Organic Design of Massively Distributed Systems: A Complex Networks Perspective

Ingo Scholtes · Claudio Juan Tessone

Introduction

Networked computing systems are becoming increasingly large, complex and – at the same time – important for our everyday lives. Many of the services we rely on, are now being collaboratively provided by thousands or millions of machines in large Peer-to-Peer (P2P) systems or data centers. Sustaining the robustness and manageability of such systems are challenging tasks. Because of the ongoing miniaturization of network devices, their price-decline as well as the proliferation of mobile and embedded computing equipment, scenarios in which billions of devices are connected to global-scale information systems become reality. Promising aspects of the coalescence of the virtual and physical world that results from the increasing incorporation of communication technology into everyday objects, as well as the associated technical and societal challenges have been highlighted in the visions of *Ubiquitous Computing* [46] or the *Internet of Things* [30].

Building services and applications in an environment of numerous dynamic and error-prone communication devices poses enormous technical challenges in terms of scalability, efficiency, manageability, and robustness. It is frequently argued that, in order to cope with these challenges, computing technologies need to adopt the remarkable self-organization, self-adaptation and self-healing qualities of biological systems. Facilitated by advances in the study of principles underlying self-organization mechanisms as well as the massively increasing complexity of technical infrastructure, in recent years the vision of *Organic Computing* has been gaining momentum. The development of associated technologies is likely to go hand-in-hand with a cutback of sophisticated algorithmic schemes and deterministically organized communication structures. Instead, adopting a heterodox approach that utilizes simple mechanisms that underlie self-organization, adaptivity, robustness, and resilience in natural systems is reasonable.

Driven mainly by the availability of massive data sets, during the last decade these principles have been studied in a variety of different contexts, including disparate fields like biology, physics, computer science, economics, and sociology. The resulting interdisciplinary strand of research is subsumed as *complex systems science* and here we argue that it offers a promising and quickly evolving methodological framework for the modeling, design, and control of *organic computing systems*. Providing a set of tools and abstractions to analyze the collective properties of systems comprised of a large number of stochastic, interacting elements, complex systems science addresses one of the key questions emerging in the scenario described above: How can we analyze, monitor, and control the structure and dynamics of massively distributed systems evolving from distributed mechanisms?

In this article we address some aspects that arise in the management of robust and adaptive overlay topologies for massively distributed systems. Here we adopt the perspective of statistical mechanics

DOI 10.1007/s00287-012-0597-4

Ingo Scholtes · Claudio Juan Tessone
ETH Zürich, Chair of Systems Design,
CH-8032 Zürich, Switzerland
E-Mail: {ischoltes, tessonec}@ethz.ch

Abstract
The vision of *Organic Computing* addresses challenges that arise in the design of future information systems that are comprised of numerous, heterogeneous, resource-constrained and error-prone components. The notion *organic* highlights the idea that, in order to be manageable, such systems should exhibit self-organization, self-adaptation and self-healing characteristics similar to those of biological systems. In recent years, the principles underlying these characteristics are increasingly being investigated from the perspective of complex systems science, particularly using the conceptual framework of statistical physics and statistical mechanics. In this article, we review some of the interesting relations between statistical physics and networked systems and discuss applications in the engineering of organic overlay networks with predictable macroscopic properties.

on the emergence of complex structures and collective dynamics in networks – an area that has been particularly active and successful during the last decade. In Sect. "Overlays, Random Graphs, and Complex Networks", we summarize structured and unstructured approaches to the management of overlay networks. Here we additionally review the relevance of random graph theory for the design of unstructured systems and introduce the relations between the study of statistical mechanics and complex networks as well as the modeling of dynamical processes. In Sect. "Managing Organic Overlays – A Thermodynamic Perspective", we discuss abstractions from statistical mechanics and statistical physics in the design of organic overlay networks. In Sect. "Conclusion and Outlook", we summarize challenges and opportunities of using complex systems science in the engineering of distributed systems with predictable and controllable self-* properties.

Overlays, Random Graphs, and Complex Networks

Overlay networks – which define virtual connections on top of physical communication infrastructures – are becoming an increasingly important issue. As argued in [45], the possibility to define communication topologies and protocols at the application layer without having to make a – potentially globally – coordinated change of existing protocols, standards, and communication infrastructures is an important factor for a quick proliferation of novel services on the Internet as well as in large-scale data centers. The research of overlay topologies as well as efficient distributed algorithms providing core functionality like search, routing, and content dissemination has received a lot of attention recently.

Most of this research has been done in the context of P2P systems, which are now increasingly used for the cost-efficient distribution of data for example by means of the BitTorrent protocol, the provision of video-telephony services like Skype, or even to face challenges emerging in large-scale scientific setups like the Large Hadron Collider[43]. One usually distinguishes *structured* and *unstructured* approaches in the management of overlay topologies. Most of the currently deployed systems belong to the former category. In such structured systems, virtual connections between machines are created in a globally consistent way to construct a particular network topology. While this allows for the development of highly efficient algorithms for distributed search, routing, or information dissemination, the major difficulty is to maintain this fine-tuned topology under dynamic conditions. Reconsidering the scenario outlined in Sect. "Introduction", maintaining fine-tuned structures will entail massive complexities due to the excessive fluctuation of participating devices and the associated concurrency. In fact, for the distributed hash table *Chord* it has been argued in [7] that in settings with very large numbers of highly dynamic participants, the communication overhead imposed by mere topology maintenance and management schemes could exceed the cost for actual data transfer operations and thus dominates performance. It has further been argued that designing, implementing, and debugging topology maintenance schemes pose a huge challenge due to the massive concurrency that is introduced by failing or joining machines. These problems of structured overlays are well known in the literature and question their usability in future scenarios like the one laid out in Sect. "Introduction". Hence, alternative approaches for dealing with large and dynamic settings are being studied.

Unstructured Topologies and Random Graph Theory

A straight-forward idea is to use *unstructured overlays* in which virtual connections between machines are created in a simple, uncoordinated fashion while still allowing all machines to communicate with each other. While this reduces the overhead of topology management, it necessitates probabilistic algorithms for example for distributed search or routing that make no – or at least less specific – assumptions about the structure of the network or the placement of data items. Such schemes are inevitably less efficient compared to those tailored for a particular network structure. Nevertheless, they are significantly simpler to implement and allow for larger degrees of freedom in terms of adapting the network structure to operational conditions.

