Feedback between node and network dynamics can produce real world network properties.

Hilla Brot[1], Lev Muchnik[2], Jacob Goldenberg[3], Yoram Louzoun[1,4,*]

[1] Gonda Brain Research Center, Bar-Ilan University, Ramat Gan, Israel, 52900

[3]School of Business Administration, NYU, NY

[4] Department of Mathematics, Bar-Ilan University, Ramat Gan, Israel, 52900

- To whom correspondence should be addressed : louzouy@math.biu.ac.il

Abstract

Real world networks are characterized by common features, including among others a scale free degree distribution, a high clustering coefficient and a short typical distance between nodes. These properties are usually explained by the dynamics of edge and node addition and deletion.

In a different context, the dynamics of nodes' content within a network has been often explained via the interaction between nodes in static networks, ignoring the dynamic aspect of the edge addition and deletion.

We here propose to combine the dynamics of the nodes content and of the edges addition and deletion, using a threshold automata framework. Within this framework, we show that the typical properties of real world networks can be reproduced with a Hebbian approach, in which nodes with similar internal dynamics have a high probability of being connected. The proper network properties emerge only if an imbalance exists between excitatory and inhibitory connections, as is indeed observed in real networks.

We further check the plausibility of the suggested mechanism by observing an evolving social network and measuring the probability of edge addition as a function of similarity between contents of the corresponding nodes. We indeed find that similarity between nodes increases the emergence probability of a new link between them.

The current work bridges between multiple important domains in network analysis, including: network formation processes, Kaufmann Boolean networks and Hebbian learning.

It suggests that the properties of nodes and the network convolve and can be seen as complementary parts of the same process.

PACS: 89.65-s, 89.2.Hh, 64.6al, 87.18.Sn, 05.40-a

Introduction

Networks are typically studied in two separate contexts: On the one hand, the generation of edges and nodes is typically studied assuming no intra-node dynamics (the nodes content is constant and unchangeable). One can find a large number of such models, including among many others models explaining the degree distribution of nodes [1-3], the typical distance distribution [4], the high clustering coefficient [5], as well as more complex scaling features of such networks [6]).

On the other hand, the dynamics of the node's content is usually studied assuming a fixed network, for example, in genetics regulatory networks [7, 8], neural networks [9, 10] and contact process networks [11-13]. One of the most fundamental approaches developed to study the state distribution in genetic regulation networks is the Kaufman networks (KN) formalism [14, 15]. The states of nodes in Kaufman networks are binary and the nodes are wired randomly. Each node is regulated by a randomly chosen logical function [15]. In Kaufmann networks, the node content is dynamic, while the edges properties are constant over time.

A third class of models studied in the context of neural networks are Hopfield Networks (HN) [16]. HNs were developed in order to describe content addressable memory [17]. HNs have an all-to-all connectivity. Each node has a binary value and the edges weights are changed according to the required memory patterns. In HNs, the network connectivity is constant, while the edges weights change.

While it is sometimes convenient to separate between the evolution of the network edges and the intra-node dynamics, the two processes occur concurrently and are likely to affect each other. Here, we propose to merge the concept of HNs and of KNs to represent this interrelation. We modify KNs so that the edges are added based on the content similarity between nodes. Specifically, instead of changing the edges weights, we change the probability that an edge would be created. We show that this simple change produces the features shared by the majority of the real world networks, in multiple domains, including social relations [18-20], biological systems [21-24], eco-systems [25-27], internet servers, WWW pages connections [3, 28, 29] and neural networks [30-32].

This is not the first analysis relating the creation of social networks and content [33, 34]. The current model, however, is the first time a simple feedback model between content and structure is used to explain structure formation. In order to show that such a model is plausible, we show that it reproduces a large set of observed networks features [1, 35, 36], such as:

- a power law total degree distribution;
- a high clustering coefficient;
- a small average distance between nodes;
- the difference between the in-degree and the out-degree distribution - a scale-free distribution of the in degree and a Gaussian for out degree distribution;
- a positive degree assortativity;
- a scaling law with a -1 power for relation between node's clustering coefficient and the node degree.

Many models have been developed to explain some of these common features (e.g. [35, 37]). However, all such models separate the edge dynamic from the nodes states dynamics. Our current goal is not to compare the quality of our model to existing models, but rather to show that a feedback between content and structure can indeed explain many observed network features.

A second more experimental proof that the proposed mechanism is plausible is based on the analysis of multiple snapshots of the Live-Journal blogs, where we show that nodes (blogs) tend to connect (link) to similar nodes.

Methods

Network simulation

The network dynamics was simulated on an N node network. Each edge in the network was assigned either a positive interaction (i.e. a weight of +1) or a negative interaction (i.e. weight of -1). The nodes were binary and could receive internal values of either 1 or -1. Each node had an initial value of 1 or -1 with equal probability. The value of each node at the next time point was a function of the weighted sum of its incoming neighbors.

The network was initiated as an Erdos-Renyi network [38] with a connectivity of *V*. A constant fraction (V_p) of the newly added edges was set to be positive (1) and the others (*V_n=1-V_p)* were set to be negative (-1) (parametric values will be introduced in the Results section). At each step the value of each node was computed using a simple threshold rule:

$$(1)\quad x_i(t+1) = Sign\left(\sum_j w_{ij} x_j(t)\right),$$

where x_i is the state of node *i,* and w_{ij} is the weight of the edge between nodes *i* and *j*. Such networks are often termed threshold automata [39]. We then introduced an external noise to the nodes states and randomized a fraction (*No)* of the nodes. The update of the network was synchronous as is typically applied in Boolean networks [14, 15].

The only change we performed compared with standard threshold automata models is the introduction of edge dynamics. The edge addition mechanism was based on the correlation between nodes. A correlation between the values of the nodes over *H* previous iterations*)* was computed for randomly selected pairs of nodes. If the correlation exceeded a threshold value (ψ), an edge was added between them. These new edges received a weight of +1 or -1 with probabilities of *V_p* and *V_n*, respectively. Edge removal was random and was at the same rate as the edge addition (i.e. one edge was removed on average for each edge added). We did not include the removal of edges between anti-correlated nodes for the sake of the model simplicity. The precise details of the deletion rate had no effect on the results.

