Ann. Zool. Fennici 45: 441–448 ISSN 0003-455X (print), ISSN 1797-2450 (online)
Helsinki 29 October 2008

Plasticity in individual choice in social network evolution

Kah Loon Ng

Department of Mathematics, 2 Science Drive 2, National University of Singapore, Singapore, 117543 (e-mail: matngkl@nus.edu.sg)

Received 16 Sep. 2007, revised version received 2 Dec. 2007, accepted 14 Nov. 2007

Ng, K. L. 2008: Plasticity in individual choice in social network evolution. — *Ann. Zool. Fennici* 45: 441–448.

Constant re-evaluation of social affiliation is known to cause populations of individuals with different predetermined affiliation preferences to diverge into different network structures. In this study, rather than assigning to each individual a fixed affiliation preference, held throughout the duration of the dynamic network evolution, individuals were allowed an initial "learning period" during which they compared their own relative success, using each of three strategies, at maximizing their social status under three different metrics. Based on the outcomes from this learning period, individuals then chose one particular strategy. The organizational success and stability of the resulting populations was seen to be higher than those of the populations of individuals whose behaviors were predetermined. This indicates that individual-level evaluation and strategy choice in social affiliation preferences can yield strong benefits to the organizational success of the population as a whole.

Introduction

The evolution of a network has been the focus in various fields of study ranging from sociology (Snijders 1997, Doreian 2006), computer science (Huberman & Adamic 1999, Vázquez *et al*. 2002), physics (Réka & Barabási 2002, Dorogovtsev & Mendes 2003, Pastor-Satorras & Vespignani 2004) and ecology (Altizer *et al*. 2003, Fefferman & Ng 2007). A general similarity in these studies is that individuals, be it humans, web pages, social insects or otherwise, are represented by vertices in a network and links exist between pairs of these vertices if there is sufficient physical contact or if they are connected at a more abstract level (e.g. links between web pages).

To examine the role of individual choice in the organizational success of dynamic networks Fefferman and Ng (2007) experimented with social networks comprised of a fixed number of individuals, where each individual consistently re-evaluates its set of existing social affiliations according to one of three centrality measures borrowed from social network theory (Freeman 1979, Wasserman & Faust 1994). These three measures, betweenness, closeness and popularity were chosen to represent the ability of an individual to assess the social status of its affiliations at both a basic (closeness and popularity) and a more complicated/evolved (betweenness) levels. In this study, the overall organizational success of the population was measured in terms of these same three measures. Besides homogeneous populations where every individual had the same type of affiliation preference, a heterogeneous population, where each of the three affiliation preference was held by approximately one third of the population, was also considered. This heterogeneous model was chosen to represent diversity in social preference, potentially due to genetic mutations directly altering the individual's preference measure, or alternatively due to phenotypic plasticity in the expression of an individual's affiliation preference.

In the context of social evolution, an individual's affiliation preferences and resulting affiliation networks can affect multiple scales of organizational success ranging from local, individual-level centrality, to overall, population-level structures. Building upon the heterogeneous population model of Fefferman and Ng (2007), the work presented here introduced an initial period of "learning" into the dynamic affiliation process. During this initial period, each individual was able to "remember" the change in its centrality measure (for all three types) and, based on that information, compute which affiliation strategy resulted in the largest average increase to its own centrality measure over the duration of the learning period. In this way, these studies examined whether or not the ability of an individual to choose its affiliation preference, rather than adhere to a predetermined strategy (as in Fefferman & Ng 2007), can significantly impact stability and organizational success of the population.

Further, given individual choice, it is possible that the organizational success of a population would be maximized under conditions in which all individuals chose uniformly, producing a relatively homogenous population by voluntary consensus. If this were the case, we would expect that the evolution of novel affiliation strategies would be difficult under any circumstances of intense selective pressure to maximize organizational success. On the other hand, should group-level success be maximized, or equivalently maintained, in cases in which individual choices did not concur, then even scenarios of intense selective pressure could be expected to produce variation in genotypically determined social affiliation behaviors.