In terms of modeling performance and robustness, most unstructured approaches to the management of overlays rely – either explicitly or implicitly – on results from the field of random graph theory which was established more than 50 years ago [20]. In order to explain the analogies between large, dynamic networked systems and statistical mechanics, we briefly recall one of the basic models of random graph theory. The so-called $G(n,p)$ model defines a probability space that contains all possible graphs or networks[1] G with n nodes. Assuming that edges between pairs of nodes are being generated by a stochastic process with uniform probability p, the $G(n,p)$ model assigns each network G with n nodes and m edges the same probability to be created:

$$P_G(n,p) = p^m \cdot (1-p)^{n(n-1)/2-m}$$

This simple stochastic model for networks has been used in the modeling of a variety of real-world networks. In particular, one can use it to make predictions about the properties of unstructured overlays, if virtual connections are assumed to be created at random with probability p or, alternatively, if an average number of $p \cdot n(n-1)/2$ connections are established between randomly chosen pairs of nodes.

In general, in the study of random networks one is particularly interested in properties that hold for a subset of network realizations whose probability measure converges to 1 as the size of the generated networks (in terms of the number of nodes) increases. In this case one can say that a property holds *asymptotically almost surely* for a randomly generated network. This is because the probability to draw a network that does not exhibit the property in question quickly vanishes. An authoritative overview of the interesting results derived from this perspective can be found in [14]. Two well-known examples of particular relevance for the design of overlay networks are results on the critical percolation threshold and the diameter. The critical percolation threshold refers to a point in the $G(n,p)$ model's parameter p above which the generated networks almost surely contain a connected component that is of the order of the network size. For the $G(n,p)$ model it has been found that connected components of a random graph are with high probability of the order $log(n)$ if $p < 1/n$. For $p > 1/n$ the connected component is of the order n [20][2]. In practical terms, this result is a crucial prerequisite for the feasibility of unstructured overlay management schemes since it tells that – if at least a certain minimum number of connections is created in a random and uncoordinated fashion – all machines will be able to communicate with each other with a high probability. Another set of results that are important for overlays with random structures relates the parameter p to the diameter of the resulting topology. It further gives a criterion for the emergence of so-called *small-world topologies* which are assumed to have a diameter of the order of the logarithm of the network size. For the $G(n,p)$ model, it has been shown that the diameter is with high probability of order $log(n)/log(np)$, if the average number of links per node is at least 1. In the design of unstructured topologies, this argument is crucial to reason about the efficiency of search and routing schemes.

Statistical Mechanics of Complex Networks

As argued in [15], the existence of so-called *critical points* in the $G(n,p)$ model's parameter p and the associated sudden change of macroscopic qualities like diameter or connectedness, highlights interesting relations to phase transition phenomena in statistical physics, i. e., sudden changes of material properties as aggregate control parameters (e. g.,

[1] Throughout this article, we will use the terms *graph* and *network* interchangeably.

[2] Interestingly this is a so-called double-jump transition, i. e., for $p = 1/n$ the size of the connected component is of the order $n^{\frac{2}{3}}$.

temperature or pressure) change slightly. In recent years, these analogies to fundamental natural phenomena have been substantially deepened by reframing the study of random graph structures in terms of statistical mechanics and statistical physics (see e. g., [2, 11, 19, 21, 23, 37]). This perspective is possible since statistical mechanics reasons about configurations of many-particle systems, just like random graph theory reasons about network configurations. Each of these particle configurations – the so-called *microstate* – fixes the exact positions and energy states of all particles present in a given volume of space at a given temperature and total energy. On the basis of energy distributions, particle positions and fluctuations induced by temperature as well as a quantum approximation, each microstate can be assigned a probability based on combinatorial arguments. The set of all possible microstates thus defines a probability space which is called a *statistical ensemble*. The study of material properties translates to reasoning about the probability measures for a subset of these particle configurations, just like random graph theory reasons – for instance – about the probability of the subset of network configurations with a given diameter.

On the basis of these similarities, it has been shown for instance in [23] that the $G(n,p)$ model of classical random graphs can be reframed in terms of the so-called *grand-canonical ensemble* of many-particle systems. In this perspective, the study of statistical ensembles with fixed thermodynamic quantities like volume, chemical potential, and temperature in statistical mechanics translates to the study of *network ensembles* with fixed aggregate statistics like a given number of nodes or edges, degree distributions, degree-degree correlations, or clustering coefficients. In the resulting ensembles all realizations with the same aggregate statistics (e. g., all networks with a particular degree sequence) are assumed to have equal probability. In statistical mechanics, this corresponds to an *adiabatic* situation in thermodynamic equilibrium while at the same time the accessible states are being constrained by certain fixed quantities. In the remainder of this article we will thus refer to such probability spaces as *constrained adiabatic ensembles*.

During the last decade such constrained adiabatic ensembles of networks have been studied extensively in the fields of *complex networks* and *statistical mechanics* [2, 34]. A particularly active strand of research in this direction is the study of ensembles with fixed degree distributions following, for instance, a power law. This is, the probability that a randomly chosen node in the network has exactly k links, is proportional to $k^{-\gamma}$ for some $\gamma \in [2,\infty)$. Since the classical $G(n,p)$ model can be viewed alternatively as an adiabatic ensemble with a fixed Poissonian degree distribution, this naturally extends earlier works on random graphs. Since power-law degree distributions have been observed for a number of real-world networks [2], the associated constrained adiabatic ensemble effectively serves as a null model for these kinds of systems.

Over recent years, a rich set of results both on the collective properties of such networks, as well as on simple local mechanisms by which they emerge have been obtained. Extensive surveys of results based on this statistical physics perspective on complex network structures can be found in [10, 12]. Prominent results for the special case of networks with heavy-tailed degree distributions include their resilience against random faults [16] and targeted attacks [17] or the performance of probabilistic distributed search schemes [1]. In the remainder of this article, we illustrate that these results and – more importantly – the underlying methodological framework are relevant for the design of robust networked systems with organic properties.

