution at the stationary state. All other higher-order properties are completely random, except for the structural correlations induced by the degree sequence \([25, 26]\). It therefore represents the fitness-dependent version of the so-called configuration model \([27, 28]\). As we show later on, structural correlations have an important impact on the dynamics. With the above choice, \( \rho_ho(x) \) can be directly computed analytically through eq.\((4)\). However we write it in a different form, which is equivalent when \( N \rightarrow \infty \), in order to solve also more complicated integrals involving it later on. We use the trick \( \langle f(x, m) \rangle \approx f(x, \langle m \rangle) \) where the angular brackets denote an average over the distribution \( q(m) \) of the minimum fitness, that is \( \tau^{-1} \int_0^\tau f(x, m)dm \approx (z\tau^2/2)/(1 + z\tau^2/2) \). As we show below, when \( N \rightarrow \infty \) this approximated expression becomes exact. Then eq.\((4)\) yields

\[
\rho_ho(x) = \begin{cases} 
(\tau N)^{-1} & x < \tau \\
(\tau N)^{-1} + 2/(zN\tau^2x) & x > \tau
\end{cases}
\]

\[(8)\]

Figure 2: Left figure: Evolution of the opinions of 25% of the population, denoted by lines, for a system of \( 10^3 \) agents with tolerance \( d = 0.15 \) and average degree \( k = 5 \), on a static network for a single run. The evolution of the opinion of a few individuals is highlighted with color. Right figure: Same as the left one, for an adaptive network when the rate of rewiring is \( w = 0.5 \).
where $\tau$ is the solution of eq.(5), which reads
\[
\frac{1}{\tau} + \frac{1}{2\tau^2} \log \frac{1}{\tau^2} = N
\] (9)

A trivial solution is $\tau \to 0$, obtained if $z$ remains finite as $N \to \infty$, or in other words if $zN \to \infty$. On the other hand, nonzero solutions exist. If $\tau \neq 0$, then for large $N$ the term $1/\tau$ in the above expression can be neglected with respect to $N$. Multiplying both sides by $z$ yields
\[
\frac{1}{\tau^2} \log \frac{1}{\tau^2} = zN
\] (10)
whose solution is $\tau = \sqrt{\frac{\phi(zN)}{zN}}$, where $\phi(x)$ is the ProductLog function defined as the solution of $\phi e^\phi = x$. Putting these results together, we have
\[
\tau = \sqrt{\frac{\phi(zN)}{zN}} \to \begin{cases} 
1 & zN \to 0 \\
\frac{\phi(d)/d}{zN = d} & zN \to \infty \\
0 & zN \to \infty 
\end{cases}
\] (11)

As for random graphs, we find a marked transition as the scaling of $z$ changes from $N^{-1}$ to more rapidly decaying. This suggests an analogous underlying percolation transition. As we show below, this is indeed the case. We can therefore still refer to the sub-critical, sparse and dense regimes. Note that as $N \to \infty$ we have $f(x,y) = zxy$ in the sparse and sub-critical regimes since $zxy < z \ll 1$, which implies that we can neglect $zxy$ in the denominator of eq.(7). Therefore the expression $\langle f(x,m) \rangle = f(x,\langle m \rangle)$ is exact. On the other hand, in the dense regime we have $\tau \to 0$, which again implies the same expression since $q(m)$ becomes the Dirac delta function $\delta(m)$. Therefore our trick to use the above expression turns out to be exact in all regimes for $N \to \infty$.

