Chapter 3
Time-Dependent Complex Networks: Dynamic Centrality, Dynamic Motifs, and Cycles of Social Interactions

Dan Braha and Yaneer Bar-Yam

Abstract We develop a new approach to the study of the dynamics of link utilization in complex networks using data of empirical social networks. Counter to the perspective that nodes have particular roles, we find roles change dramatically from day to day. "Local hubs" have a power law degree distribution over time, with no characteristic degree value. We further study the dynamics of local motif structure in time-dependent networks, and find recurrent patterns that might provide empirical evidence for cycles of social interaction. Our results imply a significant reinterpretation of the concept of node centrality and network local structure in complex networks, and among other conclusions suggest that interventions targeting hubs will have significantly less effect than previously thought.

3.1 Dynamic Centrality in Large-Scale Communication Networks

Recent advances have demonstrated that the study of universal properties in physical systems may be extended to complex networks in biological and social systems [1–6]. This has opened the study of such networks to experimental and theoretical characterization of properties and mechanisms of formation. In this chapter we extend the study of complex networks by considering the dynamics of the activity of network connections. Our analysis suggests that fundamentally new insights can be obtained from the dynamical behavior, including a dramatic time dependence of the role of nodes that is not apparent from static (time aggregated) analysis of node connectivity and network topology.

We study the communication between 57,158 e-mail users based on data sampled over a period of 113 days from log files maintained by the email server at a

D. Braha (✉)
University of Massachusetts, Dartmouth, MA 02747, USA; New England Complex Systems Institute, Cambridge, MA 02138, USA
e-mail: braha@necsi.edu

This chapter is an extension of D. Braha and Y. Bar-Yam, "From Centrality to Temporary Fame: Dynamic Centrality in Complex Networks," Complexity, Vol. 12 (2), November/December 2006.

T. Gross, H. Sayama (eds.), *Adaptive Networks,* Understanding Complex Systems,
DOI 10.1007/978-3-642-01284-6_3, © NECSI Cambridge/Massachusetts 2009

large university [7]. The time when an e-mail link is established between any pair of email addresses is routinely registered in a server, enabling the analysis of the temporal dynamics of the interactions within the network. To consider only emails that reflect the flow of valuable information, spam and bulk mailings were excluded using a prefilter. There were 447,543 messages exchanged by the users during 113 days observation. We report results obtained by treating the communications as an undirected network, where email addresses are regarded as nodes and two nodes are linked if there is an e-mail communication between them. Analysis based upon treating the network with asymmetric links (where a distinction is made between out-going links and incoming links) gave essentially equivalent results. From the temporal connectivity data, a time series of topological networks can be obtained; each represents an aggregation of links over a time scale that is short compared to the duration of observation (113 days). The edges forming each network in the time series thus represent the short time opportunity for communication as detected by the log files of the email server. Unless otherwise indicated, we set the time scale to one day, thus creating 113 consecutive daily networks.

Most studies of large social networks have accumulated data over the entire time of observation, whereas here using the smaller intervals of accumulation we can study how the network interactions change over time. Social network dynamics has historically been of interest, though data was limited [8, 9]. Recent papers have considered the times between communications [10] or the creation of temporally linked structures [11]. In this chapter we study for the first time the dynamics of individual importance and local structure (sub-graphs or motifs) in Dynamic Complex Networks [12].

Our first result is that networks obtained on different days are substantially different from each other. Figure 3.1 shows the correlation between corresponding edges of the 113 daily networks. Surprisingly, we find that all networks are weakly correlated, despite the expected routine nature of the social activity. Correlations between any two networks have a distribution that is approximately normal with a mean $\pm$ standard deviation of 0.15 ± 0.05 (we adopt this notation throughout). The low correlation implies that the existence of a link between two individuals at one time does not make it much more likely that the link will appear at another time. While all networks are weakly correlated, we find that workdays and weekends are more distinct, so that workday networks, and weekend networks are more correlated among themselves (correlations 0.17 ± 0.03 and 0.16 ± 0.05, respectively), than they are with each other (correlation 0.12 ± 0.02). Remarkably, the low correlations increase only very gradually if we form networks using data over multiple days, and never reach a high value even if networks are made from communications over a month or more (Fig. 3.1b).