Dynamical Processes in Complex Networks

So far, we have commented on the structural properties emerging in networks being drawn from constrained adiabatic ensembles. For the design of distributed algorithms which must operate in an efficient and reliable way in large dynamic overlays, it is equally important to have tools at hand that allow one to reason about dynamical processes (like e. g., information dissemination, synchronization, or distributed search) operating upon them. In order to formalize the problem, a useful representation of a network is its adjacency matrix $\boldsymbol{A}$ where each element is $a_{ij} = 1$ ($a_{ij} = 0$) if the nodes i and j are connected (respectively, disconnected). Then, the *spectrum* of such a network is given by the set of n eigenvalues of its adjacency matrix. For the $G(n,p)$ model, it is possible to characterize the spectrum of the networks in the limit of diverging network sizes. In this model, and if there is a giant cluster that spans the complete network, the probability $p(\lambda)$ of finding an eigenvalue λ^a in the spectrum follows the

so-called *semi-circle law* [31]:

$$p(\lambda^a)=\begin{cases}\frac{\sqrt{4\,n\,p(1-p)-(\lambda^a)^2}}{2\pi\,n\,p(1-p)} & \text{if } |\lambda^a|<2\sqrt{n\,p(1-p)}\\ 0 & \text{if } |\lambda^a|\geq 2\sqrt{n\,p(1-p)}\end{cases}$$

The bulk of the distribution of eigenvalues is centered around the null eigenvalue, with a characteristic size proportional to $\sqrt{n}$. However, the largest eigenvalue λ_1^a, is proportional to $p \cdot n$.

In order to study continuous dynamical processes in networks, we consider that each node i is endowed with a continuous variable x_i which describes its current dynamical state. Then, the change of its state can be thought to be given by

$$\frac{d}{dt}x_i = f_i(x_i) + C\sum_{j=1}^{n}\ell_{ij}\cdot h(x_j), \qquad (1)$$

where $f(x_i)$ is a function describing a deterministic change of state of node i given its current state, and $h(x_j)$ is a coupling function given the state of node j, and C is the coupling strength. In (1), ℓ_{ij} are the elements of the so-called *Laplacian matrix*, $\boldsymbol{L}$. Such a matrix is defined as $\ell_{ij} = -k_i\delta_{ij} + a_{ij}$, where k_i is the degree of node i, and δ_{ij} are the elements of the identity matrix. The Laplacian matrix naturally extends the Laplacian operator ∇^2 – as used in the description of dynamics in spatially extended, physical systems – into a discrete manifold.

Different kinds of collective behavior have been observed in the study of dynamical processes in networks emerging from different constrained adiabatic ensembles [5, 12]. Whenever different nodes show the same dynamics, it can be said that a synchronized state has emerged. In general, it was shown [9] that such synchronized states are stable if $\lambda_n^l/\lambda_2^l < \beta$, where λ_n^l and λ_2^l are (respectively) the largest and smallest nonzero eigenvalues of the Laplacian matrix, and β is given by the functions $f(x)$ and $h(x)$ describing the dynamical properties of the system. This means that there is a relationship between structural properties of the network (in terms of its eigenvalues) and dynamical process taking place at the level of nodes as to whether the system is able to synchronize. This runs against typical intuition on synchronization phenomena, stating that if the coupling strength is large enough, the system should exhibit a synchronized state. Furthermore, if a synchronized state emerges the question of how much time the system needs to reach such a state emerges. For arbitrary network ensembles it was shown [4] that the consensus time is of the order $T_C = (\ln C - \ln\epsilon/2)/\lambda_2^l$, where C is an integration constant (related to the initial conditions) and ϵ is a *synchronization threshold*, i. e., the difference below which nodes are assumed to be synchronized.

Statements on the impact of the Laplacian spectrum play an important role when assessing the properties of synchronization and consensus schemes applied in distributed systems and relating the emerging collective dynamics to the topology of the network. In structured and unstructured overlays, synchronization models have been proposed to provide a network-wide synchronous notion of time epochs or protocol cycles [6, 8, 29, 42]. In fact, it is even possible to infer from the synchronization dynamics at the level of individual nodes statements about the cluster structure and the Laplacian matrix of the network [41]. Finally, regarding detrimental collective phenomena like the synchronization of periodic routing messages described in [22], an analysis of the spectral properties of networks is crucial.

Dynamical processes additionally can take place in a discrete space. Perhaps the simplest of such processes are random walks in a network. The question how fast a node receives and spreads information in such a random process is important for example in the study of stochastic search and transport phenomena[35]. To quantify this, the *random-walk centrality* $\mathcal{C}_i$ was introduced which – for node i – is given by:

$$\mathcal{C}_i = \frac{k_i}{\sum_i k_i}\left(P_{ii}(t) - \frac{k_i}{\sum_i k_i}\right)^{-1}$$

Here, $P_{ii}(t)$ is the probability that a random walk which started at i at time zero, returns to the same node after t time steps, and can be readily computed by means of the adjacency matrix $\boldsymbol{A}$. This index shows that nodes with large centrality receive information faster. In the case of limited bandwidth capacities, it indicates that such nodes may become overloaded first.

Arguments about the relation between spectral properties of networks and the dynamics of random walk-related processes have been used – again implicitly and explicitly – in a variety of contexts like multicast communication [24], database replica maintenance [18], the computation of network-wide aggregates [26], or random sampling [48]. By means of the spectral perspective on the equivalent

of the Laplacian operator in networks, it is possible to study the performance of these schemes for unstructured overlays given that they are drawn from a particular statistical ensemble. This clearly demonstrates the relevance of these techniques for the design of organic networked systems, while at the same time highlighting interesting mathematical relations to fundamental natural phenomena being studied in statistical physics.

