

**Dynamic model of time-dependent complex networks**

Scott A. Hill

*Department of Physics, University of Toledo, Toledo, Ohio 43606, USA*

Dan Braha

*New England Complex Systems Institute, Cambridge, Massachusetts 02138, USA**and University of Massachusetts, Dartmouth, Massachusetts 02747, USA*

(Received 14 July 2010; published 12 October 2010)

The characterization of the “most connected” nodes in static or slowly evolving complex networks has helped in understanding and predicting the behavior of social, biological, and technological networked systems, including their robustness against failures, vulnerability to deliberate attacks, and diffusion properties. However, recent empirical research of large dynamic networks (characterized by irregular connections that evolve rapidly) has demonstrated that there is little continuity in degree centrality of nodes over time, even when their degree distributions follow a power law. This unexpected dynamic centrality suggests that the connections in these systems are not driven by preferential attachment or other known mechanisms. We present an approach to explain real-world dynamic networks and qualitatively reproduce these dynamic centrality phenomena. This approach is based on a dynamic preferential attachment mechanism, which exhibits a sharp transition from a base pure random walk scheme.

DOI: [10.1103/PhysRevE.82.046105](https://doi.org/10.1103/PhysRevE.82.046105)

PACS number(s): 89.75.Hc, 05.40.Fb, 89.65.-s, 89.75.Da

**I. INTRODUCTION**

The modern study of social networks largely presupposes static or slowly evolving complex networks, as clearly indicated by an extensive review of the pertinent literature [1–6]. In the social and behavioral sciences, this static perspective of networks has generated a wide spectrum of formal representations and tools for measuring structural and locational properties of social networks (e.g., centrality and prestige), social network roles and positions (e.g., structural equivalence), and local (subgraph) structures. Statistical models of static networks are prevalent [1,2]. The study of networks has been developed beyond the social sciences by a substantial number of papers that have extended the study of universal properties in physical systems to complex networks in social, biological, and technological systems; these recent advances have generated an extensive research effort in understanding the effect of static connectivity patterns on various dynamical processes occurring on top of networks [7]. One of the most important concepts in complex network studies is *network centrality*, which characterizes the “most connected” nodes in static complex networks [1,2]. Among other things, the concept of network centrality has led to the study of “scale-free networks” [8], where the *degree centrality* of nodes is distributed according to a power-law or other long-tail distribution, implying the existence of highly connected nodes called *hubs* [4,8]. The leading explanation for this uneven distribution of connectivity, present in many real networks, is based on growth and preferential attachment mechanisms, in which the structural changes of real networks are governed by the dynamical evolution of the system (see [8] and its many variants [9]). The long-tail characterization of complex networks has helped in understanding and predicting the behavior of social, biological, and technological networked systems, including their robustness against failures [10,11], vulnerability to deliberate attacks [10,12], and diffusion properties [7,13]. This static network centrality

perspective has also led the discussion of strategies for immunizing particular individuals to reduce disease propagation [14–17], advertising to “opinion leaders” [18], and targeting central individuals in terrorist networks [19].

The analysis of *evolving* social networks has long been regarded as a “Holy Grail” of network scientists [1,2,20]. A number of models, mostly theoretical with limited empirical backing, have been considered for analyzing longitudinal social network data [21,22], with networks that change either at discrete time points [23–26] or continuously over time [27–29]. The complex network community has also been interested in the analysis of these dynamic networks; their research has mainly examined various mechanisms of network formation (e.g., “preferential attachment” and its variants [3,9]), and the evolution of aggregate properties of the network at the macroscopic and mesoscopic levels (e.g., network size, average degree, small-world characteristics, or community structure [30,31]).

While some of the above studies allow their networks to change, in most cases the evolution is assumed to happen more slowly than any dynamical processes occurring on the network. Real network connections, however, are often more fluid than this; even when there exists an underlying fixed topological structure, connections between nodes can adaptively become active or inactive over time [20]. Consideration of the dynamical interplay between the state and the topology of the network is pertinent to a wide variety of network types. Examples include travel on a transportation network, message exchanges over communication networks, interactions between molecules that can bind to each other, and the switching of genes “on” and “off”: all are networks with an underlying structure that can change quickly and adaptively with respect to the state of the network, and whose changes are relevant to the behavior and functionality of the system. The generality and broad implications of adaptive and dynamical wirings in real complex networks has been leading to a rapidly growing study and much interest

from social scientists, physicists, and biologists (see reviews of recent advances in [32,33]). This mostly theoretical research has revealed a number of new approaches and insights including self-organization of adaptive networks [34–36], contact processes and epidemiology on adaptive networks [37–39], and social games on adaptive networks [20,40,41].