Live-Journal crawling

Live-Journal [40] is one of the earliest and still popular online blog systems hosting 8-17 million personal blogs and well over 500 thousands communities. Users periodically publish posts detailing their life experiences, political views or opinions on events. These are later discussed in threads of comments. To simplify navigation of the system and to receive timely activity updates, users maintain a list of bookmarks of the blogs they wish to follow. Since the blogs are personal, these bookmarks represent the social network emerging in the system. We have chosen the Live-Journal network because it is very dynamic and content dependent.

In addition to periodic posts and bookmarks, Live-Journal users usually maintain a detailed profile that includes their place of residence, a list of educational institutions they attended and frequently a comprehensive list of interests. These profiles represent the blog owner, outline his/her scope of interests and assist occasional visitors to determine whether they are interested in the blog. We created a crawler and collected nine periodic snapshots of the Live-Journal system of blogs. Each snapshot took about a month and a half to collect and was immediately followed by the next scan. The collected data included the complete profile of each user, most noticeably, the list of interests and the list of ties representing the blogs closely followed by the user. The data was anonymized – user names were substituted with unique numerical identifiers.

Each snapshot contains above 8 million blogs (nodes) and more than 120 million different connections between them (edges). Each consecutive snapshot contains more blogs and more edges than the previous one (roughly 100,000 more blogs and a few million more edges for each following snapshots). There are more than 8 million overlapping blogs among all the 6 snapshots. For each snapshot there are also more than 7 million different domains of interest defined by the users. We used the LJ network to check whether the model proposed here is plausible in a real world network.

Measurement of network properties

We measured multiple values on the real and simulated networks. Most values are quite standard by now, such as the indegree and outdegree, the distance distribution and the clustering coefficient. The precise definition of all these measures can be found in multiple reviews (e.g. . [5, 41, 42]).

Similarity between nodes in LiveJournal

We measured similarity between users according to their realm of interests. Some of the interests are related one to each other. For example: Microsoft is related to computer and guitar is related to music.

We thus define two levels of similarity. Similarity between concepts $d(x,y)$ is defined as the fraction of time they appear in the same user (i.e. how many time does the same user have interests x and y). Note that this is not symmetric. In order to define the similarity between nodes, with distinct domain of interest lists (U_A, U_B), we first compute the size of the overlap ($|U_A \cap U_B|$). We add to that the average similarity for all domains of interest that are *not* in the overlap, and divide by the maximum between the sizes of U_A and U_B.

$$\frac{|U_A \cap U_B| + \sum_{x \in U_A / U_B, y \in U_B} \frac{d(x,y)}{\sum_{y'} d(x,y')} + \sum_{x \in U_A, y \in U_B / U_A} \frac{d(x,y)}{\sum_{x'} d(x',y)}}{\max(|U_A|, |U_B|)}$$

This produces a symmetric similarity value between nodes. A similarity of 1 occurs when A and B have the same domains of interest.

Results

In order to study the possible feedback between the dynamics of nodes and the structure of the network, we developed a simple basic model, where the state of each node affects its connectivity, and the connectivity feedbacks on the states of the nodes. We used a standard threshold automata [39], where the state of each node is binary (either -1 or 1), and is determined by the sign of a linear combination of the states of the nodes pointing to it, with weights of either -1 or 1. In this simple form, the model is a degenerate Kaufmann network [14] that uses a very limited class of Boolean functions. This deterministic dynamics was supplemented with a stochastic noise that replaced the value of a certain fraction of the nodes (*No=15%*) to a random value (either -1 or 1 with a 50 % chance).

The feedback of the node dynamics on the network was performed through edge addition. The network was initiated as an Erdos Renyi network [38]. At each time step, some edges were added to the network, while other edges were randomly removed. The edges added to the network were not randomly chosen. Instead, at each iteration, we calculated correlation between a set of randomly chosen nodes. Any two tested nodes exhibiting a correlation of more than ψ=0.3 along the previous *H=100* time steps were connected. The added edge direction was chosen randomly. Among the newly added edges, *V_p =70%* were set to have a positive (+1) weight. Note that the value of *V_p* has a critical effect on the results. Edge removal was applied in order to maintain an approximate steady state on the number of edges and to maintain the network connectivity. For each edge added, a node was selected at random and a random edge pointing to this node was removed unless that node happened to have a degree of 1. In such a case another edge was removed, in order to maintain the connectivity of the network.

As a test for the plausibility of the proposed model, we checked if using such schematic dynamics, one can reproduce the now classical real-world network properties:

- A power law degree distribution.
- A high clustering coefficient.
- A short distance distribution [1].
- A positive degree mixing factor.
- The correct in and out degree distribution.
- The inverse ratio between node's degree and clustering coefficient [36, 43, 44].

Indeed, the simulation converged to a classical small world network with a scale free total degree distribution (Figure 1), a high clustering coefficient ($<CC>=0.115$), a characteristic distance distribution (approximately Poisson) (Figure 2), a scale free in-degree distribution and a Guasssian out-degree distribution (Figure 3), a positive correlation between the degree of neighboring nodes ($R= 0.09, p<1.e-10$), and a inverse correlation between the degree and the average clustering coefficient (Figure 4) . While this is not a proof of the correctness of the proposed feedback of the node dynamics on the network dynamics, it shows that such a feedback based model is plausible.

The model contains multiple parameters, but most of them have no effect on the results and could be chosen at any other values. The initial number of edges had no influence on the properties of the generated networks (Figure 1a). We let the edge number increase until the desired number was obtained. Similarly, the total number of nodes (N) had no effect on the results. We checked 4 different network sizes (number of nodes) with a fixed average number of edges per node, and obtained quite similar results (Figure 1c). The average degree (V) (Figure 1h), the time over which correlations were computed (H) (Figure 1g) and the number of steps in the simulation (Figure 1d) also had no effect on the results beyond a minimal value. Table 1 contains all the parameters values in the simulation and the parametric range used to check the network stability (Table 1 and Figure 1).

The emergence of the real-world network properties is sensitive to two main parameters: The imbalance between the fraction of positive and negative edges and the noise level. These two elements affect the positive feedback loop of the degree on itself:

- The fraction of excitatory edges (*V_p* edges with weight of 1 vs -1): This parameter defines the fraction of the edges with positive weights (Figure 1f). A random assignment (*V_p=0.5)* or an assignment of all edges to be positive (*V_p=1)* does not produce a power law degree distribution. A random assignment, as well as an assignment of only positive edges, ruins the scale free distribution *(*Figure 5). Values in between lead to a power law degree distribution. Interestingly, a fraction of approximately 70 % excitatory edges has been observed in multiple real world networks [45-49]. The sensitivity to this parameter will be further discussed.