Managing Organic Overlays – A Thermodynamic Perspective

From an engineer's perspective, being able to assess the structural and dynamical properties of networks with complex structures sounds appealing. However, since most of the findings are based on simple stochastic models for complex systems one needs to be careful when applying them to real-world systems: Statistical ensembles of complex networks should be viewed as mere *null models* for networked systems with complex structures and it is unlikely that they accurately reproduce the properties of sophisticated infrastructures like the Internet, which are subject to numerous technological constraints. Some of the fallacies that can arise when imprudently applying oversimplified complex network models to sophisticated technical infrastructures have been summarized for instance in [47].

Nevertheless, given that the statistical mechanics perspective on networks is able to reason about structural and dynamical properties that are relevant for the design and operation of overlay networks, it is reasonable to study how one can use models for complex networks *in a constructive rather than in an explanatory way*. Overlay management schemes can explicitly be designed based on a stochastic model that gives rise to a class of topologies whose properties are advantageous for a particular setting. By means of distributed probabilistic protocols – like for instance suitably configured random walk sampling or rewiring schemes – reproducing a stochastic model is often much simpler than implementing and debugging complex algorithms that precisely control the topology.

The perspective of statistical mechanics actually allows one to contrast this approach with traditional structured and unstructured overlays: From this point of view structured approaches give rise to states of small (statistical) entropy in the sense that only a small subset of all possible network realizations are accessible. This maximizes the amount of information one has about the detailed structures of the network. This information can then be used to design algorithms that utilize the network structure to provide efficient key lookups, routing, and information spreading. However, maintenance schemes are required to prevent a gradual loss of structure – and thus increase of entropy – due to the dynamics of users and devices, hard- and software failures or communication errors. As such, the overhead induced by topology maintenance mechanisms can be viewed in analogy to the input of energy that is used by nonequilibrium biological systems to prevent the entropy increase that is due to the second law of thermodynamics. Analogously, unstructured approaches can be related to states of maximum statistical entropy in which all network realizations are equally likely, as the topology is constructed in a completely uncoordinated fashion. In this case, distributed algorithms can not use *a priori* information about the detailed structures of the network structure and therefore flooding or exhaustive search are the only viable options.

One of the most interesting aspects of complex network science is that it allows one to explore the interesting middle-ground of *statistically* or *thermodynamically* structured topologies with intermediate levels of entropy. These networks are neither completely random nor deterministic. They rather introduce a statistical structure (like a particular type of degree sequence, a certain clustering coefficient, or correlations between data location and network structures) that facilitates adaptivity and that allows one to solve algorithmic tasks more efficiently than in unstructured systems. For distributed search in P2P systems, it has been shown in [39] that generic structures of randomly generated networks with a power law degree distribution can be exploited in order to improve search performance, while being oblivious to the details of the topology. Similar results have been obtained for routing schemes making use of correlations between node addresses and network structure, particular clustering structures, or an embedding into Euclidean or Hyperbolic coordinate spaces [13, 27, 36, 38, 44].

Enforcing Ensembles

In Sect. "Statistical Mechanics of Complex Networks", we commented on the rich body of results on collective properties like connectedness, diameter,

resilience against faults and attacks, as well as on the performance of dynamical processes like information dissemination, synchronization, consensus and gossiping schemes. Since these statements are derived based on statistical ensembles, they are necessarily stochastic. However, statements on properties that hold *asymptotically almost surely* become more reliable as the size of the network topology (in terms of the number of nodes) increases, just like statements on properties of thermodynamic systems become more reliable as the volume being considered is increased. This applies for instance to stochastic guarantees on the diameter or connectedness of overlay networks emerging from suitable stochastic processes. Again, this is due to the fact that it is becoming increasingly unlikely to generate a network realization from a set whose probability measure converges to zero under the given construction process[3]. This often contrasts with the kind of guarantees one can obtain for structured overlay networks.

So far, we have discussed the link between the abstraction of constrained adiabatic ensembles of networks and collective network properties. However, from a practical perspective this link is useless if we cannot relate the ensemble description of a system to the distributed processes shaping the overlay at the level of individual nodes. In real-world systems, one may be confronted with situations in which the actual stochastic dynamics constructing the overlay depend on external factors that cannot easily be influenced. Here one can often use tools from statistical mechanics (like e. g., master equations or mean-field approximations) to derive aggregate statistical quantities of interest – and thus the ensemble description of the system – from a stochastic description of individual nodes (see e. g., the application of this procedure in [3]). The situation is different when we wish to actually design a topology management scheme that constructs a network topology drawn from a particular constrained adiabatic ensemble which has desirable properties. In this case, one can often employ a *configuration approach* to analytically derive the local stochastic dynamics from the ensemble description. In the following we will briefly demonstrate this approach for the particular case of random scale-free networks. The actual distributed mechanism has been presented and evaluated in much detail in [40, 41].

The basic idea is to start with an arbitrary connected topology that has been generated for instance by a bootstrapping process. One can then progressively rewire all connections by means of a biased random walk sampling scheme like the one proposed in [48]. Again referring to [40, 41] for a more detailed algorithmic description of the protocol, each rewiring of an existing connection e between two nodes i and j in the topology is initiated by one of their endpoints. This node deletes the connection e and creates a sampling message that contains the addresses of i and j. By means of two consecutive biased random walks of length l, two new nodes v and w (the final nodes at which the two random walks reside after l steps) are sampled. In a distributed setting, the address of the first node v sampled by the first random walk can be incorporated in the sampling message passed along in the second random walk. In this case, the final node w has all information needed to establish a new overlay link $e' = (v, w)$ if it does not exist already. The process is illustrated in Fig. 1, which shows a random walk rewiring of edge e initiated by node 0. A first random walk takes three steps and finds node $v = 3$. After another 3 steps a second endpoint $w = 6$ is found and the new edge $e' = (3, 6)$ is created. From a statistical mechanics perspective, this scheme results in the fact that edges

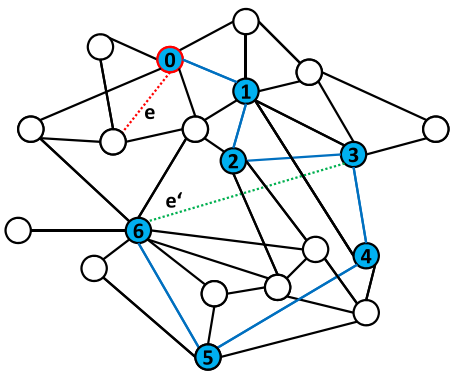


Fig. 1 Rewiring of edge e initiated by node 0 by means of two consecutive random walks with step length three. The red dotted edge is replaced by the green dotted one

[3] Since in practice one necessarily deals with finite-size systems, it is important to note that technically one also needs to consider how fast the associated probability converges as the network size increases. It is often possible to give analytical expressions for these so-called finite-size effects. While we refer the interested reader to [19] for more details, at this point it is sufficient to note that the properties mentioned above hold for networks whose size is reasonably small (of the order of a few hundred to a few thousand nodes) to be of practical value for the envisioned scenario.

mimic a random particle motion according to an engineered energy landscape which in this particular case is defined according to the adiabatic ensemble of scale-free networks.