Despite these important theoretical advances, there have been very few empirical studies related to the dynamics of large-scale network connections, or theoretical models that are validated and informed by empirical results [42–44]. Clearly, the study of data collected from real-world time-dependent networks, accompanied by theoretical models which address and reproduce the properties of said networks, will provide new, and perhaps fundamental insights into the mechanisms of dynamic networks. For example, the experiments detailed in Ref. [42] detected a dramatic time dependence in network centrality and the role of nodes, something that is not apparent from static (i.e., time-aggregated) analysis of node connectivity and network topology. These experiments studied the time-dependent structure of a large-scale email network of over 57 000 users, based on data gathered over 113 days from log files maintained by the email server at a large university. They found, among other things, that the daily networks (as well as the total aggregate network) were scale-free, but the “hubs” (that is, the most well-connected nodes) of these networks changed from day to day. In other words, a node which is among the most connected on one day, might only have a couple links the following day (popular today, anonymous tomorrow), and may not even be among the hubs of the aggregate network. This phenomenon, called *dynamic centrality*, is not accounted for by the standard explanation of hub formation: normally, hubs in networks develop in a system where popular nodes tend to attract more links (a process called *preferential attachment*), so that when one particular node becomes slightly more popular than the rest, it attracts slightly more links from other nodes, making it more popular, and so forth in a process of positive feedback. This process typically works over a long period of time, as news of a node’s new “popularity” is allowed to disseminate through the system before the next connection is made. In Ref. [42], however, hubs develop over the course of a single day, and vanish just as quickly: too fast for such a process to occur. Thus, the implications of dynamic centrality are in sharp contrast to previous complex network research, which depend on such preferential attachment-type mechanisms to create power-law distributions in empirical complex networks. This result draws attention to the inherent dynamic nature of adaptive networks which has long been neglected in the literature [32]. Similar results [43] were found in the interactions caused by the spatial proximity of personal Bluetooth wireless devices, recording the interactions between pairs of students over the period of 31 days. Results that are consistent with dynamic centrality behavior have also been seen in the social dynamics of free and open-source (FLOSS) development projects [45], peer prestige in academic hiring networks [46], trading networks [47], protein-interaction networks [48], and transcriptional networks [49]. The availability of complete (i.e., nonsampled) data that tracks the temporal dynamics of the interactions within the network will allow further study of dynamic centrality.

The existence of dynamic centrality and a dynamic interaction structure has strong implications for existing work related to the attack and error tolerance of complex networks as well as in network transport. Research in epidemics has suggested that an effective disease or computer-virus prevention strategy would be to identify and vaccinate the high-degree nodes of a network, which would inhibit the spread of infection [14–17]. Similarly, “popular influencer” marketing techniques (closely related to word-of-mouth or viral marketing) are based on the premise that focusing marketing activities on the hubs of social networks increases the likelihood of a cascading adoption of products or services, forming a type of social epidemic [18]. The underlying assumption of both models is that the network topology is basically static: a node which is well-connected or popular now will continue to be a hub later in time. In a dynamic network, however, well-connected nodes can quickly become only weakly connected (or even disconnected) tomorrow, which invalidates such strategies. Thus, the existence of dynamic centrality calls for a radical rethinking of the interplay between link/structural dynamics and the dynamical processes underlying the time-dependent complex networks, and may call for modified strategies in predicting and preventing (or encouraging) epidemics.

Our paper contributes to the understanding of adaptive and dynamic networks by demonstrating an approach to explain and reproduce real world dynamic centrality phenomena. More specifically, we will investigate the type of network seen in Refs. [42,43], which we will call *dynamic scale-free* (DSF) networks; these are a series of scale-free networks with a scale-free aggregate and dynamic centrality. We will show that DSF networks arise from a series of vertex-reinforced random walks on a finite scale-free network (called the *underlay*), governed by a “limited” preferential-attachment reinforcement rule—limited in the sense that, when making a connection, a node only takes into account the popularity of its neighbors in the underlay.

The study of random walks on graphs, and reinforced random walks in particular, has a long fascinating history in its own right [50,51]. In the context of social and complex networks, the nature of random walks and diffusion over small-world and scale-free networks is also a much investigated topic [52–59]. Random walks have been used in analyzing search and navigation on networks [60], calculating the “betweenness centrality” of nodes in a network [61], extracting community structure from a network [61], and sampling a large network as a means of estimating its topological properties [62,63]. The close interplay between the degree of a node and the proportion of time visited by a uniform random walk is a well-known result [50], and has also been examined for various real world networks [64]. This body of literature largely focuses on the behavior of *nonreinforced* random walks on static network topologies. In contrast to dynamics *on* networks, a different approach for modeling the dynamics *of* networks, which uses reinforced random walks or other stochastic processes that incorporate reinforcement mechanisms, has been considered [20,22]. Most pertinent to our study are models of dynamic networks where individuals choose whom to interact with depending on how “popular” their choice is [20,28]. Despite these theoretical advances,

application to real world dynamic networks has been limited to parameter estimation with very small networks that are observed at several time points [22]. Here we broaden the scope of this research in several distinctive ways. First, to the best of our knowledge, the use of a reinforced random walk has not yet been used as a mechanism of dynamical network connectivity that is informed and supported by actual high-frequency measurements of large-scale social network link dynamics. This shift in perspective leads to new insights including evidence that nonreinforced random walk schemes cannot serve as a plausible explanation for experimental data of dynamic networks, in contrast to reinforced random walk schemes. Second, random walk mechanisms (or other stochastic models) have not been used so far to study real-world dynamic centrality phenomena. Finally, new analysis methods, including various “ranking” and “centrality overlap” measures, are incorporated to facilitate and enable effective evaluation of models against large-scale data.