Methods

The three social evaluation measures discussed here are identical to those previously described (Fefferman & Ng 2007), where readers are referred to for precise definitions of various social network theory terminologies used. Also as mentioned in Fefferman and Ng (2007), the use of betweenness, closeness and popularity as the three measures and affiliation preferences was not meant to be definitive or exhaustive. They merely represent a diverse set of measures through which we have begun to investigate the complex problem of social network dynamics and further work in this area, including expansion to a broader variety of social network centrality measures, will certainly be needed.

Building on the notations already defined in the aforementioned article, a population of n individuals with directed relationships is represented by a digraph G with vertices $V = \{v_1, ..., v_n\}$. The directed graph resulting from G after t time steps of affiliation shifts is denoted by G_t. The Popularity (Closeness and Betweenness, respectively) measure of vertex v_i in G_t is denoted by $P(v_i^t)$ ($C(v_i^t)$ and $B(v_i^t)$, respectively) while the population-wide Popularity (Closeness and Betweenness, respectively) measure of G_t is denoted by $P(G_t)$ ($C(G_t)$ and $B(G_t)$, respectively). A vertex that prefers Popularity (Closeness and Betweenness, respectively) as a measure of social affiliation is said to be 'a P- (C- and B-, respectively) vertex', or else is 'of type P' (C and B, respectively).

For all experimental models (*see* Table 1), the vertex affiliation preference for each v_i was determined at the outset of computation. With the exception of experiment 5, an individual assigned a particular preference would keep that preference throughout the entire simulation. Each vertex began with affiliations to five other randomly chosen vertices and, in each time step would evaluate the centrality of each of its five out-neighbors, then replace the two with the lowest measure (according to its affiliation prefence) at random from the rest of the population. (For the details of the process of re-evaluating and changing its set of out-neighbors, *see* Fefferman & Ng 2007). Experiments 1 to 3 represent populations with homogeneous affiliation preference, while experiment 4 models a population with approximately 1/3 of the vertices of each type.

To model a population in which individuals go through an initial phase of "learning", where affiliation preferences can be changed from one time step to another, a digraph G from experiment 5 begins with approximately 1/3 of the vertices of each type and then each vertex keeps track of the change in its centrality measure corresponding to its affiliation preference for that time step during the first twenty time steps of the iteration. More precisely, suppose v_i is initially assigned to be a P vertex with popularity measure $P(v_i^1)$ at time step 1. The largest popularity measure among all vertices of type P during the current time step

$$P_{\max}^1 = \max\left\{P\left(v_k^1\right)\middle| v_k \text{ is of type P during time step } 1\right\}$$

is also recorded. After v_i, as well as other vertices, have changed their affiliations according to their respective preferences, $P(v_i^2)$ and $P_{\max}^2$ are computed at time step 2. The change in v_i's popularity measure from time step 1 to 2, scaled respectively by $P_{\max}^1$ and $P_{\max}^2$ is computed as

$$\Delta P\left(v_i^1\right) = \frac{P\left(v_i^2\right)}{P_{\max}^2} - \frac{P\left(v_i^1\right)}{P_{\max}^1}.$$

Vertex v_i then randomly chooses its affiliation preference before changing its affiliations according to this new preference at the end of time step 2. The change in v_i's centrality measure corresponding to this new preference is again computed and recorded at time step 3. Similar computations are performed for each vertex and their respective affiliation preferences from time steps 1 through 20, resulting in the computations of $\Delta C(v_i^t)$ (respectively $\Delta B(v_i^t)$) for vertex v_i at time steps t where v_i is of type C (respectively B). At time step 21, each vertex v_i chooses an affiliation preference permanently for the remainder of the iteration by comparing the values of $\Delta P(v_i^t)$, $\Delta C(v_i^t)$ and $\Delta B(v_i^t)$ averaged over those time steps t between 1 to 20 where v_i was a vertex of that particular type. In other words, if $T_i^{\mathrm{P}}, T_i^{\mathrm{C}}$ and T_i^{B} are the number of times v_i is a type P, C or B vertex respectively during the first 20 time steps, we calculate

$$\overline{\Delta X}\left(v_i\right) = \frac{\sum \Delta X\left(v_i^t\right)}{T_i^{\mathrm{X}}}$$

for $X = \{P,C,B\}$ where the summation in the numerator is summed over all such t between 1 and 20 where v_i was of type X. In the event that v_i was never a type X vertex during the first 20 time steps, $\overline{\Delta X}(v_i)$ was set to be zero. Vertex v_i's permanent affiliation preference is determined by choosing the preference with the largest $\overline{\Delta X}(v_i)$.