- The other essential element is the relation between the noise level and the correlation threshold for edge addition. Noise decreases the correlation between edges. If the noise level is too high, then the edge addition process will freeze. If the noise level is too low, then the threshold automata will freeze [50, 51]. An intermediate noise level is required.

To summarize, only two parameters affect the model results: The relation between the noise level and the correlation cutoff and the imbalance between positive and negative edges.

A classical mechanism inducing a scale free degree distribution is a “rich get richer” [52] mechanism where a positive feedback of the degree on itself exists. In the current model, such a mechanism would imply that high degree nodes are more probable than low degree nodes to have an above-threshold correlation, and thus to induce the creation of a new edge. We computed the average degree of unconnected node pairs with a supra-threshold correlation and the average degree of node pairs with correlations below cutoff (Figure 6). If the chances for positive and negative edges are equal, there is no advantage to high degree nodes. However, when 70 % of the edges have a positive weight, there is a maximal advantage to high degree nodes. In such a case (*V_p=0.7*), there is indeed a “rich get richer” mechanism.

The high clustering coefficient can be explained by the transitive advantage of nodes with a common first neighbor (A->B and A->C). Node pairs that are both connected to another node, but not connected one to each other have a slightly higher fraction of supra-threshold pairs compared with node pairs that have no common neighbor (Figure 7). Interestingly, following the synchronous update mechanism, nodes connected in the typical transitive path (A->B,B->C) are less correlated than nodes connected by a common neighbor (A->B,A->C). Such a behavior was actually observed in real networks, where a common network was shown to increase the probability of an edge addition, while transitivity does not [53]. To summarize, the scale free distribution emerges in the limited parameter range where correlations are not random on the one hand and the nodes are not fixed (up to the noise level) on the other hand. In this regime many other realistic features appear, such as the preference for a feed forward type attachment mechanism and a high clustering coefficient.

Delayed correlation

The relation between similarity between nodes and the creation of edges was initially proposed in the context of Hebbian learning [40]. However, the typical time between the activation of a pre-synaptic and a post-synaptic neuron can be significantly longer than the length of a typical spike, introducing a delay. We have thus checked if the proposed mechanism can induce realistic network properties, even when the nodes correlation was calculated with a delay (i.e. the correlation between the current activity of a node and the activity of a different node a few steps previously). Note that the introduction of a delay between the origin and target of an edge automatically introduces directionality to the network, beyond the one induced by the node dynamics. The early node precedes the late one, and thus the emerging edge will be from the early node to the late one.

We tested the effect of the offset between the time series of states used to compute the correlation and to decide whether to introduce a new edge between the nodes. This kind of edge creation between nodes according to correlation between delayed state resembles a Hebbian learning rule, where the strength of chemical synapses is modified by the synchronization of the signal transmission between two neurons (synaptic plasticity) [54, 55], which leads to a process called LTP (Long Term Potentiation) considered to be a major process underlying memory and learning. Figure 8 shows the degree distribution of a network obtained by the simulation with the standard set of parameters (as in Table 1). While these networks exhibit a scale free degree distribution (Figure 8), and a short distance distribution (Figure 2), they have a lower clustering coefficient of 0.012 (which is still on the borderline of the range of realistic networks). As was the case for the instantaneous network, the emergence of the network properties is not very sensitive to parameters (data not shown).

To summarize, while the model is robust to changes in a wide range of parameters, a crucial condition for it to reproduce real world like networks is a positive feedback between the correlation and the degree. This positive correlation is translated to an autocatalytic process of the degree on itself, leading to the scale free distribution. For this feedback to occur, the fraction of edges with a positive weight must be larger than the number of edges with a negative state. A node in a +1 state will increase the probability of other nodes being in a +1 state if the edge between them has a positive weight and decrease it, if the edge has a negative weight.

In order to check whether the assumption that similarity and the probability of adding an edge are correlated, we checked the edges addition in the Live-Journal blogs network. This is obviously only one possible example, but the goal of this analysis is to show the plausibility of the assumption in real networks, and not its general applicability. We have collected detailed profiles of the Live-Journal users, and produced a sequence of snapshots representing the changes in the underlying social network. We then computed the relation between the similarity of node pairs and the chance that a new link would emerge between them (see Methods). We computed the similarity of unconnected node pairs preceding the production of a new edge between them, and compared it to the similarity of node pairs not preceding such an addition. The node-pairs that will acquire in the following snapshot an edge indeed had a higher similarity than the one that will not acquire an edge (T test $p<1.e-50$) (Figure 9).

Discussion

The relation between similarity and edge addition has been addressed in the literature from at least two perspectives. One proposed mechanism suggested that similar nodes tend to connect [56, 57]. The alternative mechanism claimed that connected nodes tend to become more similar over time [58]. While evidence for the existence of each of the mechanisms has been presented, these two opposing positions stem from a short term view of the network dynamics. We here propose a feedback loop between these two processes.

Similar nodes tend to accumulate common edges and these edges contribute to the similarity between the nodes.

The proposed feedback loop explains many observed features of real world networks. The power law degree distribution is explained by this autocatalytic mechanism, since the feedback loop induces an autocatalytic edge addition mechanism, which is known to lead to power law distributions [59]. Another important feature of real world networks explained by this mechanism is the high clustering coefficient. This high clustering coefficient results from the effect of common origins. If A is similar to B and to C, then B and C have an above random probability to be similar. Since similar nodes have a higher probability of being connected, this automatically leads to a high clustering coefficient. Interestingly, transitivity (A->B,B->C) was not shown to induce a high correlation. The short network diameter is explained by the attachment of random nodes. Other properties explained by such a mechanism are the positive degree assortativity, the difference between the indegree and the outdegree distributions, and the inverse relation between the degree and the clustering coefficient.