**II. MODEL**

In our approach we start with the aggregate network, which is called the *underlay*; it is reasonable to assume that the aggregate network represents the long-term pattern of contact between individuals, which is often well approximated by a power-law distribution [3]. Underlays are generated using the Barabási-Albert algorithm [8] with parameters  $m_0=5$  (initial number of nodes; also the minimum degree of every node in the underlay) and  $N=1\,000\,000$  (the size of the network). We then generate a series of DSF networks using a vertex-reinforced random walk upon this underlay network: starting from a randomly chosen node, we add weight (according to a scheme described below) to the vertices every time they are visited, constantly altering the jump probabilities so that the walk will tend to revisit vertices already visited [20,51]. This proposal is inspired by the way messages on actual communication networks are part of larger-scale conversations. For instance, suppose Alice sends an email to Bob, who then responds to Alice and also forwards her message to Claire. Claire could in turn either forward it to someone new, reply to either Alice or Bob, or do nothing. Assuming these actions are roughly equal in likelihood, there is a greater probability that Claire sends an email to either Alice or Bob, than to some other specific individual; thus, people who are already part of the email chain are more likely to become part of the chain again. Here, each “daily” network is a subnetwork of the underlay, built on top of the underlay via the node-reinforced random walk mechanism.

The specific algorithm is as follows: starting with a randomly chosen node, we consider that node’s neighbors  $i$  in the underlay, each neighbor being weighted according to

$$w_i = 1 + CV_i, \tag{1}$$

where  $V_i$  is the number of visits the node has received so far, and  $C$  is a parameter of the model. Neighbor  $i$  is then chosen as a target with probability  $\Pi_i = w_i / \sum_j w_j$ . Once a target is chosen, the link between the current node and the target is added to the sub-network, the algorithm steps to the target, and the process continues until the chain reaches the pre-

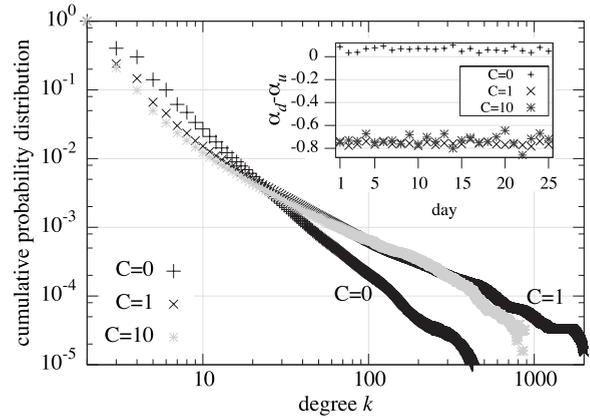


FIG. 1. A log-log plot of the cumulative degree distributions of daily networks, averaged over 25 days, built with single chains of 1 million steps: each histogram shows the probability that a member node of the daily network has degree  $k$  or larger. All curves show power-law behavior; the inset shows the difference between the power-law exponent of each daily network’s histogram, and the exponent of its underlay (approximately  $-1.9$ , though a different underlay was used for each value of  $C$ ).

scribed number of steps  $S$ . The parameter  $C$  captures the tendency of the random walk to revisit vertices—that is, the strength of dynamic preferential attachment. When  $C=0$ , one has a pure (i.e., unreinforced) random walk, a case we will refer to as a control, to demonstrate the importance of preferential attachment to our results. It should be noted that the vertex-reinforced random walk is known [51] to have significantly different behavior from the pure random walk. For example, it is well known that in a pure random walk the probability of visiting a particular node becomes proportional to the degree of the node, after  $\log N$  steps. This is not so for vertex-reinforced random walks which can become “trapped” within a particular subnetwork [51]: nodes within the subnetwork are visited frequently while nodes outside the subnetwork, even high-degree nodes, remain unvisited. This effect is critical to this model’s ability to recreate dynamic centrality, but from a practical point of view it can be too powerful, trapping the walk in a very small subnetwork (even one of only two points). To mollify this effect somewhat, we modify the vertex-reinforced rule to prohibit links from being traversed two times in a row: nodes and links can be revisited multiple times during the simulation, but a walk may not jump immediately backward to the node it came from.

**III. RESULTS AND DISCUSSION**

We will now show that the above reinforcement scheme is able to create scale-free networks with dynamic centrality and a scale-free aggregate (i.e., the underlay, scale-free by construction), reproducing the main qualitative features of DSF networks seen in [42,43]. We begin by constructing daily subnetworks with  $S=1$  million steps atop underlays of  $N=1$  million nodes. Figure 1 shows the cumulative probability degree distribution (averaged over 25 daily runs) for the cases of no ( $C=0$ ), weak ( $C=1$ ), or strong ( $C=10$ ) reinforce-