Using the nodal "degree" (the number of nodes a particular node is connected to) we characterized the centrality of nodes in the daily networks. Each of the daily networks has a distribution of nodal degrees well described by a power-law [13], with exponents in the range . Thus a small number of highly connected nodes have great importance in the connectivity of the network. However, while each daily network has highly connected nodes, we found that they were not the same nodes. The degree of a node varied dramatically over time. For each identified "local hub," we

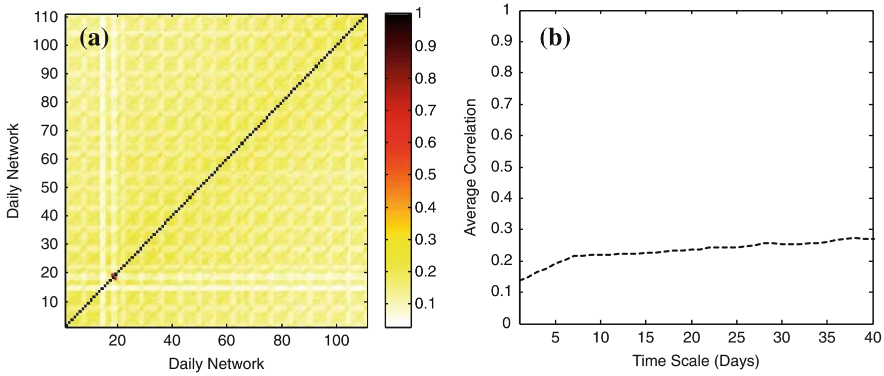

Fig. 3.1 **(a)** Matrix of correlations between pairs of daily networks sampled July 29th, 2001 (Sunday) to November 18th, 2001 (Sunday). Days 55 and 56 were excluded from further analysis due to lack of email communication. **(b)** Correlation between pairs of daily networks aggregated over times ranging from 1 to 40 days

measured its degree from day to day over the duration of observation. Surprisingly, we find that a large number of "local hubs" exhibit a highly fluctuating time-series (Fig. 3.2). The corresponding distribution of degrees over time itself follows a scale-free power-law distribution [13, 23] over two orders of magnitude (Fig. 3.2). The degree distribution of a hub over time implies that the node's degree does not have a characteristic value. The degree is small most of the time, but we only need to wait long enough to encounter degrees of any size.

A broader characterization of which nodes are important on a given day was made by comparing how the nodes were ranked in importance. We identified the top 1000 nodes, about 1.7% of the network according to their degree, for each of the daily networks. We then determined, for each pair of daily networks, the percentage of nodes that appear in both top-ranking lists ("centrality overlap," Fig. 3.3). The centrality overlap between any two networks is small, around 0.27 ± 0.06. When considering separately workday and weekend networks, the overlap values are around 0.33 ± 0.03 and 0.20 ± 0.04, respectively; consistent with the bimodal nature of the social activity. The distinctiveness of the top 1,000 nodes between daily networks is also typical for other top-ranking list sizes. By varying the percentage of nodes in the top-ranking list, it is found that the mean centrality overlap, which is already small for small percentages (0.3), actually decreases to a value of 0.2 at around 4%, before increasing slowly to 1 when the list includes all the nodes. The distributions of ranking overlaps are well behaved, having a standard deviation much smaller than the mean.

We compared daily networks with the aggregate network, as would be considered by other contemporary studies, by aggregating over the entire 113 day observation. Our previous results suggest, and direct analysis confirms, that daily networks deviate significantly from the aggregate network. We determined which nodes in the daily 1,000 top-ranking list also appear in the top-ranking list of the aggregate network, obtaining the binary image in Fig. 3.4a. Though some nodes that are ranked