At each time step t, in all experimental models, $P(G_t)$, $C(G_t)$ and $B(G_t)$ were computed. Each model was run for 200 time steps in order for the $P(G_t)$, $C(G_t)$ and $B(G_t)$ values to stabilize. For each of the 5 experiments, we performed Monte Carlo simulations with 100 iterations to ensure accuracy.

Due to the stochastic nature of experiment 5, each Monte Carlo iteration of the experiment resulted in slightly different compositions of vertices of each type. In order to investigate the effect of these different compositions of vertices on the overall network centrality measures, and how they compared with those previously reported (Fefferman & Ng 2007, reproduced here as experiments 1 to 4), the 100 Monte Carlo iterations of experiment 5 were divided into two classes. First, the 9 correlation coefficients were computed for each network centrality measure (Betweenness, Closeness and Popularity) and

Table 1. The different experiment models.

Experiment	Vertex affiliation preference	Total number of vertices	Number of B vertices	Number of C vertices	Number of P vertices
1	Betweenness	100	100	0	0
2	Closeness	100	0	100	0
3	Popularity	100	0	0	100
4	Heterogeneous	100	33	34	33
5	Heterogeneous	100	Varies	Varies	Varies

each of the number of each type of vertices (B, C and P) there are in the network for a particular Monte Carlo iteration. Thus for each centrality measure-vertex type pair, there were one hundred data points on which to perform a simple regression using the model $y = A + Bx$ where y is the network centrality measure and x is the number of vertices of a particular type.

Each Monte Carlo iteration i can then be given an index

$$V(i,x) = R_{b,x}n_b + R_{c,x}n_c + R_{p,x}n_p$$

for each centrality measure $x = B,C,P$ where in the equation above, $R_{b,x}$ ($R_{c,x}$ and $R_{p,x}$, respectively) is the correlation coefficient between the network centrality measure x and the number of vertices of type B (C and P, respectively) in Monte Carlo iteration i. For each centrality measure x, the 50 Monte Carlo iterations with the highest $V(i,x)$ were grouped into experiment 5H, while the remaining 50 iterations were grouped together as experiment 5L. This division in grouping was performed for experiment 5 only, since only in this experimental scenario was it pertinent to examine the possible effect the number of individuals of each affiliation preference type had on the total network centrality. Finally, the network centrality measures at each time step t were averaged over all 50 iterations in each of the two groups 5H and 5L to produce a representative value for experiment 5 itself at each t.

Results

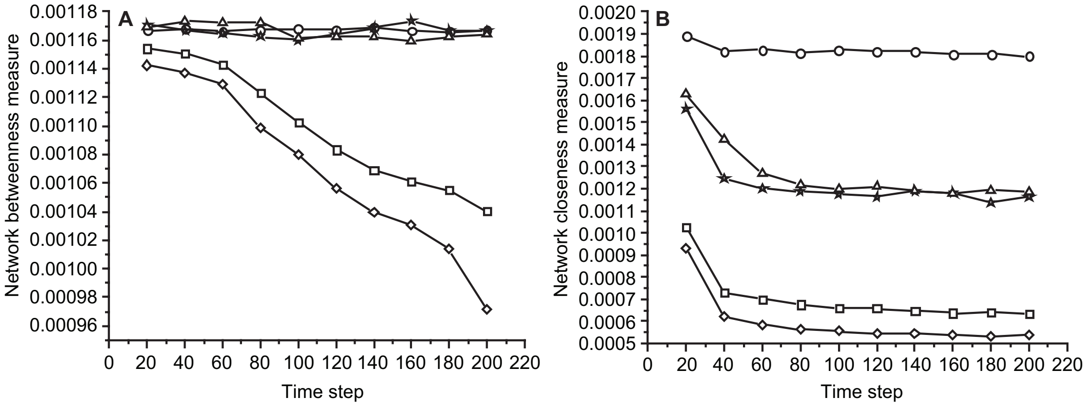
In experiment 5, though the population was initially equally divided among the three types of affiliation preferences, after the end of the learning period, there were significantly more type C vertices than either type B or P vertices while the number of type B and P vertices were not significantly different (*see* Table 2). This illustrates that within a population with non-uniform affiliation preferences, significantly more vertices experienced the greatest increase in their closeness centrality measure during the 20 time-step learning period, during which they changed and evaluated the success of their affiliation strategies. This might also imply that the closeness centrality measure of a vertex is least dependent on (and thus more robust to) the affiliation preferences of other vertices in the population. This difference in behavior of a C vertex (and consequently that of a C population) as compared with that of both the B and P vertices (B and P populations, respectively) mirrors the different resulting network structure that a C population has as compared with both the B and P populations (*see* Fefferman & Ng 2007). For example, the C population was observed to attain little organizational success when success was measured in terms of popularity whereas B and P populations converged to somewhat similar structures with high network popularity measure.