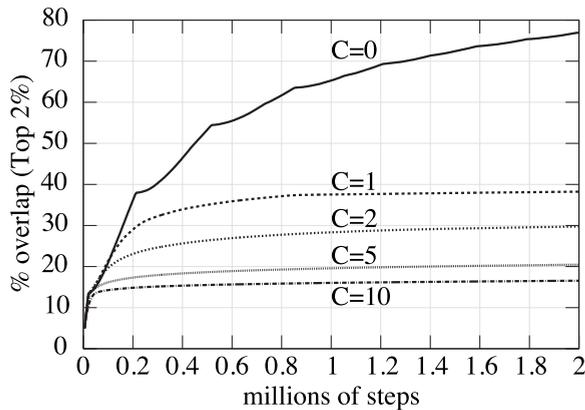


FIG. 2. The Top-2% overlap, as defined in the text, as a function of the number of steps  $S$  in the daily network, for various values of  $C$ . The overlap for  $C > 0$  appears to increase at a much slower rate, suggesting that the highly connected nodes in the underlay are only modestly represented in the daily networks.

ment. All three distributions show power-law behavior. However, the networks with reinforcement ( $C > 0$ ) have smaller slopes and longer tails, which suggests a more heterogeneous degree distribution and a more pronounced scale-free behavior. Furthermore, if we fit the histograms to a power-law  $Ax^{-\alpha}$ , the exponent  $\alpha$  for  $C=0$  is very close to the power-law exponent ( $\alpha_u \approx 2$ ) of the underlay [65], suggesting that the random-walk network is merely taking on the scale-free nature of the underlay. The inset shows this more clearly: considering 25 daily networks constructed for each value of  $C$ , we fit the cumulative histogram to a power law, found the exponent  $\alpha_d$  for each, and subtracted the exponent  $\alpha_u$  of that network's underlay. (In this case, we used three different underlays, one for each value of  $C$ .) We see that the  $C=0$  curve decays at the same rate as the underlay (or even a little more quickly); the difference is  $\alpha_d - \alpha_u = 0.06 \pm 0.02$ . By contrast, the  $C > 0$  cases have  $\alpha_d - \alpha_u = -0.76 \pm 0.01$  (for  $C=1$ ) and  $-0.73 \pm 0.04$  (for  $C=10$ ). With reinforcement, the daily exponents are smaller by about 0.75, resulting in a slower rate of decay and relatively more nodes of high degree. Note also that these results are relatively consistent from day to day.

Having shown that the subnetworks created by our model are scale-free subnetworks of a scale-free aggregate, we now demonstrate that they exhibit dynamic centrality. Following [42,43], we define the “Top- $n$  overlap” for two subnetworks of a given underlay as follows: after ranking the nodes according to their degree in each subnetwork, we count the number of nodes which appear among the top  $n$  nodes in both networks, dividing the count by  $n$  to express it as a fraction (or percentage). If the overlap is close to 100%, then the two subnetworks' hubs are largely identical; thus we expect that daily subnetworks exhibiting dynamic centrality (where hubs change from day to day) have a relatively low Top- $n$  overlap, both with the underlay and with each other. (The exact method and formula for calculating the overlap, which deals with the issue of ties for  $n$ th place, is described in the Appendix.)

We first used this “top  $n$  overlap” to compare individual daily networks with their underlays. In Fig. 2, we see that the

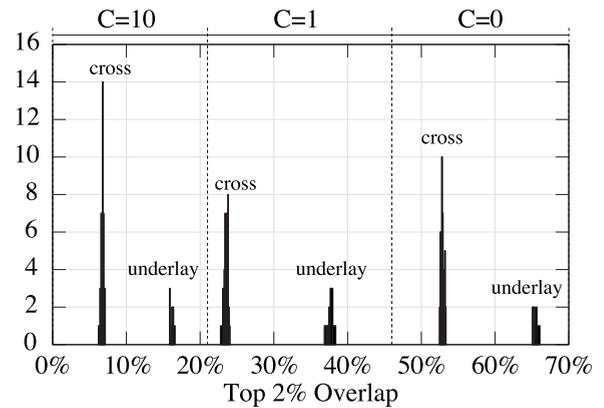


FIG. 3. The distributions of the Top-2% overlap between 10 daily networks and their common underlay (the peaks labeled “underlay”), and between the 45 pairs of those ten networks (labeled “cross”). The bar graphs were constructed by binning, with a bin size of 0.01%. The results suggest that “local” hubs vary greatly from day to day.

overlap of the top 2% of nodes (i.e., the 20 000 nodes with the highest degree) is dramatically smaller with the introduction of dynamic preferential attachment, once the walk extends further than roughly 200 000 steps, and the overlap decreases monotonically as the preference weight  $C$  increases. While the overlap of the  $C=0$  case grows steadily as the random walk fully explores the underlay, the overlap for  $C > 0$  increases at a much slower rate, suggesting that the highly connected nodes in the underlay are only modestly represented in the daily networks.

In Fig. 3, we look at the distribution of the Top-2% overlap between the underlay and 10 different daily subnetworks, generated with walks of  $S=1$  million steps, and for  $C=0, 1$ , and 10; we also look at the Top-2% overlap for the 45 pairs one can make among the ten daily networks for each value of  $C$ . All the distributions are sharply peaked, showing that the overlap does not vary much from day to day, and the overlap is again greatly decreased by the introduction of reinforcement, even from day to day. This demonstrates that the set of hubs of the daily network varies greatly from day to day, a result consistent with empirical dynamic scale-free networks [42,43].