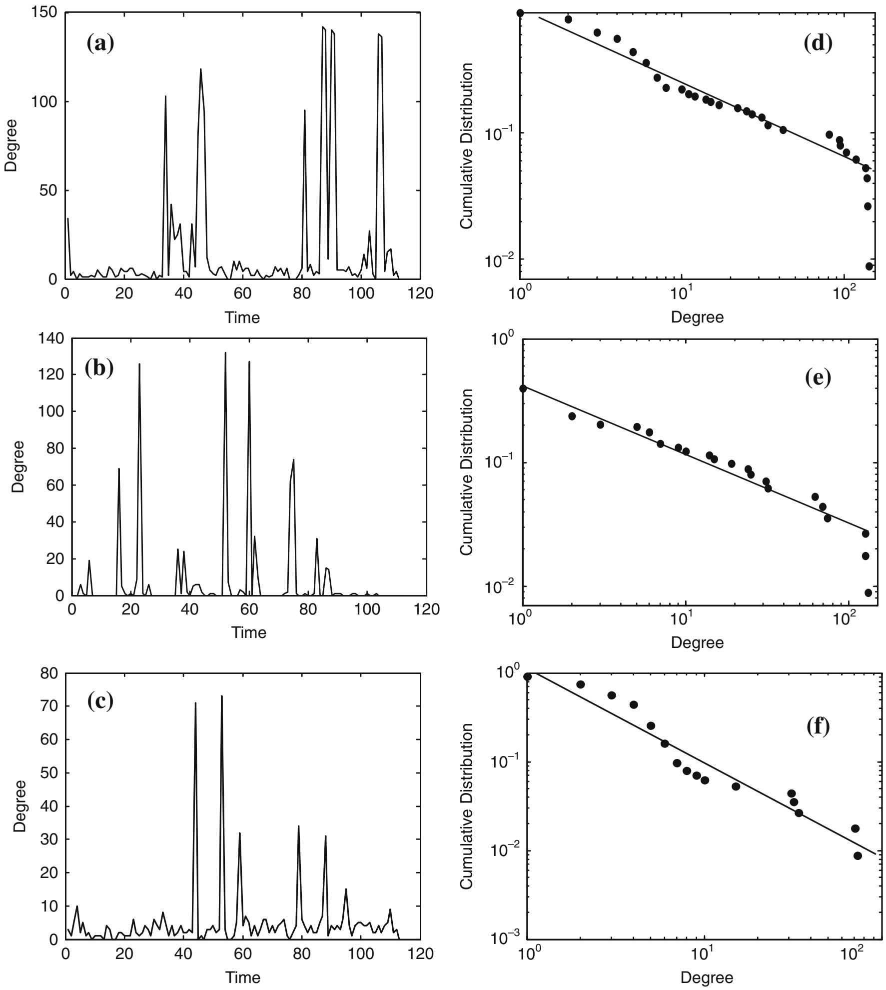

Fig. 3.2 Degree variations over time associated with the most connected node ("local hub") identified for a particular daily network. **(a–c)** Time series of degrees associated with nodes 724 (hub in day 34), 4,631 (hub in day 52), and 450 (hub in day 44), respectively. Small and very large node degrees are observed. **(d–f)** The corresponding log–log plots of the cumulative distributions of degrees over time associated with "local hubs" 724, 4,631, and 450, respectively. The distributions follow a power law ($p < 0.001$)

high in the daily networks are also ranked high in the aggregate network, a significant number are not. In particular, we find that the centrality overlap is 0.41 ± 0.03 and 0.27 ± 0.04, for weekday and weekends respectively. Comparing other sizes of the top ranked nodes gives similar results. Perhaps even more surprisingly, the nodes that are highly ranked in the aggregate network are not even *on-average* important in daily networks. To show this we calculated the average ranking position of the top 1,000 highly connected nodes in the aggregate network for each daily network.

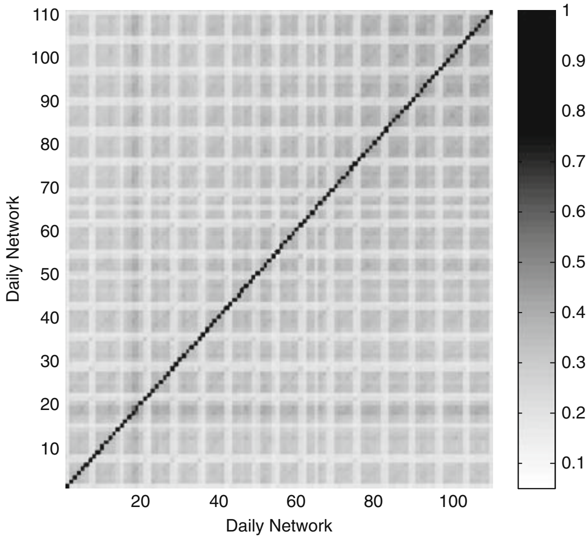

Fig. 3.3 Top-ranking list overlap between pairs of daily networks. For each pair of networks, the color code of the matrix denotes the percentage of nodes that appear in the 1,000 top-ranking list of the networks

The average ranking position over time (normalized to a fraction so that 1 is the highest and 0 is the lowest) exhibits a weekly oscillation from about 0.40 to 0.65. In the aggregate network these nodes have an average ranking of 0.99. This shows that highly connected nodes in the aggregate network only play a moderate role in the daily networks.