The main panel of fig.3 shows the cumulative density function (CDF) of the fitness $\tilde{\rho}_0(x) \equiv \int_x^1 \rho_0(x')dx'$, while the inset shows a plot of $\tau(zN)$. The theoretical results are in excellent agreement with numerical simulations. As predicted by eq.(8), $\tilde{\rho}_0(x)$ is the superposition of a uniform distribution and a power-law with exponent $-1$. For $z \ll 1$ we have $f(x,y) \approx zxy$ and $\hat{\rho}(x) \propto x^{-1}$ for $x > \tau$. This purely power-law behaviour, that becomes exact in the sparse regime $z = d/N$ for $N \to \infty$, results in a logarithmic CDF looking like a straight line in log-linear axes. Note that, despite the value of the exponent, the presence of a nonzero lower threshold ensures that $\tilde{\rho}_0(x)$ is normalizable. This mechanism may provide a natural explanation for the onset of Pareto distributions with a finite minimum value in real systems. By contrast, for large $z$ the uniform part is non-vanishing and $\tilde{\rho}_0(x)$ deviates from the purely power-law behaviour. The decay of $\tilde{\rho}_0(x)$ for $x > \tau$ is a completely novel outcome of the extremal dynamics due to the feedback with the topology: now the fittest species at a given time is also the most likely to be connected to the least fit species and to mutate at the following time-step. Being more connected also means being more subject to changes. This enriches the coexistence patterns displayed on static networks. Even if one of the most studied properties of the BS model on regular lattices is the statistics of avalanches characterizing the SOC behaviour [17], we do not consider it here. This is because, as shown in ref.[29], when considering long-range [18] instead of nearest-neighbour connections, it can lead to a wrong assessment of the SOC state, which is put into question by the absence of spatial correlations even if the avalanches are power-law distributed. Rather, we further characterize the topology at the stationary state by considering the degree distribution $P(k)$ and the degree correlations. We do this by studying, as a function of $x$, the average
degree \( k(x) = N \int f(x,y) \hat{\rho}_z(y) dy \) of a vertex with fitness \( x \):

\[
k(x) = \frac{2}{z^2} \ln \frac{1 + zx}{1 + z^2} + \frac{zx - \ln(1 + zx)}{2zx}
\]  

(12)

The inverse function \( x(k) \) can be used to obtain the cumulative degree distribution \( P_>(k) \equiv \int_k^{k(1)} P(k')dk' = \hat{\rho}_>(x(k)) \). As shown in Fig.4, \( k(x) \) is linear for small \( z \) since \( f(x,z) \approx zx \), while for large \( z \) it saturates to the maximum value \( k_{\text{max}} = k(1) \). This implies that in the sparse regime \( P(k) \) mimics \( \hat{\rho}_o(x) \) and is characterized by the threshold value \( k(\tau) \) and by a power-law decay \( P(k) \propto k^{-1} \) above it (see Fig.4). Note that here \( \tau \) remains finite even if \( P(k) \propto k^{-\gamma} \) with \( \gamma < 3 \), in striking contrast with what obtained on static scale-free networks [20]. Differently, for large \( z \) the saturation \( k \to k_{\text{max}} \) translates into a cut-off that makes \( P(k) \) deviate from the pure power-law behavior for \( k > k(\tau) \).

As shown in refs. [25, 30], this saturation determines anticorrelation between the degrees of neighbouring vertices (disassortativity) and a hierarchy of degree-dependent clustering coefficients as observed in many real-world networks (this is not shown here for brevity).

As \( N \to \infty \), these correlations vanish in the sparse regime \( (\tau > 0) \), while they survive in the dense regime \( (\tau \to 0) \). Structural correlations and a nonzero threshold \( \tau \) are then mutually excluding in this model, which is another interesting effect of the feedback we have introduced. Our results represent a first step into the unexplored domain of systems with generic self-organized coupling between dynamics and topology. A huge class of such processes needs to be studied in the future, to further understand the unexpected effects of this coupling.

4.2 Evolving consensus model

Let us first focus on the case \( p = 0 \) and compare the results of the opinions evolution on a static and a dynamically adaptive network. Figure 4.2 displays for both cases the size of the largest \((S_{\text{max}})/N\) and second largest \((S_2)/N\) opinion clusters in the final
state, where a cluster is defined as a connected group of agents sharing the same opinion. In both cases, at large tolerance $d$, a global consensus is achieved, with a single cluster containing all agents. A jump of $\langle S_{\text{max}} \rangle / N$ from a value close to 1 to a value close to 1/2 is observed at a critical value $d_{c1}(w)$ which depends on the rewiring rate. Interestingly, $d_{c1}(w)$ increases with $w$ (see Fig. 4.2). If the rewiring is more frequent, it allows more easily to break the network in two pieces, since agents can more easily search for other agents with whom they can communicate, and break ties with the ones with too different opinions: the formation of different clusters is favored, and larger tolerance values are necessary to achieve consensus.