Many models could have been adopted for the node dynamics. We have here used a simplified threshold automata model, which is a direct simplification of Kaufmann Networks [14, 15], with a limited set of Boolean function. More realistic models could have been used with similar results, as long as the two elements of the autocatalytic positive feedback loop and the need for a high correlation to add edges are maintained. In the context of our model, these conditions are expressed through the requirement of an imbalance between the number of edges with positive and negative weights. Another important element is the rate differences between the internal node dynamics and the edge addition/deletion dynamics. The former must be much faster than the latter. We have confirmed in the simulation that when the edge addition and removal dynamics is too fast, real world network properties do not emerge.

Similarity is reflexive and thus does not define a direction for the edge addition. This can be solved by assuming a delay. Such a delay can occur for example in neurons, where the pre-synaptic neuron must be activated before the post-synaptic neuron. If such a delay is introduced, the real world network properties are conserved, and causality is incorporated in the model.

The model presented here links Boolean networks, Hebbian learning and network dynamics. The current analysis only shows the plausibility of the basic mechanism driving the model. We have shown in parallel that in blogs, the mechanism of edge addition as a function of the nodes similarity is indeed observed. Thus, we show that it is possible that the dynamics of edge addition and removal is determined by the content of the nodes and not directly by the structure of the network.

Bibliography

1. Albert, R. and A. Barabsi, *Statistical mechanics of complex networks.* Reviews of modern physics, 2002. **74**(1): p. 47-97.

2. Callaway, D.S., et al., *Are randomly grown graphs really random?* Physical Review E, 2001. **64**(4): p. 41902.

3. Kleinberg, J., et al. *The web as a graph: Measurements, models, and methods*. 1999: Springer-Verlag.

4. Newman, M.E.J. and D.J. Watts, *Renormalization group analysis of the small-world network model.* Physics Letters A, 1999. **263**(4-6): p. 341-346.

5. Newman, M.E.J., *The structure and function of networks.* Computer Physics Communications, 2002. **147**(1-2): p. 40-45.

6. Song, C., S. Havlin, and H.A. Makse, *Origins of fractality in the growth of complex networks.* Nature Physics, 2006. **2**(4): p. 275-281.

7. Basso, K., et al., *Reverse engineering of regulatory networks in human B cells.* Nature Genetics, 2005. **37**(4): p. 382-390.

8. Chaves, M., R. Albert, and E.D. Sontag, *Robustness and fragility of Boolean models for genetic regulatory networks.* Journal of Theoretical Biology, 2005. **235**(3): p. 431-449.

9. Cowley, M.A., et al., *Leptin activates anorexigenic POMC neurons through a neural network in the arcuate nucleus.* Nature, 2001. **411**(6836): p. 480-484.

10. Diesmann, M., M.O. Gewaltig, and A. Aertsen, *Stable propagation of synchronous spiking in cortical neural networks.* Nature, 1999. **402**(6761): p. 529-533.

11. Castellano, C. and R. Pastor-Satorras, *Non-mean-field behavior of the contact process on scale-free networks.* Physical Review Letters, 2006. **96**(3): p. 38701.

12. Yang, R., et al., *Epidemic spreading on heterogeneous networks with identical infectivity.* Physics Letters A, 2007. **364**(3-4): p. 189-193.

13. Zhou, T., et al., *Behaviors of susceptible-infected epidemics on scale-free networks with identical infectivity.* Physical Review E, 2006. **74**(5): p. 56109.

14. Kauffman, S., *Homeostasis and differentiation in random genetic control networks.* Nature, 1969. **224**(5215): p. 177-8.

15. Kauffman, S.A., *Metabolic stability and epigenesis in randomly constructed genetic nets.* J Theor Biol, 1969. **22**(3): p. 437-67.

16. Hopfield, J.J. and D.W. Tank, *Computing with neural circuits: A model.* Science, 1986. **233**(4764): p. 625.

17. Hopfield, J.J., *Neural networks and physical systems with emergent collective computational abilities.* Proc Natl Acad Sci U S A, 1982. **79**(8): p. 2554-8.

18. Barabsi, A.L., et al., *Evolution of the social network of scientific collaborations.* Physica A: Statistical Mechanics and its Applications, 2002. **311**(3-4): p. 590-614.

19. Feld, S., *The focused organization of social ties.* American Journal of Sociology, 1981. **86**(5): p. 1015-1035.

20. Galaskiewicz, J., *Social organization of an urban grants economy: A study of business philanthropy and nonprofit organizations*. 1985: Academic Press.

21. Jeong, H., et al., *The large-scale organization of metabolic networks.* Nature, 2000. **407**(6804): p. 651-654.

22. Farkas, I., et al., *The topology of the transcription regulatory network in the yeast, Saccharomyces cerevisiae.* Physica A: Statistical Mechanics and its Applications, 2003. **318**(3-4): p. 601-612.

23. Ito, T., et al., *A comprehensive two-hybrid analysis to explore the yeast protein interactome.* Proceedings of the National Academy of Sciences of the United States of America, 2001. **98**(8): p. 4569.

24. Jeong, H., et al., *Lethality and centrality in protein networks.* Nature, 2001. **411**(6833): p. 41-42.

25. Baird, D. and R. Ulanowicz, *The seasonal dynamics of the Chesapeake Bay ecosystem.* Ecological Monographs, 1989. **59**(4): p. 329-364.

26. Cohen, J., et al., *Community food webs: data and theory*. 1990: Springer.

27. Jordano, P., J. Bascompte, and J. Olesen, *Invariant properties in coevolutionary networks of plant–animal interactions.* Ecology Letters, 2003. **6**(1): p. 69-81.

28. Chen, Q., et al. *The origin of power laws in Internet topologies revisited*. 2002: IEEE.

29. Faloutsos, M., P. Faloutsos, and C. Faloutsos. *On power-law relationships of the internet topology*. 1999: ACM.

30. Bonifazi, P., et al., *GABAergic hub neurons orchestrate synchrony in developing hippocampal networks.* Science, 2009. **326**(5958): p. 1419.

31. Sporns, O., *Network analysis, complexity, and brain function.* Complexity, 2002. **8**(1): p. 56-60.

32. White, J., et al., *The structure of the nervous system of the nematode Caenorhabditis elegans.* Philosophical Transactions of the Royal Society of London. B, Biological Sciences, 1986. **314**(1165): p. 1.

33. Axelrod, R., et al., *Coalition formation in standard-setting alliances.* Management science, 1995: p. 1493-1508.

34. Guimera, R., et al., *Self-similar community structure in a network of human interactions.* Physical Review E, 2003. **68**(6): p. 065103.