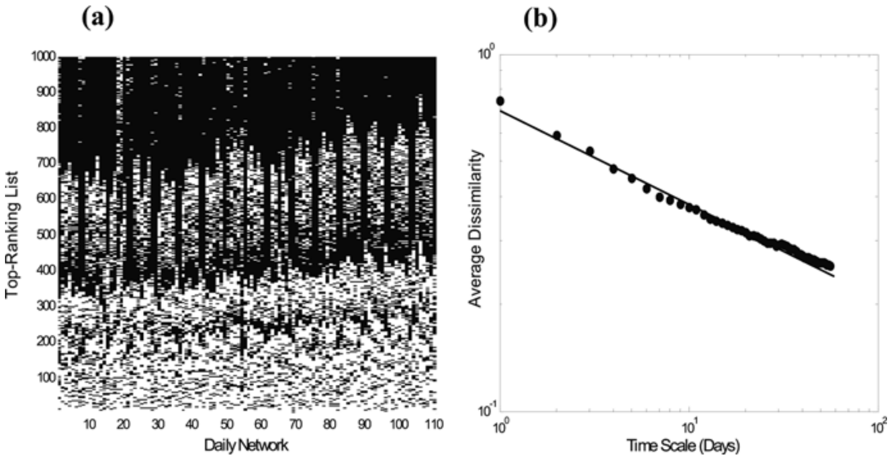

Fig. 3.4 **(a)** Comparison of the aggregate network with daily networks. A binary overlap matrix describing whether a node, included in the 1,000 top-ranking list of a daily network, also appear (*colored white*) in the 1,000 top-ranking list of the aggregate network. **(b)** Average dissimilarity of networks aggregated over times ranging from 1 to 56 days. Dissimilarity is measured as one minus the fractional overlap of the 1,000 top-ranking nodes. The plot follows a power law ($p < 0.001$) indicating that networks formed over longer time periods do not converge to a well-defined structure

Finally, we considered a full range of networks formed by aggregating links over time scales that are longer than a day and shorter than the full time period (Fig. 3.4b). Similar relationships between smaller and larger time scales to those found above are observed. Moreover, the similarity between networks at a particular time scale increases as a power-law, so there is no particular time scale at which a converged structure is achieved. Thus, the network dynamics follows a "multiscale" structure with networks at each scale forming scale-free topologies, but the specific links in existence vary dramatically between observation time scales as well as over time.

3.2 Dynamic Centrality in Spatial Proximity Social Networks

In addition to the e-mail network studied here, we have found similar results when analyzing social network data about interactions found from the spatial proximity of personal Bluetooth wireless devices, recording the interactions between pairs of students over the period of 31 days of October 2004 [12]. The spatial proximity network records dynamic interactions among 80 students who are socially related in some way (students in the same school or class), and thus reliably approximates social ties.

As before, we start our analysis by testing the association between the 31 temporal sequence of networks. We computed the Pearson's correlation coefficients as well as simple matching coefficients between corresponding edges of the 31 data networks. Consistent with the results previously reported in Sect. 3.1, we find that all networks are weakly correlated (see Fig. 3.5) despite the expected routine nature of the social activity. Despite the weak correlations, Fig. 3.5 suggests that neighboring workday networks tend to be more correlated than networks that are far apart in time. We also find that networks representing workday social interactions are significantly more correlated among themselves than they are with networks representing weekend social interactions, indicating a periodicity in the link dynamics. The relatively strong interactions among the workday networks over the fourth week suggest that unique patterns of social interactions might show up over time in response to both internal and external spikes ("external stimuli"). Overall, the above initial analysis implies that a static network analysis, which is based on aggregating the interaction data over all various time periods, will lose a lot of valuable information that is embedded in a dynamic social network.

While the identification of the "most important" nodes in networks has help to understand or predict the behavior of networked systems, it is based on the assumption that node centrality is a time-invariant property. We have demonstrated in Sect. 3.1 that, for a dynamic email network, centrality measures are time-dependent and might fluctuate over time. That is, a highly central node at one time point might be the least central in a different time point. In general, a node centrality index in a dynamic network may be better defined as a probability distribution over the total sampling time of the network. This is in contrast to static network, where node centrality index is a single measure. To illustrate the above argument in the context of the spatial proximity network, we computed the ranking of each node (student) according to its degree for each of the 31 data networks. Figure 3.6 shows that the