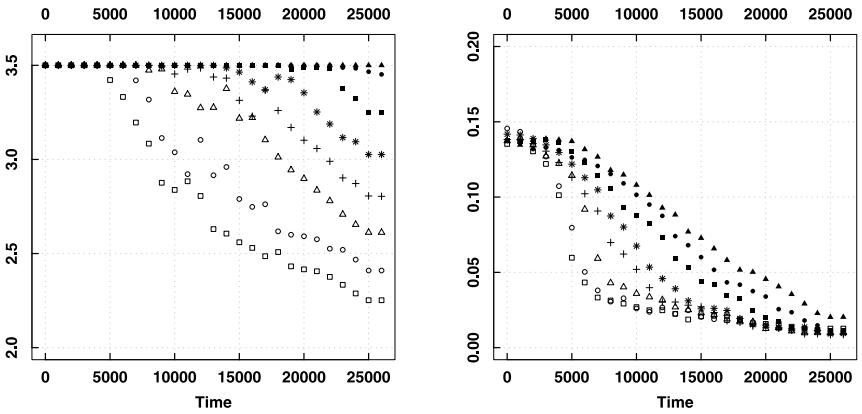
If the transition kernel of the random walk sampling is configured appropriately and the random walk is sufficiently long, the topology is guaranteed to be drawn from the desired adiabatic ensemble once all connections have been rewired. Here, the number of random walk steps needed to allow the system to equilibrate (i. e., forget about the arbitrary initial topology) is crucial to apply equilibrium arguments about the targeted ensemble. In experiments described in [40], we have found a length of $\log(n)$ steps to be sufficient where n is the number of nodes. An analytical approach that is based on the framework described in Sect. "Dynamical Processes in Complex Networks" can be used to derive an upper bound for the required random walk length (see details in [40]). On the basis of a configuration model approach for a constrained adiabatic ensemble of scale-free networks that was introduced in [28] as well as the Metropolis–Hastings algorithm [25] we can actually derive the required transition kernel for the random walk. In order to effectuate an ensemble with a particular exponent γ one has to configure the random walk such that each node i forwards a sampling message to a neighbor j with probability

$$P_{i,j} = \frac{d_i}{d_j}\left(\frac{i}{j}\right)^{\frac{1}{\gamma-1}} \qquad (2)$$

where i, j are numeric, not necessarily unique node identifiers chosen uniformly at random from an identifier space. The effect of the resulting rewiring process on the degree distribution is shown in Fig. 2 in terms of the evolution of the fitted degree distribution exponent as well as the Kolmogorov–Smirnov statistic D which quantifies the goodness of the fit. The decreasing values for D show that the hypothesis that the topology is indeed a random network with a power-law degree distribution becomes more reasonable. Since the network is sampled at random, for this simple topology management mechanism one can thus safely rely on all analytical results that hold for an adiabatic ensemble of random scale-free networks with the chosen exponent.

The Micro-Macro Link

An interesting aspect of the procedures described above is that they allow one to analytically link a stochastic model for the micro-scale dynamics of an overlay (at the level of individual nodes and protocol messages) with the collective properties that are observable at the system level. The three levels that naturally arise in this perspective are depicted in Fig. 3. At the microscopic level, one assumes a stochastic model that captures the behavior of individual nodes with respect to the way connections in the overlay are being constructed. In the previous section, we discussed that such a stochastic micro-model can be related to an adiabatic statistical ensemble – which defines the next level of description – and vice-versa. At the level of statistical ensembles, aggregate network statistics like average



(a) Average fitted exponent γ_f (b) Average Kolmogorov-Smirnov statistic D

Fig. 2 Time evolution of 5000 node networks during adaptation runs with $\gamma \in [2.1, 3.5]$

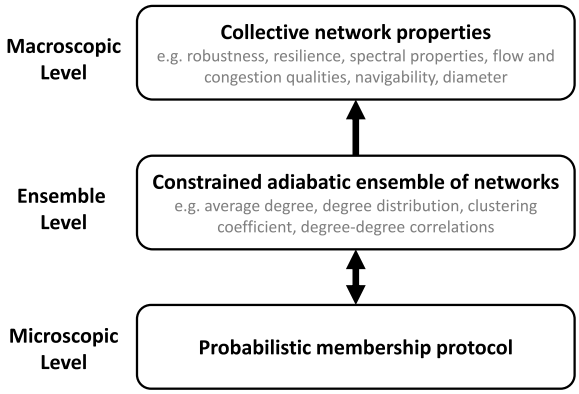


Fig. 3 Micro-macro link provided by a thermodynamic approach to the design of overlay networks

degree, system size, degree distributions or other correlations effectively serve as *thermodynamic quantities* that constrain the accessible states and thus determine the structural and dynamical properties of the system. By measuring the probability of network configurations with particular macroscopic characteristics, one can then make strong probabilistic statements about collective properties like diameter, connectedness, or spectral properties which emerge at the macro-level of the system.

At the intermediate level of statistical ensembles, rather than a single ensemble, one may further consider a set of ensembles with different constraints. By means of the micro-macro link exemplified in the previous section, it is possible to provide distributed adaptation strategies which switch between overlay topologies drawn from different ensembles based on the environmental conditions. Reconsidering the particular distributed sampling scheme exemplified above, the exponent parameter γ actually defines a point in a continuum of adiabatic ensembles of scale-free networks that can be tuned by changing the parameter in the random walk bias, thus effectively changing the stochastic connection rewiring dynamics.