At $d < d_{c1}$, an apparently polarized state is entered, with a first and second-largest clusters of similar extensive sizes, and apparently similar behavior for static and adaptive networks. A difference appears however at small tolerance values: for static networks, $\langle S_{\text{max}} \rangle / N$ vanishes for $d < d_2$ in the thermodynamic limit. The final state of the system is then *fragmented*, with no cluster of extensive size. This polarized-fragmented transition is due to the finite connectivity of the agents and corresponds to a percolation phenomenon. An agent $i$ with $k$ connections and tolerance range $d$ will indeed have on average $2dk$ neighbours with whom to communicate. For an average degree $k < 1/(2d)$, there will not be any percolating paths of agents with close enough opinions and only very small clusters of agreeing agents can be formed. In the case of adaptive networks on the other hand, the fragmented phase disappears as soon as the rewiring of the links is enabled. The size of the largest component decays smoothly but remains extensive as the tolerance decreases. Rewiring processes thus allow the small clusters, initially non-percolating, to create links between them and reach extensive sizes even below the polarized-to-fragmented transition appearing on static networks.

Further insight into the differences between static and adaptive networks is provided by the number of opinion clusters in the final state, $N_{\text{clusters}}$, shown in Fig. 4.2. For $d_2 \leq d \leq d_{c1}$, an extensive number of clusters is indeed obtained in the static case, saturating at $O(N)$ at $d_2$. The system presents therefore a “false”-polarized state, with a coexistence of macroscopic opinion clusters with an extensive number of finite size clusters. As $d$ decreases, more and more macroscopic clusters appear, as in mean-field [21], but
there is also an extensive proliferation of finite size “microscopic” clusters. For adaptive networks, the number of clusters is much smaller, and decreases as \( w \) increases. In fact, the precise investigation of the cluster size distribution reveals that the density of nodes in non-extensive clusters vanishes in the thermodynamic limit \([12]\). The polarized phase on adaptive networks differs therefore strongly from the one on static networks: thanks to the possibility of link rewiring, agents who would remain isolated (or in very small groups) on a static network may manage to find agents with whom to communicate and thus enter a macroscopic cluster. Without rewiring on the other hand, a macroscopic number of agents remain in fragmented components which coexist with few macroscopic clusters.

4.3 Group structure

In the final state of the system, an interesting question concerns the structural differences between the various opinion groups that have been formed. The most basic property one can investigate is the average degree of an agent. It turns out that the average degree of an agent inside a group is strongly correlated with the size of the group. On static networks, the average degree of a cluster is a linear function of its size (left plot of Fig. 6), which can be explained as follows: for a cluster of size \( S \), the probability for a node to have a link pointing towards this cluster is simply \( S/N \), and a node of degree \( k \) will have on average \( kS/N \) links pointing towards other nodes in the cluster. The average “in-degree” of the nodes in a cluster of size \( S \) is therefore \( Sk/N \) (assuming that there is no correlation between the degree of a node in the network and the cluster to which it finally belongs).
Figure 6: Left figure: average degree of the agents sharing the same opinion as a function of the size of the uni-opinion groups, $S$, on static networks. The dashed line corresponds to a linear behaviour. Right figure: same in the case of adaptive networks when $w = 0.5$ and $p = 0$ for different average degrees and tolerance values. The dashed line shows a power-law $S^{0.6}$. Inset: the binned average of the same measurement for $k = 20$, $d = 0.01$, $w = 0.5$, and different values of $p$. The lines correspond to the power laws $S^{0.6}$, $S$, and $S^{1.3}$. In both figures, $N = 10^4$ and data points were collected from 10 realizations for each set of parameter values.

On adaptive networks on the other hand, the linear relationship is no longer valid, as shown in Fig. 6. At small tolerance values, many clusters are obtained, with very different sizes. A power-law-like relationship appears between the clusters’ size and degree. The behaviour depends on $p$, the parameter of the rewiring rule: a sub-linear relationship holds for small $p$ values while a super-linear behaviour appears for $p$ close to 1. Since these cases correspond to the situations when the development of the opinion clusters takes place at a much shorter timescale than that of the topological clusters [12], an analytical treatment of the problem is possible, investigating the diffusion of the links between the clusters of constant opinions at a mean-field level [31]. For large tolerance values, the cluster’s average degree becomes less correlated with its size (right graph of Fig. 6): this is due to the fact that the clusters correspond to large fractions of the original network and their average degree saturates at $\bar{k}$.