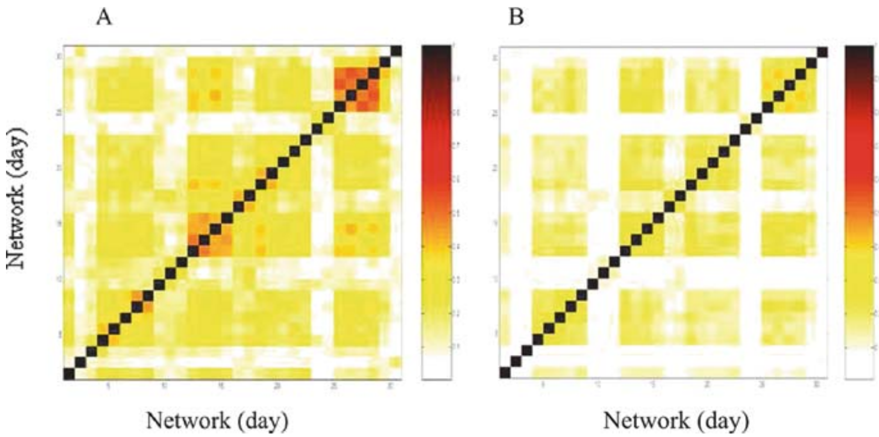

Fig. 3.5 Correlation profiles of the time-series networks sampled over 31 days. **(a)** The Pearson's correlation coefficient between corresponding links of the 31 temporal sequences of networks. **(b)** Correlation based on Jaccard's coefficient of similarity. The Jaccard's coefficient measures the degree of overlap between two networks by computing the ratio of the number of shared links to the number possessed by both networks. The color code of each matrix denotes the degree of correlation shown in the matrix

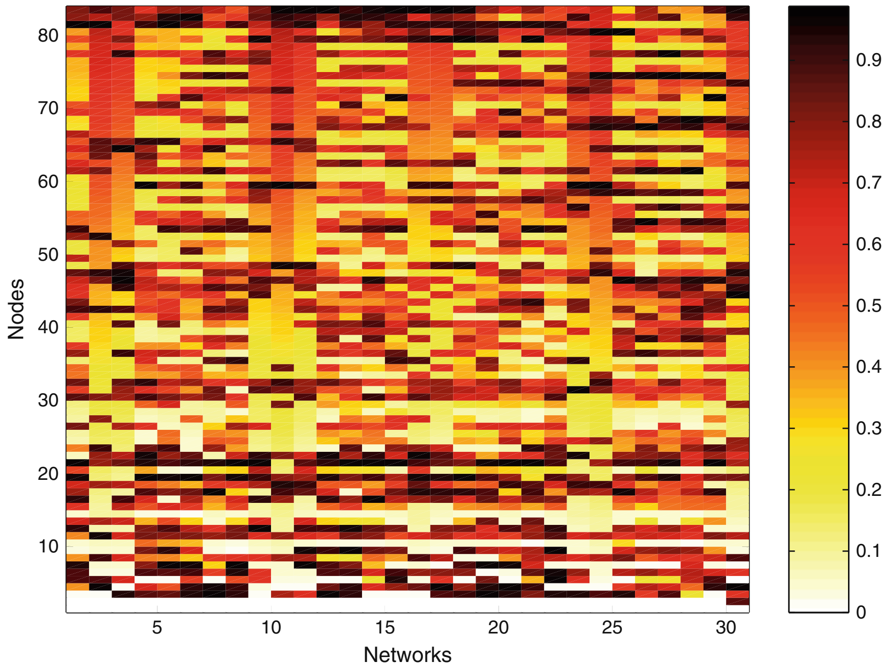

Fig. 3.6 Normalized degree ranking of each actor in each sampled network. The degree ranking is obtained by sorting actors in ascending order by their degree centrality measure. The normalized ranking centrality is the ranking position divided by the number of actors. The color code of the matrix denotes the degree of normalized degree ranking shown in the matrix

prominence of actors is “spread out” quite consistently over time suggesting that the prominence of actors embedded in a network is time-based.

3.3 Dynamic Network Motifs and Cycles of Social Interaction

We have found above that the time-series social networks are weakly correlated (Figs. 3.1 and 3.5). Notwithstanding the weak correlation, it is of interest to analyze and compare the local structure of the various temporal sequence of networks. It has been shown that many complex networks include some sub-graphs (motifs) that are significantly abundant as compared to randomized versions of the same networks, while others are strongly suppressed [14, 15]. The presence or absence of a given sub-graph presumably encapsulate information about the system-level function the network performs [14, 15]. Moreover, *a sub-graph significance profile* – a set of counts of the different kinds of sub-graphs that arise in a real-world network compared to randomized networks – serves as a distinctive signature of the network [14, 15]. For a dynamic network, the local structure is time-dependent and might evolve over time. Analyzing the time-based local structure might provide important information about the dynamics of system-level task and functionality.