35. Barabsi, A. and R. Albert, *Emergence of scaling in random networks.* Science, 1999. **286**(5439): p. 509.

36. Watts, D. and S. Strogatz, *Collective dynamics of 'small-world'networks.* Nature, 1998. **393**(6684): p. 440-442.

37. Watts, D., *Small worlds: the dynamics of networks between order and randomness*. 2004: Princeton Univ Pr.

38. Erd s, P. and A. R nyi, *On random graphs I.* Publ. Math. Debrecen, 1959. **6**: p. 290–297.

39. Ilachinski, A., *Cellular automata: a discrete universe*. 2001: World Scientific Pub Co Inc.

40. Hebb, D., *The organization of behavior: A neuropsychological theory*. 2002: Lawrence Erlbaum.

41. Newman, M.E.J., *The structure and function of complex networks.* SIAM review, 2003: p. 167-256.

42. Newman, M.E.J., *Assortative mixing in networks.* Physical Review Letters, 2002. **89**(20): p. 208701.

43. Barab si, A. and R. Albert, *Emergence of scaling in random networks.* Science, 1999. **286**(5439): p. 509.

44. Bollob s, B. and O. Riordan, *Mathematical results on scale-free random graphs.* Handbook of graphs and networks, 2002: p. 1–34.

45. Contreras, D., *Electrophysiological classes of neocortical neurons.* Neural Networks, 2004. **17**(5-6): p. 633-646.

46. Hendry, S., et al., *Neuropeptide-containing neurons of the cerebral cortex are also GABAergic.* Proceedings of the National Academy of Sciences of the United States of America, 1984. **81**(20): p. 6526.

47. Mason, A. and A. Larkman, *Correlations between morphology and electrophysiology of pyramidal neurons in slices of rat visual cortex. II. Electrophysiology.* Journal of Neuroscience, 1990. **10**(5): p. 1415.

48. Parnavelas, J., et al., *Separate progenitor cells give rise to pyramidal and nonpyramidal neurons in the rat telencephalon.* Cerebral Cortex, 1991. **1**(6): p. 463.

49. Rockel, A., R. Hiorns, and T. Powell, *The basic uniformity in structure of the neocortex.* Brain, 1980. **103**(2): p. 221.

50. Qu, X., M. Aldana, and L.P. Kadanoff, *Numerical and theoretical studies of noise effects in the Kauffman model.* Journal of Statistical Physics, 2002. **109**(5): p. 967-986.

51. Miranda, E. and N. Parga, *Noise effects in the Kauffman model.* EPL (Europhysics Letters), 1989. **10**: p. 293.

52. Simon, H.A., *Models of man.* 1957.

53. Itzhack, R., et al., *Empirical extraction of mechanisms underlying real world network generation.* Physica A: Statistical Mechanics and its Applications, 2010.

54. Guillery, R., *Binocular competition in the control of geniculate cell growth.* The Journal of Comparative Neurology, 1972. **144**(1): p. 117-129.

55. Miller, K., *Synaptic economics: Competition and cooperation in correlation-based synaptic plasticity.* Neuron, 1996. **17**: p. 371-374.

56. Lazarsfeld, P. and R. Merton, *Friendship as a social process: A substantive and methodological analysis.* Freedom and control in modern society, 1954. **18**: p. 66.

57. McPherson, M., L. Smith-Lovin, and J. Cook, *Birds of a feather: Homophily in social networks.* Annual Review of Sociology, 2001. **27**: p. 415-444.

58. Friedkin, N., *A structural theory of social influence*. 1998: Cambridge Univ Pr.

59. Biham, O., et al., *Generic emergence of power law distributions and Lévy-Stable intermittent fluctuations in discrete logistic systems.* Physical Review E, 1998. **58**(2): p. 1352.

Figure captions

Figure 1. Degree distribution of networks resulting from simulations with different parameter sets. The X axis is the degree. The Y axis is the probability of each degree. The basic network contains the same parameters (last column Table 1). Each subplot shows the degree distribution of networks resulting from different simulations, which differ in one of the parameters. The parameter range used is described in the third column of Table 1. The parameters varied in each subplot are (decreasing order, left column first): Initial connectivity a, correlation cutoff b, nodes number c, iteration steps d, noise e, positive edges rate f, history size g and average degree h. The two parameters that influence the network power law degree distribution are the correlation cutoff and positive edge rate.

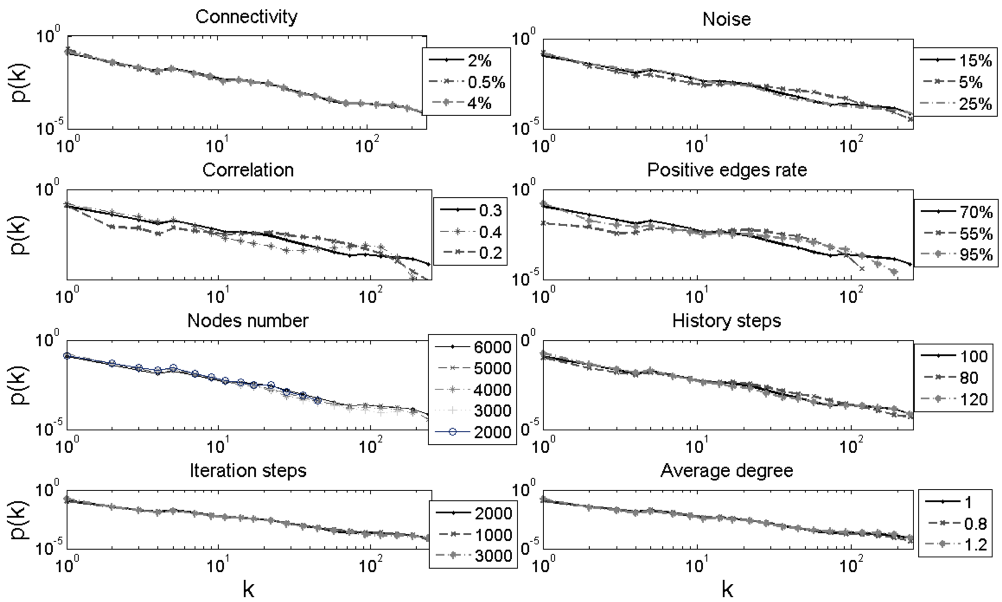

Figure 2. Average distance histogram for three networks resulting from simulations with parameter as described in Table 1. One network has no delay and the two other have delays of 1 and 10 iterations. All these networks have short average distances.