An explanation for dynamic centrality in the  $C > 0$  case can be found by examining the number of unique nodes which are visited by the random walk, which is equivalent to the number of nodes in the resulting daily network, as is done in Fig. 4. For the pure random walk ( $C=0$ ), the underlay network (due to its small-world nature) is rapidly covered as the number of steps  $S$  increases, as is seen in the inset. The Top-2% overlap (shown in the main portion of the figure) is closely correlated with the number of nodes in the daily network, so as the daily network grows, it increasingly adopts the underlay's hubs as its own, ruining dynamic centrality because the hubs remain constant from day to day. With reinforcement ( $C > 0$ ), however, the random walk is more likely to revisit nodes it has visited before, and so it takes much longer to explore the underlay. This slow rate of exploration appears to be the primary reason that the overlap is suppressed in Fig. 2; in fact, Fig. 4 shows that if we compare

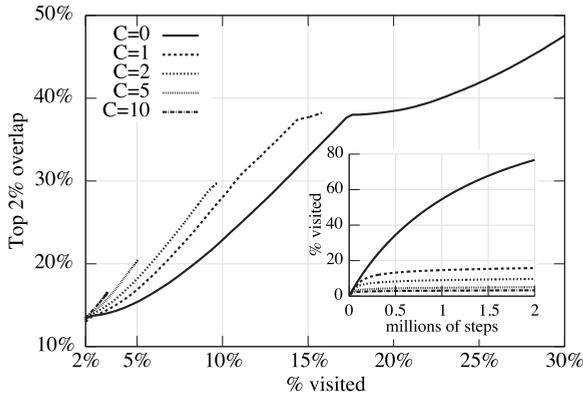


FIG. 4. The importance of the number of unique visited nodes. The inset shows the fraction of nodes in the underlay that are included in a daily network generated with different numbers of steps  $S$  and with  $C=0,1, 2, 5,$  and  $10$ . Higher percentages of unique visited nodes correspond monotonically to lower values of  $C$ . The main graph shows the same curves plotted versus the Top-2% overlap instead of the number of steps. The kink in the  $C=0$  curve is a finite-size effect.

daily networks with the same number of *nodes*, rather than the same number of steps, reinforcement actually results in a slightly *higher* overlap.

We have so far looked at the behavior of the top 20 000 nodes in the underlay and their modest role in the daily networks, but we can also compare the rankings of *all* nodes in the underlay with their ranking in the daily subnetworks. We did so as follows: for a given underlay, we ranked each node according to its degree, with the most highly connected node being given rank 0; nodes with the same degree were grouped together and given a rank which was the average of the range they covered. We then generated 10 daily subnetworks of  $S=1$  million steps each, and averaged the rank each node had over all ten subnetworks, dealing with ties in the same way as above. We did this for  $C=0$  and  $C=10$ , and Fig. 5 shows the underlay rank plotted versus the average daily rank so calculated. When multiple nodes have the same underlay rank, the average daily rank has itself been averaged over *all* nodes with the same underlay rank. In the  $C=0$  case, we see that the underlay and daily ranks have a strong positive correlation: the daily rank of a node always depends strongly on its rank in the underlay. The  $C=10$  curve is much different: the average daily rank is roughly *constant*, once we consider nodes below the top 50 000 (5%) or so. In addition to the average daily rank, we can also consider the *best* (i.e., minimum) daily rank that a node reaches over the ten daily networks; in doing so, we see (in the inset to Fig. 5) that *every* node in the  $C>0$  case has an opportunity to be among the top 1500 nodes (0.15%) in *some* daily network; without preferential attachment, most nodes cannot crack the top 10 000 (1%). The  $C>0$  case is extremely *egalitarian* in this regard: any node, even those with the smallest degree in the underlay, can be among the most important nodes in some daily network. This behavior agrees with the empirical results reported in [42,43].

One last empirical result seen in [42] involves the aggregate of  $M$  daily subnetworks, which are formed by connect-

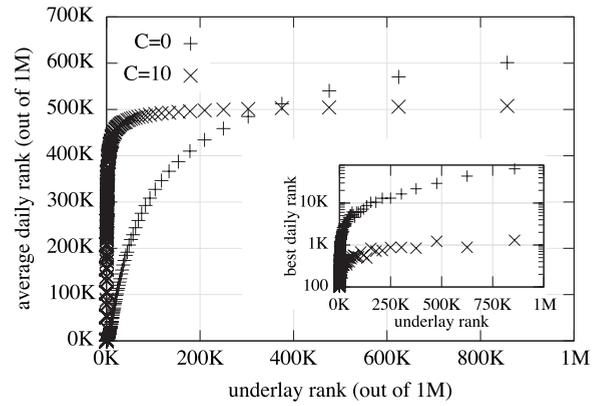


FIG. 5. Comparison of the underlay network with daily networks. We show the average rank of a node in ten daily subnetworks (with 1 million steps) versus their rank in the underlay network; nodes with higher degree have lower rank, and the top-ranked node has rank 0. In the case of ties, all tied nodes are assigned the rank that is the average value of the range they cover. The inset shows, not the average, but the best daily rank that node has over the 10 days. The  $C>0$  case is remarkably egalitarian—every node has an opportunity to become a hub in some daily network.