We present in Fig. 3.7 the significance profiles of the 6 types of connected tetrads for the different time-series networks. Despite the overall similarity of significance

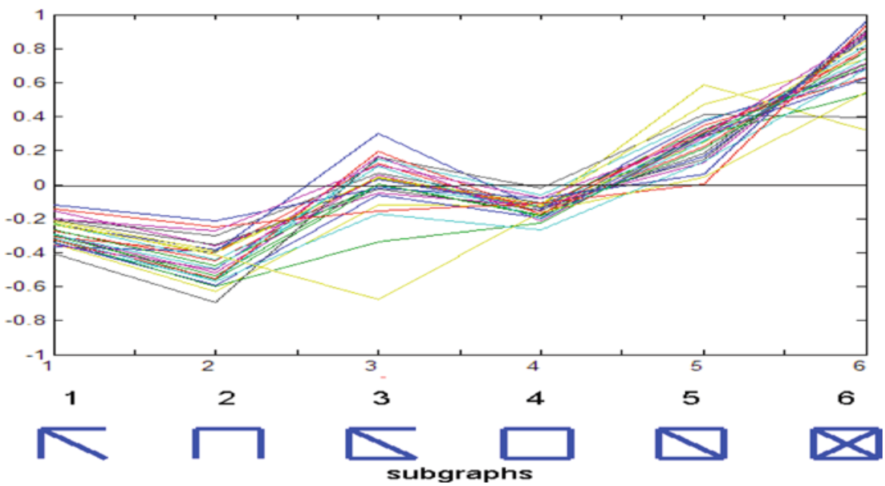

Fig. 3.7 The tetrad significance profile of networks for the different time-series networks. Continuous lines are drawn as guide to the eye. The tetrad significance profile shows the normalized significance level (Z score) for each of the 6 connected tetrads. Tetrad significance (Z score) is the difference between numbers of occurrence in the given network and in an ensemble of randomly rewired surrogate networks with the same degree sequence, divided by the standard deviation. The tetrad significance profile is the normalized vector of tetrad significances (see [14, 15] for details). The wide fluctuations of the network local structure reproduce the dynamic of social interaction at the system level

profiles, the figure indicates a fluctuation of the network local structure over time. To examine for possible cyclic behavior embedded in the time-series networks, we calculated the similarities between them by looking at the correlation between the significance profiles of the 6 types of connected tetrads for the different networks. Next, we applied an average-linkage hierarchical clustering algorithm [16] to the significance profile correlations (Fig. 3.8). Several families of networks – each includes networks separated by time intervals varying 2–20 days – with very similar significance profiles ($c > 0.995$) emerge from this analysis. The relation between the local structures of networks within tightly interconnected families is further visualized in Fig. 3.9, which presents the significance profiles of the 6 types of connected tetrads for several families of networks. The recurrent patterns of network local structure over time might provide empirical evidence for cycles of social interaction despite being only the aggregate of distinctive behaviors and preferences of individuals.