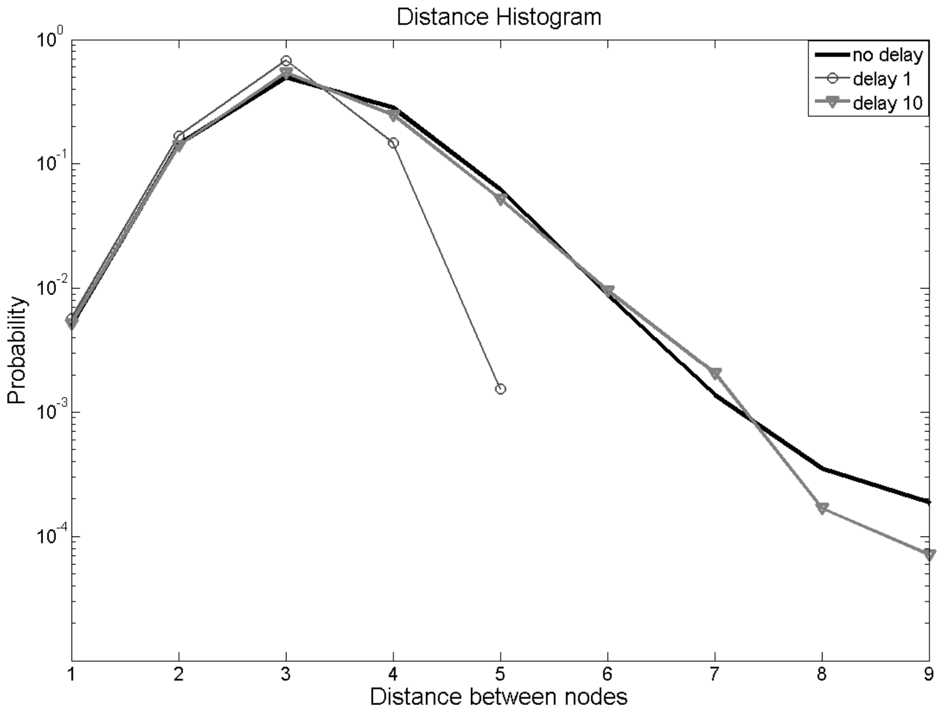

Figure 3. Indegree and outdegree distributions in simulation. The indegree has approximately a scale free distribution, while the outdegree has an approximate Gaussian distribution, as is indeed observed in many realistic networks.

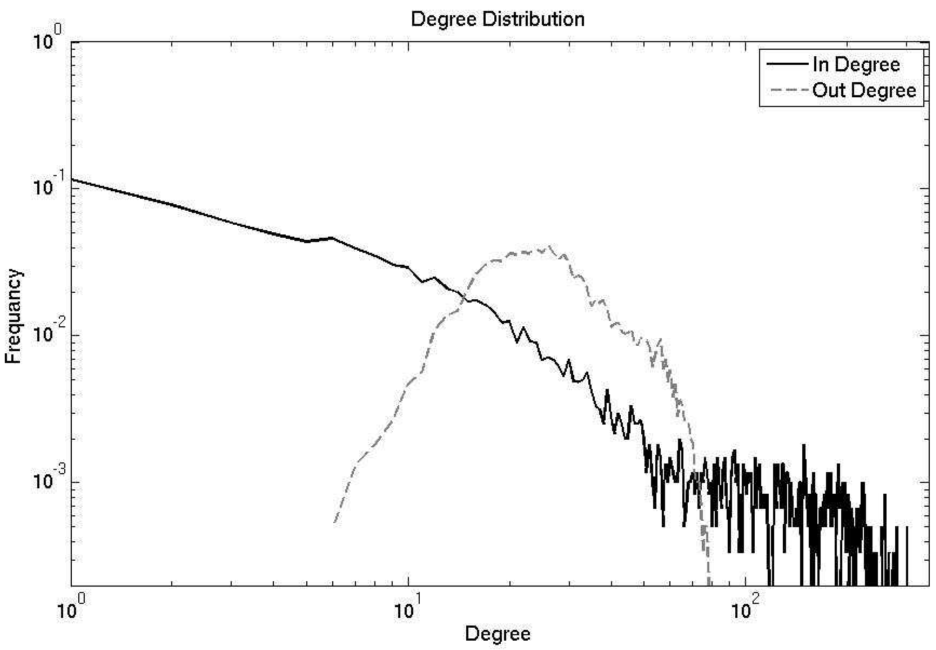

Figure 4. Relation between the degree of nodes and their clustering coefficient, and a curve representing the -1 power fit. While the relation does not have a precise -1 power, it is not far from it.

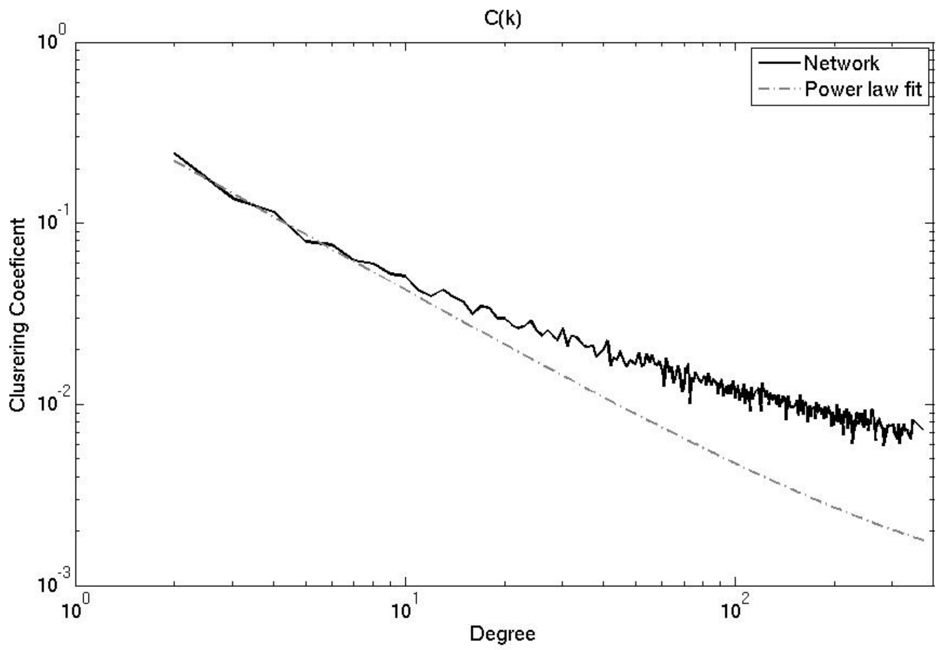

Figure 5. Degree distribution for the network obtained in the simulation with equal probability of the edge to be either positive or negative (*V_p=V_n*). All other simulation parameters are as listed in Table 1. The X axis shows the node degree and Y axis is the probability of each degree. The network does not have a power law degree distribution. The degree distribution stay Gaussian as can be shown in the fitting curve in the graph. The full line is a Gaussian fit.