For both network betweenness and closeness centrality measures, the number of C vertices was seen to be positively correlated with the network measure while the number of B and P vertices correlated negatively with the network measure (*see* Table 3). This behavior was reversed when considering the network popularity measure (*see* Table 3). Furthermore, the correlation coefficients were generally larger (in magnitude) when considering network popularity measure than were those for the other two network measures.

Also from Table 3, we see that only the number of type C vertices was positively cor-

Table 2. Mean number of vertices of each type after each vertex chooses its permanent affiliation preference. For each pair-wise post test, the inequality denotes which type of vertices has significantly more than the other, the number in each bracket is the p value for that pair-wise test.

p value for 3-way test	Number of B vertices	Number of C vertices	Number of P vertices
	30.79	36.6	32.61
		Dunn's Post Test	
< 0.0001	C > B (< 0.001)	B ≈ P (> 0.05)	C > P (< 0.05)



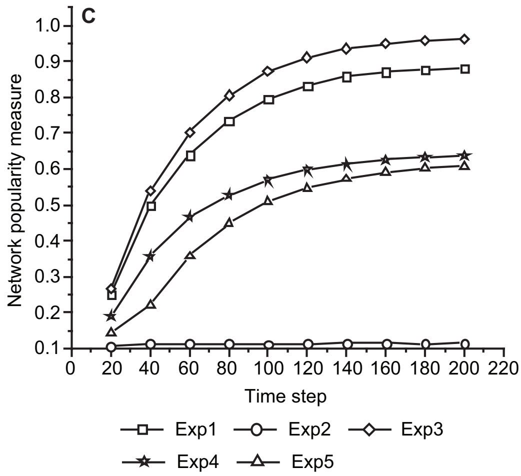


Fig. 1. Population-wide organizational success measured by (**A**) betweenness, (**B**) closeness and (**C**) popularity. As predicted by correlation coefficients shown in Table 3, vertices with different affiliation preferences have different effect on the network organization success under different measures. For example, considering network betweenness measure (panel **A**), homogeneous populations of type B or P causes the network measure to decrease rapidly while a homogeneous population of type C and all the heterogeneous populations stabilizes at a much higher network measure. On the other hand, for network popularity measure, heterogeneous populations only achieves about 60% of the success attained by homogeneous populations of both type B and P (panel **C**).

related with the network betweenness measure. This observation is further supported by the homogeneous populations consisting of type B and P individuals (experiments 1 and 3, respectively) have the lowest betweenness centrality (*see* Fig. 1A). A homogeneous population of type C-individuals (experiment 2) attained approximately similar network betweenness centrality than did a population with equal proportion of individuals of each type (experiment 4). Comparing both experiments 5L and 5H with experiment 4, all three experiments initially showed nearly equal network betweenness measures but rapidly diverged, with experiment 5H attaining a consistently higher centrality measure than experiment 4 (*see* Fig. 2A). However, though experiment 5L initially also showed a higher network betweenness measure than experiment 4, after the hundredth time step, experiment 4 had the greater network measure. Experiment 5L had

Table 3. Correlation coefficients from regressing the number of vertices of each type with each network centrality measure averaged over the last 20 time steps (180 to 200). The number in the brackets represents p from the simple regression.