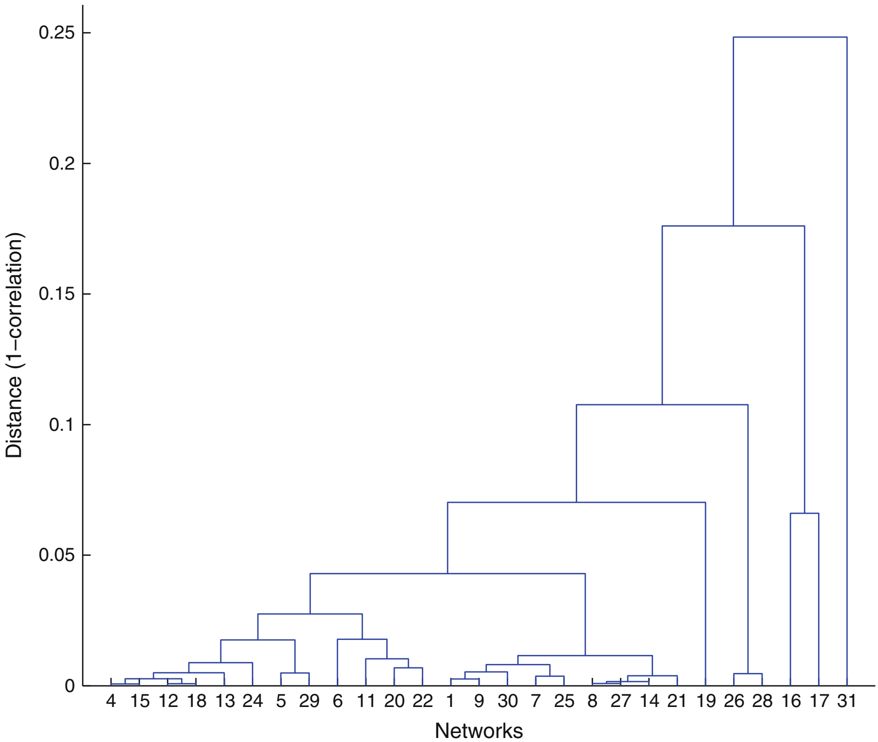

Fig. 3.8 Identifying families of networks over time with similar local structure. The hierarchical cluster tree that quantifies the relation between the different networks, based on the correlation coefficients between the tetrad significance profiles of the temporal sequence of networks

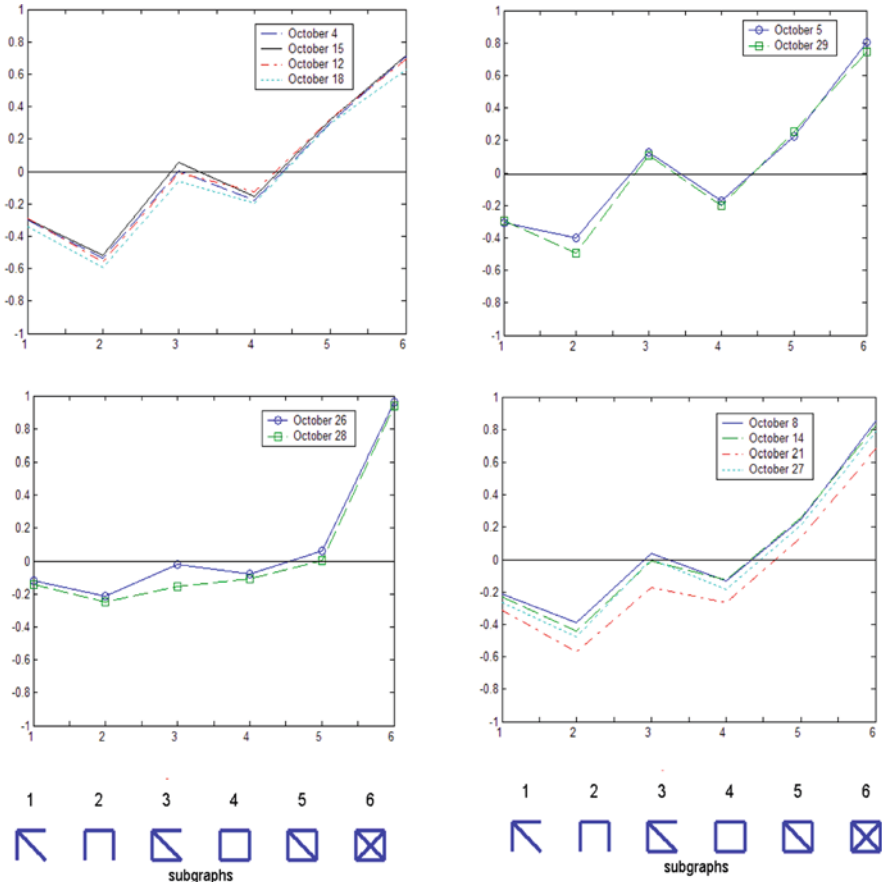

Fig. 3.9 Evidence for cycles of social interaction in dynamic social networks. Families of networks with highly similar local structure are identified, as suggested by the average-linkage hierarchical clustering tree (Fig. 3.8)

3.4 Summary

In summary, we have demonstrated that the static topology does not capture the dynamics of social networks. The prominence of nodes (as measured by degree) within the networks fluctuates widely from day to day, and a high degree in the aggregate network does not predict a high degree for individual days. Our conclusions are in sharp contrast to previous complex network research, which emphasizes the importance of aggregate nodal centrality in a static network topology [1–5, 7, 11, 14, 15, 17–20, 24].