Adaptation in Organic Overlays – Triggering Phase Transitions

While the possibility to switch between different ensembles by tuning a sampling bias can be useful *per se* (e. g., to switch between random and scale-free topologies), particularly interesting aspects that can be utilized are phase transition phenomena, i. e., the existence of critical points in the space of control parameters at which the macroscopic properties of the resulting topology change suddenly. This may involve smooth (*second-order*) or abrupt (*first-order*) changes as well as phase transitions with hysteresis effects. Such critical points are of primary interest for the complex networks community and a number of analytical results have been obtained that relate these points to a qualitative change of macroscopic network properties. For scale-free networks, such phenomena have been studied in detail. When considering a continuum approximation of scale-free networks (justified in the thermodynamic limit of infinite systems), the probability for a node to have exactly k connections is given by a Zeta (also called *scale-free*) distribution,

$$P(k) = \frac{k^{-\gamma}}{\zeta(\gamma)}$$

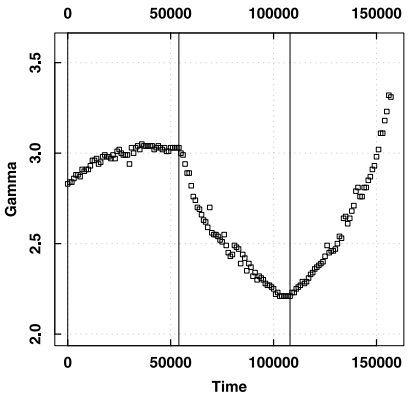
with $\zeta : \mathbb{R} \to \mathbb{R}$ being the real-valued Riemann Zeta function, $\zeta(\gamma) = \sum_{i=1}^{\infty} i^{-\gamma}$. Many results for critical phenomena in networks with fixed degree distributions are due to the so-called Molloy–Reed criterion [32] which links the relation of the distribution's first two moments to the existence of a giant connected component. This has been used successfully to study – as mentioned before – the error and attack tolerance of random power-law networks. For this special case, in [16] it was found that at least a fraction

$$r := 1 - \left(\frac{\zeta(\gamma - 2)}{\zeta(\gamma - 1)} - 1 \right)^{-1}$$

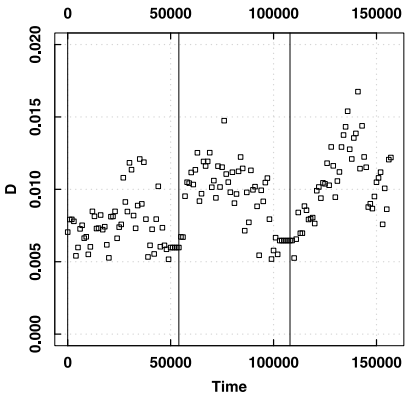
of nodes need to fail at random for the giant connected component of a power-law network to be destroyed. Regarding the control parameter γ – which determines the behavior of the distributed rewiring scheme described in Sect. "Enforcing Ensembles" – the convergence behavior of the Zeta function results in a phase transition once the parameter crosses the critical point $\gamma = 3$. For $\gamma \geq 3$, the terms $\zeta(\gamma - 2)$ and $\zeta(\gamma - 1)$ are constants, thus resulting in a constant nonzero value for the critical fraction r. For $\gamma \in (2, 3)$, the term $\zeta(\gamma - 2)$ diverges. In this range $r \to 1$, i. e., almost all nodes can be removed at random without destroying the giant connected component. The parameter γ thus effectively allows one to produce a continuum of constrained adiabatic ensembles while at the critical point $\gamma = 3$, the resilience properties of the resulting networks undergo a qualitative change.

The same argument about the change in the convergence behavior of the Zeta function has actually been used to derive phase transitions in terms of diameter, attack resilience [17] or the efficiency of spreading phenomena [33]. Given that – depending on the used distributed algorithms, as well as the current operational conditions – these properties can be both desirable or detrimental. So, the knowledge about these effects can actually be used for an active adaptation of collective network qualities. For the random walk rewiring described in Sect. "Enforcing Ensembles", all one has to do is to change the control parameter γ in (2). Figure 4 shows simulation results that have been obtained based on this idea and that were originally presented in [40]. Here, the random rewiring protocol described above has been applied continuously, while at certain times (indicated by vertical lines in Fig. 4a and b) the parameter γ was changed. After each modification, the network was allowed to equilibrate by progressively resampling all connections according to the new constrained adiabatic ensemble. Figure 4a shows the evolution of the fitted degree distribution exponent γ. The Kolmogorov–Smirnov statistic D depicted in Fig. 4b shows the goodness of the assumption that the connectivity distribution indeed follows a power law with the fitted exponent. Smaller values for D again represent a larger reliability of the fit. Here, it can clearly be seen that near the end of the adaptation cycles (which was at the same time the beginning of a new one with a different γ), the degree distribution of the network does indeed follow a power law.

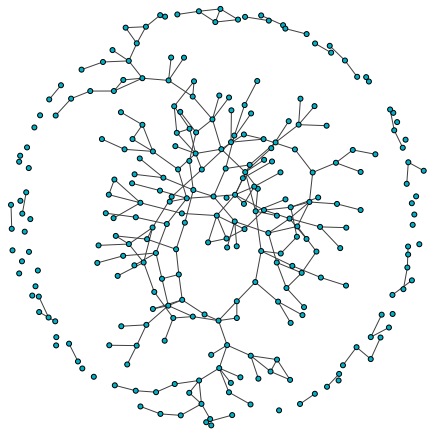
On the basis of the theoretical findings regarding the critical point of $\gamma = 3$ for the robustness of the topology against attacks, one would expect that the robustness changes qualitatively between the last two adaptation rounds, i. e., when the ex-