ing two nodes together if they are connected in any of the  $M$  subnetworks. Given their independence and random starting points, the aggregate of  $M$  daily networks must eventually converge to the underlay as  $M$  grows large, but whether this convergence occurs over a particular time scale  $M_0$ , or whether it converges in a scale-free manner, is an interesting question. One way to measure this convergence is by calculating the Top-2% overlap  $\mathcal{O}$  between the aggregate and the underlay, and subtracting this from 100%; if this quantity (called the “dissimilarity” reaches zero, then both the aggregate and the underlay both have the same set of hubs, and are roughly (though not precisely) equivalent. Figure 6 shows this dissimilarity as a function of the aggregate size  $M$ . For

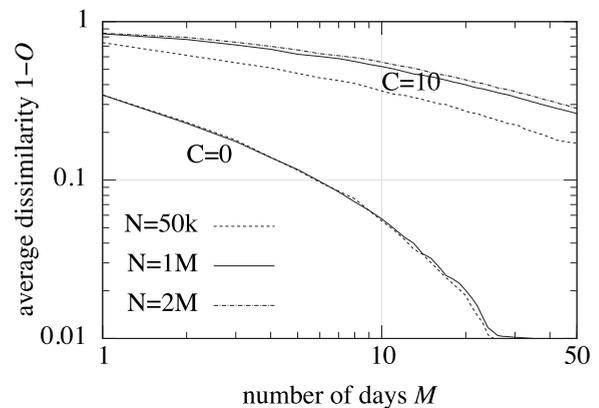


FIG. 6. Dissimilarity of networks aggregated over times ranging from 1 to 50 days. Dissimilarity is defined as the Top-2% overlap  $\mathcal{O}$  [Eq. (2)] subtracted from 100%, and is shown for both pure random walks ( $C=0$ ) and strongly weighted walks ( $C=10$ ). We show results for networks with 50 thousand, 1 million, and 2 million nodes to show how the system scales; all walks have as many steps as there are nodes ( $S=N$ ).

the pure random walk case ( $C=0$ ), the dissimilarity decreases exponentially ( $\sim e^{-M/M_0}$ , with  $M_0 \approx -7.5$ ), so that after 20 days the top nodes in both networks are basically identical. In contrast, the aggregate generated with a strong dynamic preferential attachment ( $C=10$ ) has a dissimilarity which decays much more slowly, and possibly without a characteristic time scale. This is the behavior seen in empirical DSF networks [42,43], and hints at a “multiscale” structure for the network dynamics, where networks at each time scale form scale-free topologies while the specific links in existence vary dramatically between observation time scales as well as over time. To alleviate concerns that this disparity may be a finite-size effect, we include data in Fig. 6 for networks of  $N=50\,000$  nodes and  $N=2\,000\,000$  nodes (where  $S=N$  in all cases): increasing the network size does nothing to change the  $C=0$  case, and actually results in an even slower rate of decay for the reinforced  $C=10$  networks.

#### IV. CONCLUSION

We have embarked on a research program designed to develop universal models that can recreate empirically observed phenomena in dynamic complex networks. We have shown that, using a suitable reinforced random walk on a “long-term” underlay network, one is able to produce instantaneous networks which reproduce qualitatively characteristic features of real world dynamic networks. This includes, in particular, the construction of scale-free subnetworks of a scale-free “underlay” network, whose local hubs substantially differ from subnetwork to subnetwork and from those of the underlay. We have presented evidence that dynamic preferential attachment (as opposed to a pure random walk) is a necessary requirement of dynamic scale-free networks: the preferential attachment mechanism limits the walk’s exploration of the network, which gives early visited nodes higher ranks than they would have with a more random traversal. We hope our model will stimulate further empirical and theoretical work, and provide a framework for analyzing the influence of link/structural dynamics on dynamical processes on complex networks.

#### APPENDIX: CALCULATING THE TOP- $n$ OVERLAP

If there are no ties for  $n$ th place in the network, then the top- $n$  overlap of subnetworks  $\mathcal{A}$  and  $\mathcal{B}$  is given by the formula

$$\mathcal{O} = \frac{1}{n} |A_{\text{top}} \cap B_{\text{top}}|, \quad (\text{A1})$$

where  $A_{\text{top}}$  is a set of the top  $n$  nodes in network  $\mathcal{A}$ , as  $B_{\text{top}}$  is for  $\mathcal{B}$ , and  $|\cdot|$  denotes the number of elements in a set. Given the finite size of our networks, however, there are generally large groups of nodes with the same degree, making the choice of the top  $n$  nodes somewhat arbitrary. To account for this, we modify Eq. (A1) to prorate those nodes which are tied for the  $n$ th slot, as follows: let  $a$  and  $b$  be the degree of the  $n$ th highest node in networks  $\mathcal{A}$  and  $\mathcal{B}$ , respectively. We define the set  $A_{>}$  to be the set of nodes in  $\mathcal{A}$  with degree greater than  $a$ , while  $A_{=}$  is the (guaranteed nonempty) set of nodes in  $\mathcal{A}$  with degree equal to  $a$ ; we define  $B_{>}$  and  $B_{=}$  similarly. Imagine constructing the list  $A_{\text{top}}$  of the top  $n$  nodes: it will automatically contain the entire set  $A_{>}$ , and will have room for  $n - |A_{>}|$  of the nodes in  $A_{=}$ . If we choose nodes at random from  $A_{=}$  to fill those available slots, then each node from  $A_{=}$  has probability