Implications of a dynamic node centrality contrast with existing analyses that consider targeting nodes with the highest degrees to disrupt network communication or transport [25]. Dynamic centrality implies that targeting nodes with the highest degrees at one time only weakly affects the nodes that are highly connected at another time. The approach of targeting high-degree nodes has been suggested, for example, to be an effective disease and computer virus prevention strategy; i.e. identification and "vaccination" of those nodes, would inhibit the spread of infection or computer viruses [21, 22]. Similarly, popular influencer marketing techniques (closely related to word-of-mouth or viral marketing) are based on the premise that a large number of people are connected to everyone else through a small number of hubs. Thus, identifying and focusing marketing activities around these hubs could increase the likelihood of initiating a cascading adoption of products or services – a type of social epidemic. The perspective of such marketing techniques presupposes that an individual who is an influencer now is likely to be a social hub later. In other words, the topology of a social network is quite static. Our work implies that, at the very least, a more agile strategy of monitoring, vaccinating nodes, or focusing marketing activities based upon centrality over time is necessary. Otherwise a treatment based upon aggregate connectivity information will miss the impact of a node that otherwise has a low connectivity, becoming highly connected. More generally, our findings call for a radical rethink of the mechanisms underlying the processes of link dynamics and diffusion on Dynamic Complex Networks – both experimentally and theoretically.

The type of dynamic analysis of networks we performed is pertinent to a wide range of network types. Whether or not there exists an underlying fixed topological structure, the question of which links are actually used is a relevant one. Thus, actual travel on a transportation network, and actual interactions that occur between molecules that can bind to each other, are both examples of networks that have an underlying structure but whose dynamic structure is relevant to the behavior and functionality of the system over time. This study demonstrated the potential role of time in complex networks. Ultimately our goal is to understand the role of both time and space in complex networks, leading to a Spatio-Temporal Complex Network Theory.

References

1. D. Braha and Y. Bar-Yam, Phys. Rev. E **69**, 016113 (2004).
2. D. Braha and Y. Bar-Yam, Manage. Sci. **53** (No. 7), 1127 (July 2007).
3. R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).
4. L. A. N. Amaral, A. Scala, M. Barthélémy and H. E. Stanley, Proc. Nat. Ac. Sci USA **97**, 11149 (2000).
5. Y. Bar-Yam and I. R. Epstein, Proc. Nat. Ac. Sci USA **101**, 4341 (2004).
6. Y. Bar-Yam, *Dynamics of Complex Systems* (Perseus Books, Reading, 1997).
7. H. Ebel, L.-I. Mielsch and S. Bornholdt, Phys. Rev. E **66**, 035103 (2002).
8. S. Wasserman and K. Faust, *Social Network Analysis* (Cambridge University Press, Cambridge, 1999).
9. P. Doreian and F. N. Stokman (Eds.) *Evolution of Social Networks* (Gordon and Breach, New York, 1997).
10. A.-L. Barabási, Nature **435**, 207 (2005).
11. J. P. Eckmann, E. Moses and D. Sergi, Proc. Natl. Acad. Sci. U.S.A. **101**, 14333 (2004).

12. D. Braha and Y. Bar-Yam, NECSI Technical Report 2005-February-01 (2005). The spatial proximity raw data was collected by the MIT Media Lab.
13. A.-L. Barabási and R. Albert, Science **286**, 509 (1999).
14. R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii and U. Alon, Science **298**, 824 (2002).
15. R. Milo, S. Itzkovitz, N. Kashtan, R. Levitt, S. Shen-Orr, I. Ayzenshtat, M. Sheffer and U. Alon, Science **303**, 1538 (2004).
16. S. C. Johnson, Psychometrika **32**, 241 (1967).
17. H. Jeong, B. Tombor, R. Albert, Z. N. Oltavi and A.-L. Barabási, Nature **407**, 651 (2000).
18. H. Jeong, S. Mason, A.-L. Barabási and Z. N. Oltvai, Nature **411**, 41 (2001).
19. R. Ferrer, C. Janssen and R. V. Solé, Phys. Rev. E **63**, 32767 (2001).
20. R. Guimerà, L. Danon, A. Díaz-Guilera, F. Giralt and A. Arenas, Phys. Rev. E **68**, 65103 (2003).
21. A. L. Lloyd and R. M. May, Science **292**, 1316 (2001).
22. R. Pastor-Satorras and A. Vespignani, Phys. Rev. Lett. **86**, 3200 (2001).
23. D. J. de S. Price, Science **149**, 510 (1965).
24. W. Aiello, F. Chung and L. Lu, Proc. ACM STOC, 171 (2000).
25. R. Albert, H. Jeong and A.-L. Barabási, Nature **406**, 378 (2000).