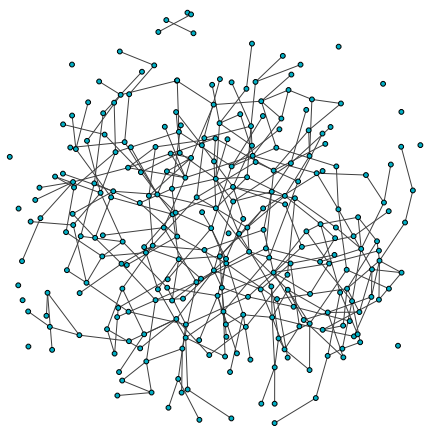
(a) Power Law Exponent γ



(b) Kolmogorov-Smirnov statistic D



(c) Result of attack for network with $\gamma \approx 2$



(d) Result of attack for network with $\gamma \approx 3.5$

Fig. 4 Time Evolution of scale-free network during multiple adaptation cycles. Start/end times of adaptation cycles are indicated by vertical lines

ponent of the topology is increased from 2. 1 to 3.5. To exemplify this theoretical result, Fig. 4c,d shows the network topology that remains after 10% of the most connected nodes have been removed from a 300 node network at these two different points in the adaptation process. The large number of isolated nodes and clusters in Fig. 4c and the comparison with the topology shown in Fig. 4d clearly show the practical implications of this theoretical finding in terms of resilience. In summary, this particular phase transition effect in equilibrium ensembles of scale-free networks can be used to make a trade-off between desirable and detrimental properties. In phases where efficient spreading is needed, the topology can be sampled from an adiabatic ensemble with $\gamma \in (2, 3)$. Similarly, according to [39] $\gamma \in (2, 2.3)$ should be chosen to maximize the efficiency of a random walk-based distributed search in scale-free topologies. In situations where attacks or a spreading of failures are being detected, the local connection sampling can be instrumented such that the resulting topology is much more resilient against these effects, while at the same time reducing the efficiency of distributed search and information dissemination.

Conclusion and Outlook

In this article, we have summarized the statistical mechanics and the statistical physics perspective on the modeling of complex network structures. We then outlined some ideas on how this perspective can constructively be used in the management of overlay networks for very large, dynamic systems. In particular, we argue that – at least in very large and highly dynamic systems – it can be easier to enforce a particular statistical ensemble which will give rise to desirable macro-level properties and performance of distributed schemes than using sophisticated topology maintenance schemes. By means of active randomization of protocols, one can thus obtain strong, *thermodynamic* guarantees for the structure and dynamics emerging in sufficiently large systems. For the design of robust and adaptive organic computing systems, we thus argue that randomization and *loosening precise control* are crucial ingredients.

While first examples of actual distributed mechanisms that allow one to actively use some of the intriguing results of complex network science have been given, this work is necessarily incomplete. For the future, we envision for instance mechanisms that make use of the natural dynamics of Peer-to-Peer systems for the efficient construction and adaptation of an overlay topology with complex, yet predictable structures and properties. For this, rather than actively rewiring connections, one can use the natural turnover of machines and users (an effect usually called churn) and apply random connection sampling schemes only as nodes join the system, machines change their characteristics, or existing connections fail. By assuming that the system is in a state of equilibrium, the collective properties of the overlay network can then be predicted in analogy to the analysis of many-particle systems by means of statistical mechanics. Moreover, within this framework it may be possible to reduce the costs induced by structure maintenance protocols massively, thus making such an approach suitable for large and dynamic systems. Referring again to the scenario depicted in Sect. “Introduction”, one may thus be tempted to summarize the challenges of future systems, as well as the idea of addressing them in the framework presented in this article in the following way:

> *As network devices become more akin to particles in terms of number, size, and stochastic behavior, can we design distributed systems using models, methods, and abstractions from statistical mechanics and thermodynamics?*

Considering recent advances in the statistical physics’ study of complex systems in general, and complex networks in particular, one can argue that this is true at least for some aspects of large dynamic systems. An interesting prospect of this perspective is the fact that – in contrast to particles in thermodynamic systems – we can actually program devices to change their local dynamics in a meaningful way, for instance to actively trigger transitions between phases in which the network is more resilient or allows for more efficient spreading or distributed search.

Given the current surge of interest in the statistical mechanics approach to the modeling of complex networks, we thus think that it is reasonable to foresee a number of interesting applications in the engineering of organic computing systems and in the design of mechanisms for sustainable future techno-social systems.