4.4 Convergence time

On static networks, the time to converge to the final state of the system, $t_{\text{conv}}$, is determined by the topological properties of the opinion clusters. In turn, the behavior of this topology is mostly determined by the distance of the tolerance of the agents from that of the polarized-to-fragmented transition, $d_{c2}$ (see Fig. 7): For $d_{c2} < d$, $t_{\text{conv}}$ grows linearly with the system size and increases as $d$ decreases: the clusters formed by agents who can communicate become more and more tree-like, which slows down the convergence to a common opinion. As $d \rightarrow d_{c2}$, $t_{\text{conv}}$ diverges as a signature of the phase transition. For $d < d_{c2}$, the clusters of agents with close enough opinions become small and the convergence time decreases as $d$ decreases.

On adaptive networks, two scenarios are possible: on the one hand, if the network evolution is slow compared to the timescale of opinion formation, $t_{\text{conv}}$ is mostly determined
Figure 7: A) convergence time, measured in the number of simulation steps, on static networks for different system sizes as a function of the tolerance, \(d\), when \(\bar{k} = 10\) (The data points were generated averaging over 10 to 100 realizations). B) rescaled convergence time on adaptive networks for small rewiring rates as a function of \(d\) for different system sizes, average degrees, and rewiring rates (Data points averaged over 30 realizations.). The dashed line is proportional to \(1/d\). C) Rescaled convergence time on adaptive networks as a function of \(d\) for different system sizes and rewiring rates (\(\bar{k} = 10\)) averaged over 100 realizations.

by the characteristic time of topological cluster formation, \(t_l\). It is important to note that, even in this limit, the network topology cannot be considered static and adaptivity plays an important role in the early time evolution of the system too, as shown in [12]. The scaling of \(t_l\) and therefore \(t_{\text{con}}\) can be estimated by considering a typical opinion cluster: its size is proportional to the tolerance range of the agents, \(2d\); the number of its links which need to be rewired is proportional to the total number of links (\(\propto \bar{k}N\)), and to the amount of opinions outside of the tolerance range (\(\propto (1 - 2d)\)); the probability to rewire a link towards an agent with a close enough opinion is moreover \(\propto d\) and the time between two link updates is \(\propto 1/w\) so that

\[
t_l \propto \bar{k}(1 - 2d)/(wd).
\]

Figure 7B) shows the rescaled convergence time \(wt_{\text{con}}/\bar{k}\) as a function of \(d\) for different parameter values, in good agreement with Eq. (13). If the network evolution is fast, on the other hand, it is possible for an agent to rewire most links towards agents with close opinions in a short time. This scenario takes place when \(t_l\) is less than or comparable to \(1/(1 - w)\), the average time between two fruitful discussions. In this case, the convergence time is expected to scale simply as \(1/(1 - w)\), as indeed shown in Fig. 7C).

5 Conclusions

We investigated the evolution of adaptive network topology in the case of a continuous change of agents and in the case of the Deffuant model where agents are endowed with bounded confidence in their neighbours and seek consensus with them only if their opinion is similar to some extent. In the first case we have been able to show numerically and analytically that it is possible to attain a scale-free steady-state distribution for the degrees in the networks after some simple hypothesis are made on the network evolution.
As regards the opinion dynamics model, we found that adaptivity has different effects on clusters of different sizes. Adaptivity makes non-extensive groups disappear by letting agents find friends with close-enough opinions and form extensive groups. Macroscopic size groups, however, tend to break to smaller ones since agents break connections with agents having too-distant opinions. We also saw that rewiring of the links introduces correlations between the inner structure of the groups and their size, depending on the specific rewiring rule. The scaling properties of the convergence time was also determined by the rewiring events in contrast to static networks where \( t_c \) was rather a function of the percolation properties of the social network. Finally, the investigation of a variant of the model further indicated that consensus formation exhibits more robust features on adaptive than on static networks.

These studies have been published on Physical Review E and Nature Physics.

References

Washington, DC, USA, ACM Press.
DELIS-TR-0360
11-14 May, Oakland, California, USA.
DELIS-0410.
[12] B. Kozma B and A. Barrat Consensus formation on adaptive networks preprint
DELIS-TR-0496