$$f_A = \frac{n - |A_{>}|}{|A_{=}|}. \quad (\text{A2})$$

of appearing in the list  $A_{\text{top}}$ . (The same applies for network  $\mathcal{B}$ .) Thus, when calculating the overlap between  $\mathcal{A}$  and  $\mathcal{B}$ , we will count each node in  $A_{=}$  and  $B_{=}$  as being worth a fraction of a regular node— $f_A$  and  $f_B$ , respectively—define the modified overlap  $\mathcal{O}(\mathcal{A}, \mathcal{B})$  to be

$$\begin{aligned} \mathcal{O} = \frac{1}{n} [ & |A_{>} \cap B_{>}| + f_A |A_{=} \cap B_{>}| + f_B |A_{>} \cap B_{=}| \\ & + f_A f_B |A_{=} \cap B_{=}|. \end{aligned} \quad (\text{A3})$$

Note that if there is no tie for  $n$ th place in either network, then  $f_A = f_B = 1$ ,  $A_{\text{top}} = (A_{>} \cup A_{=})$ , and  $B_{\text{top}} = (B_{>} \cup B_{=})$ ; and Eq. (A3) reduces to the simpler Eq. (A1) described above.

- 
- [1] S. Wasserman and K. Faust, *Social Network Analysis* (Cambridge University Press, Cambridge, England, 1999).
- [2] *Models and Methods in Social Network Analysis*, edited by P. J. Carrington, J. Scott, and S. Wasserman (Cambridge University Press, Cambridge, England, 2007).
- [3] R. Albert and A.-L. Barabási, *Rev. Mod. Phys.* **74**, 47 (2002).
- [4] L. A. N. Amaral, A. Scala, M. Barthélemy, and H. E. Stanley, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 11149 (2000).
- [5] S. Boccaletti, V. Latora, Y. Moreno, M. Chaves, and D.-U. Hwang, *Phys. Rep.* **424**, 175 (2006).
- [6] L. da Fontoura Costa, F. A. Rodrigues, G. Travieso, and P. R. V. Boas, *Adv. Phys.* **56**, 167 (2007).
- [7] A. Barrat, M. Barthélemy, and A. Vespignani, *Dynamical Processes on Complex Networks* (Cambridge University Press, Cambridge, England, 2008).
- [8] A.-L. Barabási and R. Albert, *Science* **286**, 509 (1999).
- [9] S. N. Dorogovtsev and J. F. F. Mendes, *Adv. Phys.* **51**, 1079 (2002).
- [10] R. Albert, H. Jeong, and A.-L. Barabási, *Nature (London)* **406**, 378 (2000).
- [11] R. Cohen, K. Erez, D. ben-Avraham, and S. Havlin, *Phys. Rev. Lett.* **85**, 4626 (2000).
- [12] R. Cohen, K. Erez, D. ben-Avraham, and S. Havlin, *Phys. Rev. Lett.* **86**, 3682 (2001).
- [13] R. Pastor-Satorras and A. Vespignani, *Phys. Rev. E* **65**, 035108 (2002).
- [14] Z. Dezsö and A.-L. Barabási, *Phys. Rev. E* **65**, 055103 (2002).
- [15] R. Pastor-Satorras and A. Vespignani, *Phys. Rev. E* **65**,