References

1. Adamic LA, Lukose RM, Puniyani AR (2001) Search in power-law networks. Phys Rev E 64(4):046135
2. Albert R, Barabási A-L (2002) Statistical mechanics of complex networks. Rev Mod Phys 74(1):47–97
3. Barabási A-L, Albert R, Jeong H (1999) Mean-field theory for scale-free random networks. Physica A 272(1–2):173–187
4. Almendral JA, Díaz-Guilera A (2007) Dynamical and spectral properties of complex networks. New J Phys 9(6):187–187
5. Arenas A, Díaz-Guilera A, Kurths J, Moreno Y, Zhou C (2008) Synchronization in complex networks. Phys Rep 469(3):93–153
6. Babaoglu Ö, Binci T, Jelasity M, Montresor A (2007) Firefly-inspired heartbeat synchronization in overlay networks. In: First International Conference on Self-Adaptive and Self-Organizing Systems, SASO'07, Boston, Mass., USA, July 9-11, pp 77–86, IEEE, 2007
7. Balakrishnan H, Kaashoek MF, Karger D, Morris R, Stoica I (2003) Looking up data in P2P systems. Commun ACM 46(2):43–48
8. Baldoni R, Corsaro A, Querzoni L, Scipioni S, Piergiovanni ST (2009) Coupling-Based Internal Clock Synchronization for Large-Scale Dynamic Distributed Systems. IEEE T Parall Distrib 99(RapidPosts):607–619
9. Barahona M, Pecora LM (2002) Synchronization in small-world systems. Phys Rev Lett 89(5):54101
10. Barrat A, Barthélemy M, Vespignani A (2008) Dynamical processes on complex networks. Cambridge University Press, New York, NY, USA
11. Berg J, Lässig M (2002) Correlated random networks. Phys Rev Lett 89(22):228701
12. Boccaletti S, Latora V, Moreno Y, Chavez M, Hwang DU (2006) Complex networks: Structure and dynamics. Phys Rep 424(4-5):175–308
13. Boguñá M, Krioukov D, Claffy KC (2008) Navigability of complex networks. Nat Phys 5(1):74–80
14. Bollobas B (2001) Random graphs. Cambridge Univ. Press, Cambridge
15. Cohen JE (1988) Threshold Phenomena in Random Structures. Discrete Appl Math 19(1–3):113–128
16. Cohen R, Erez K, Ben-Avraham D, Havlin S (2000) Resilience of the internet to random breakdowns. Phys Rev Lett 85(21):4626–4628
17. Cohen R, Erez K, Ben-Avraham D, Havlin S (2001) Breakdown of the Internet under intentional attack. Phys Rev Lett 86:3682
18. Demers A, Greene D, Hauser C, Irish W (1987) Epidemic algorithms for replicated database maintenance. In: Proceedings of the sixth annual ACM Symposium on Principles of distributed computing, Vancouver, BC, Canada, August 10–12, 1987, pp 1–12
19. Dorogovtsev SN, Ferreira Mendes JF, Samukhin AN (2003) Principles of statistical mechanics of random networks. Nucl Phys B 666:396
20. Erdős P, Rényi A (1959) On random graphs I. Publ Math Debrecen 6(290–297):156
21. Farkas I, Derényi I, Palla G, Vicsek T (2004) Equilibrium Statistical Mechanics of Network Structures. Lect Notes Phys 650:163–187
22. Floyd S, Jacobson V (1994) The Synchronization of Periodic Routing Messages. IEEE/ACM T Netw 2:122–136
23. Park J, Newman MEJ (2004) Statistical mechanics of networks. Phys Rev E 70(6): 066117, http://pre.aps.org/abstract/PRE/v70/i6/e066117
24. Gupta I, Kermarrec AM, Ganesh AJ (2006) Efficient and adaptive epidemic-style protocols for reliable and scalable multicast. IEEE T Parall Distrib 17(7):593–605
25. Hastings WK (1970) Monte Carlo sampling methods using Markov chains and their applications. Biometrika 57(1):97
26. Jelasity M, Montresor A, Babaoglu Ö (2005) Gossip-based aggregation in large dynamic networks. ACM T Comput Syst 23(3):219–252
27. Kleinberg JM (2006) Complex Networks and Decentralized Search Algorithms. In: International Congress of Mathematicians (ICM), 22–28 August 2006, Madrid, Spain
28. Lee D-S, Goh K-I, Kahng B, Kim D-H (2004) Evolution of scale-free random graphs: Potts model formulation. Nucl Phys B 696:351–380
29. Lucarelli D, Wang I-J (2004) Decentralized synchronization protocols with nearest neighbor communication. In: SenSys '04: Proceedings of the 2nd international conference on Embedded networked sensor systems, New York, NY, USA, ACM, pp 62–68
30. Mattern F, Flörkemeier C (2010) Vom Internet der Computer zum Internet der Dinge. Informatik-Spektrum 33(2):107–121
31. Mehta ML (1991) Random Matrices, 2 edn. Academic Press, New York
32. Molloy M, Reed B (1995) A critical point for random graphs with a given degree sequence. Random Struct Algor 6(2–3):161–180
33. Moreno Y, Nekovee M, Vespignani A (2004) Efficiency and reliability of epidemic data dissemination in complex networks. Phys Rev E 69(5):055101
34. Newman MEJ (2003) The structure and function of complex networks. SIAM Rev 45(2):167–256
35. Noh JD (2004) Random Walks on Complex Networks. Phys Rev Lett 92(11):4
36. Papadopoulos F, Krioukov D, Boguna M, Vahdat A (2010) Greedy Forwarding in Dynamic Scale-Free Networks Embedded in Hyperbolic Metric Spaces. In: Proceedings of the IEEE Infocom Conference, March 19, 2010, San Diego, CA, USA
37. Reichardt J, Bornholdt S (2006) Statistical mechanics of community detection. Phys Rev E 74(1):1–14
38. Sandberg O (2006) Distributed routing in small-world networks. In: Proceedings of the eighth Workshop on Algorithm Engineering and Experiments and the third Workshop on Analytic Algorithmics and Combinatorics, Society for Industrial Mathematics, January 21, 2006, Miami, Florida, p 144
39. Sarshar N, Boykin PO, Roychowdhury VP (2004) Percolation search in power law networks: Making unstructured peer-to-peer networks scalable. In: Proceedings of the Fourth International Conference on Peer-to-Peer Computing, IEEE Computer Society, 25–27th August 2004, Zurich
40. Scholtes I (2010) Distributed Creation and Adaptation of Random Scale-Free Overlay Networks. In: Proceedings of the Fourth IEEE Conference on Self-Organizing and Self-Adaptive Systems (SASO), September 27–October 1, 2010, Budapest, Hungary, IEEE, pp 51–63
41. Scholtes I (2011) Harnessing Complex Structures and Collective Dynamics in Large Networked Computing Systems. Dissertation, University of Trier
42. Scholtes I, Botev J, Esch M, Sturm P (2010) Epidemic Self-Synchronization in Complex Networks of Kuramoto Oscillators. Adv Complex Syst 13(1):33–58
43. Scholtes I, Kolos S, Zema PF (2008) The ATLAS Event Monitoring Service – Peer-to-Peer Data Distribution in High-Energy Physics. IEEE T Nucl Sci 55(3):1610–1620
44. Thadakamalla HP, Albert R, Kumara SRT (2007) Search in spatial scale-free networks. New J Phys 9(6):190–190
45. Waldhorst OP, Bless R, Zitterbart M (2010) Overlay-Netze als Innovationsmotor im Internet. Informatik-Spektrum 33(2):171–185
46. Weiser M (1991) The computer for the 21st century. Scientific American
47. Willinger W, Alderson DL, Doyle JC (2009) Mathematics and the Internet: A Source of Enormous Confusion and Great Potential. Internet Res 56(5):586–599
48. Zhong M, Shen K, Seiferas J (2008) The Convergence-Guaranteed Random Walk and Its Applications in Peer-to-Peer Networks. IEEE T Comput 57(5):619–633