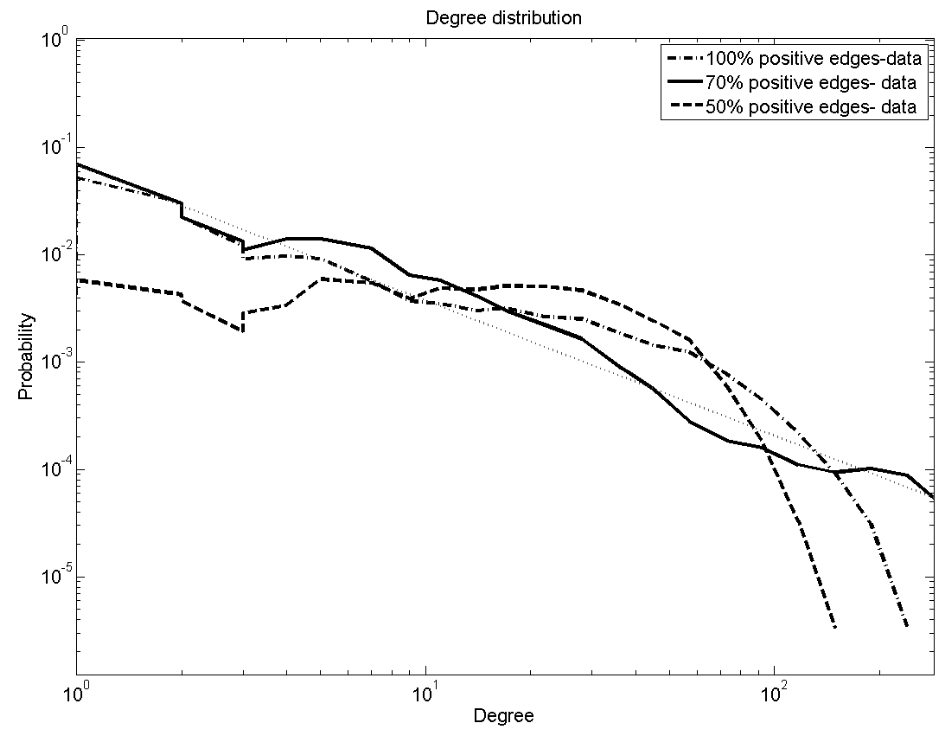

Figure 6. Average degree of nodes above and below the correlation cutoff as a function of the fraction of positive edges in the network. The average is over the two nodes in the pair. One can clearly see that the maximal difference between the degree of high and low correlation pairs is in the case of 70% positive edges. When the chances of positive and negative edges are equal, the degree of high and low correlation node pairs is equal.

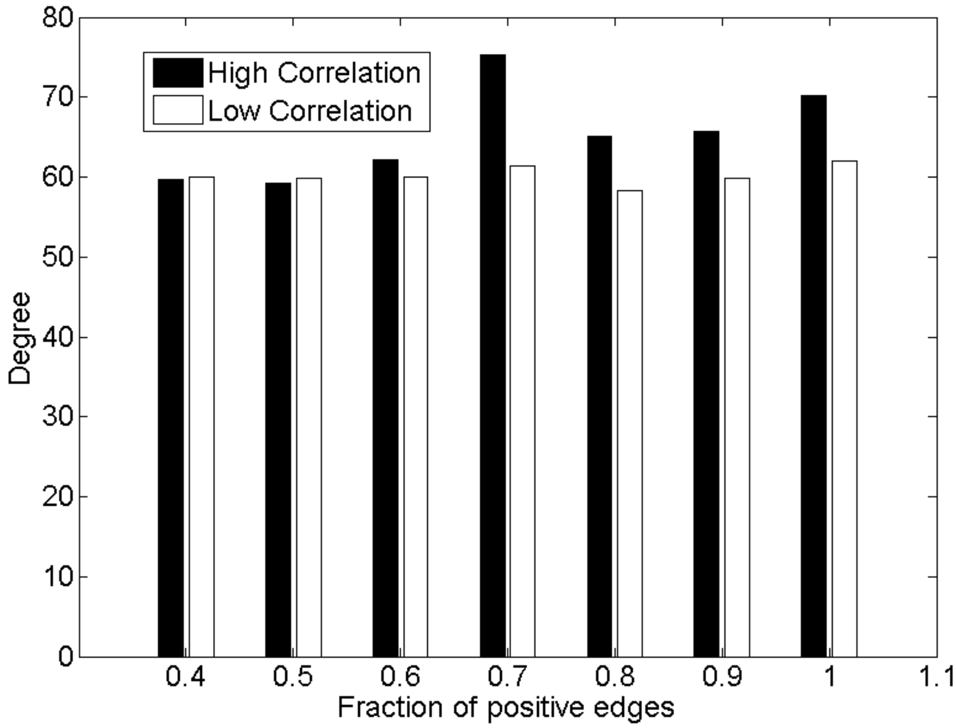

Figure 7. Fraction of above threshold node pairs for different node pair categories: Nodes that are not connected one to each other or to a common neighbor, nodes connected one to each other and to a common neighbor, nodes only connected to a common neighbor. In this third group, we separated the node pairs based on the connection pathway. A and B are the pair of nodes and C is the common neighbor.