- 036104 (2002).
- [16] R. Cohen, S. Havlin, and D. ben-Avraham, *Phys. Rev. Lett.* **91**, 247901 (2003).
- [17] J. Gómez-Gardeñes, P. Echenique, and Y. Moreno, *Eur. Phys. J. B* **49**, 259 (2006).
- [18] E. M. Rogers, *Diffusion of Innovations*, 5th ed. (Free Press, New York, 2003).
- [19] V. Latora and M. Marchiori, *Chaos, Solitons Fractals* **20**, 69 (2004).
- [20] B. Skyrms and R. Pemantle, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 9340 (2000).
- [21] *Evolution of Social Networks*, edited by P. Doreian and F. N. Stokman (Gordon and Breach, New York, 1997).
- [22] T. A. B. Snijders, in *Models and Methods in Social Network Analysis*, edited by P. J. Carrington, J. Scott, and S. Wasserman (Cambridge University Press, Cambridge, England, 2007).
- [23] L. Katz and C. H. Proctor, *Psychometrika* **24**, 317 (1959).
- [24] S. Wasserman and D. Iacobucci, *Psychometrika* **53**, 261 (1988).
- [25] D. L. Banks and K. M. Carley, *J. Math. Sociol.* **21**, 173 (1996).
- [26] G. Robins and P. Pattison, *J. Math. Sociol.* **25**, 5 (2001).
- [27] P. W. Holland and S. Leinhardt, *J. Math. Sociol.* **5**, 5 (1977).
- [28] S. Wasserman, *J. Am. Stat. Assoc.* **75**, 280 (1980).
- [29] T. A. B. Snijders, *Sociol. Methodol.* **31**, 361 (2001).
- [30] Y. Y. Ahn, S. Han, H. Kwak, S. Moon, and H. Jeong, in *WWW '07: Proceedings of the 16th International Conference on World Wide Web* (Association for Computing Machinery, New York, 2007), pp. 835–844.
- [31] P. Holme, C. R. Edling, and F. Liljeros, *Soc. Networks* **26**, 155 (2004).
- [32] *Adaptive Networks: Theory, Models and Applications*, edited by T. Gross and H. Sayama (Springer, Heidelberg, 2009).
- [33] T. Gross and B. Blasius, *J. R. Soc., Interface* **5**, 259 (2008).
- [34] S. Bornholdt and T. Rohlf, *Phys. Rev. Lett.* **84**, 6114 (2000).
- [35] P. Holme and G. Ghoshal, *Phys. Rev. Lett.* **96**, 098701 (2006).
- [36] J. Ito and K. Kaneko, *Phys. Rev. Lett.* **88**, 028701 (2001).
- [37] S. Gil and D. H. Zanette, *Phys. Lett. A* **356**, 89 (2006).
- [38] T. Gross, Carlos J. Dommar D’Lima, and B. Blasius, *Phys. Rev. Lett.* **96**, 208701 (2006).
- [39] L. B. Shaw and I. B. Schwartz, *Phys. Rev. E* **77**, 066101 (2008).
- [40] M. G. Zimmermann, V. M. Eguiluz, and M. San Miguel, *Phys. Rev. E* **69**, 065102 (2004).
- [41] J. M. Pacheco, A. Traulsen, and M. A. Nowak, *Phys. Rev. Lett.* **97**, 258103 (2006).
- [42] D. Braha and Y. Bar-Yam, *Complexity* **12**, 59 (2006).
- [43] D. Braha and Y. Bar-Yam, in *Adaptive Networks: Theory, Models and Applications*, edited by T. Gross and H. Sayama (Springer, New York, 2009).
- [44] A.-L. Barabási, *Nature (London)* **435**, 207 (2005).
- [45] A. Wiggins, J. Howison, and K. Crowston, *Open Source Development, Communities and Quality: 20th World Computer Congress, Working Group 2.3 on Open Source Software*, edited by B. Russo, E. Damiani, S. Hissam, B. Lundell, and G. Succi (Springer, New York, 2008).
- [46] A. Wiggins, M. McQuaid, and L. A. Adamic, iConference 2008: iFutures: Systems, Selves, Society ([http://www.ischools.org/conference08/ic08\\_PC\\_Papers/html](http://www.ischools.org/conference08/ic08_PC_Papers/html)).
- [47] L. Adamic, C. Brunetti, J. H. Harris, and A. A. Kirilenko, “On the informational properties of trading networks,” Working Paper. School of Information & Center for the Study of Complex Systems, University of Michigan (2009).
- [48] J. D. J. Han, N. Bertin, T. Hao, D. S. Goldberg, G. F. Berriz, L. V. Zhang, D. Dupuy, A. J. M. Walhout, M. E. Cusick, F. P. Roth, and M. Vidal, *Nature (London)* **430**, 88 (2004).
- [49] N. M. Luscombe, M. M. Babu, H. Yu, M. Snyder, S. A. Teichmann, and M. Gerstein, *Nature (London)* **431**, 308 (2004).
- [50] P. G. Doyle and J. L. Snell, *Random Walks and Electric Networks* (The Mathematical Association of America, Washington, D.C., 1984).
- [51] R. Pemantle, *Probability Surveys* **4**, 1 (2007).
- [52] S. Jespersen and A. Blumen, *Phys. Rev. E* **62**, 6270 (2000).
- [53] S. Jespersen, I. M. Sokolov, and A. Blumen, *Phys. Rev. E* **62**, 4405 (2000).
- [54] J. Lahtinen, J. Kertész, and K. Kaski, *Phys. Rev. E* **64**, 057105 (2001).
- [55] S. C. Manrubia, J. Delgado, and B. Luque, *EPL* **53**, 693 (2001).
- [56] S. A. Pandit and R. E. Amritkar, *Phys. Rev. E* **63**, 041104 (2001).
- [57] F. Jasch and A. Blumen, *Phys. Rev. E* **64**, 066104 (2001).
- [58] E. Almaas, R. V. Kulkarni, and D. Stroud, *Phys. Rev. Lett.* **88**, 098101 (2002).
- [59] B. Luque and O. Miramontes, *EPL* **63**, 8 (2003).
- [60] L. A. Adamic, R. M. Lukose, A. R. Puniyani, and B. A. Huberman, *Phys. Rev. E* **64**, 046135 (2001).
- [61] M. E. J. Newman and M. Girvan, *Phys. Rev. E* **69**, 026113 (2004).
- [62] S. Bilke and C. Peterson, *Phys. Rev. E* **64**, 036106 (2001).
- [63] L. F. Costa and G. Travieso, *Phys. Rev. E* **75**, 016102 (2007).
- [64] L. da Fontoura Costa, O. Sporns, L. Antiqueira, M. G. V. Nunes, and O. N. Oliveira, Jr., *Appl. Phys. Lett.* **91**, 054107 (2007).
- [65] The underlay is a standard Barabasi-Albert network: the exponent is not the familiar  $\alpha=3$  because we are dealing with the cumulative degree distribution, which decreases the exponent by one.