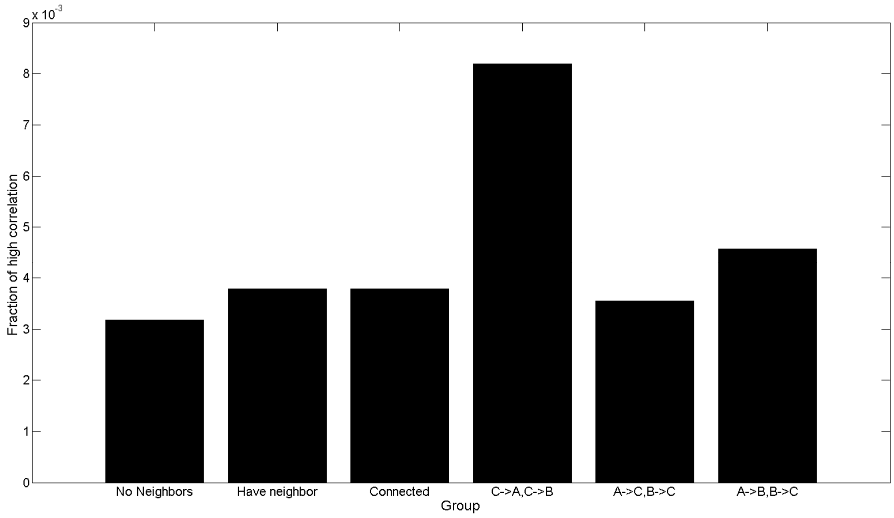

Figure 8. Degree distribution for 1 and 10 time step delayed correlation. Simulation parameters are as in Table 1. The X axis represents graph is log-log scale. The two networks have an approximate power law degree distributions, although the power law is not as clean as in the simulation without a delay.

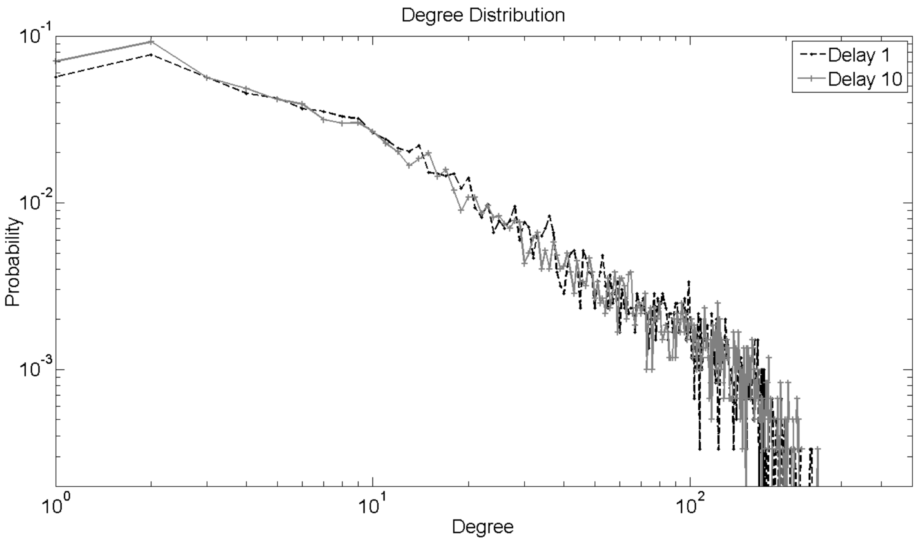

Figure 9. Average similarity in unconnected nodes in five different snapshots. In each snapshot, we analyzed node pairs that are not connected in the current snapshot, and we separated them into nodes that will be connected in the next snapshot and nodes that will not be connected in the next snapshot. Nodes that will be connected have a much higher (p<1.e-100) similarity than nodes that will not be connected in all snapshots.

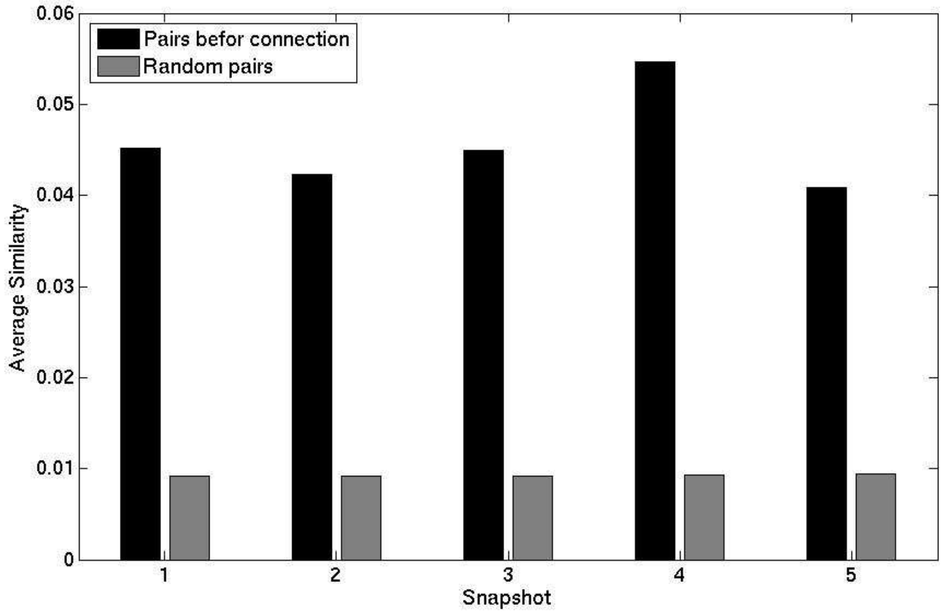

Table 1. Parameters values used during the simulation. The first row contains the parameters symbol, the second row a short description, the third row is the parameter range used to check the network stability. The upper and lower value results are shown in Figure 1. The fixed values in the last row are the values used for the parameters during all of the simulations, except for the parameter that was varied. The number of time steps used for the node correlation calculation and the number of edges chosen at every iteration, in order to check edge addition, were not part of the stability analysis and were fixed in all simulations.

Parameter	Description	Range	Fixed value
N	Number of nodes	2000-6,000	6,000
V	Connectivity	0.5%-4%	2%
V_p	Positive edges rate	55%-95%	70%
No	Noise rate	5%-25%	15%
Cor	Correlation threshold	0.2-0.4	0.3

H	Time step for correlation calculation	Const during the simulations	100
V_c	Number of edges checked in each iteration	Const during the simulations	160000