	Number of B vertices	Number of C vertices	Number of P vertices
Network betweenness measure	–0.162 (0.106)	0.438 (< 0.0001)	–0.431 (< 0.0001)
Network closeness measure	–0.335 (0.0006)	0.683 (< 0.0001)	–0.597 (< 0.0001)
Network popularity measure	0.621 (< 0.0001)	–0.966 (< 0.0001)	0.711 (< 0.0001)

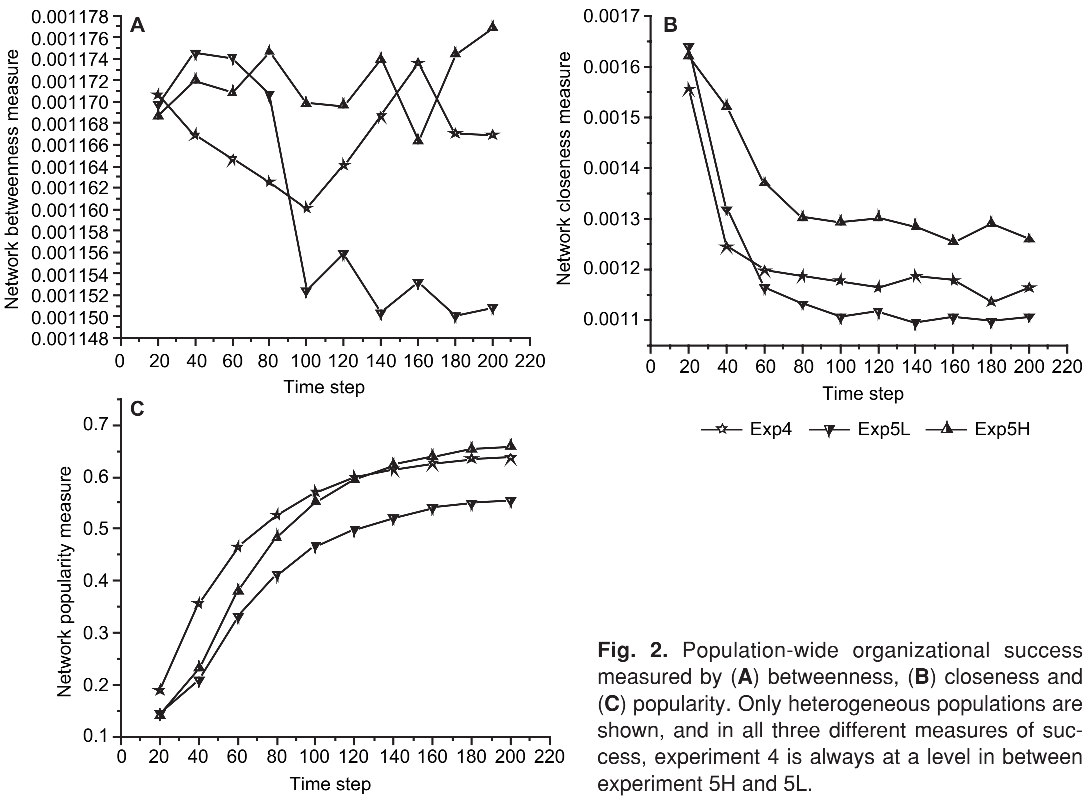


Fig. 2. Population-wide organizational success measured by (**A**) betweenness, (**B**) closeness and (**C**) popularity. Only heterogeneous populations are shown, and in all three different measures of success, experiment 4 is always at a level in between experiment 5H and 5L.

a lower proportion of type C vertices (an average of 32.7 out of the total network size of 100) and a higher proportion of type B and P vertices (an average of 32.7 and 36.4, respectively) than in the total averages over all experiments contributing to Experiment 5 as a whole (yielding an average of 30.8, 36.6, and 32.6 vertices for types B, C and P, respectively). This skew caused the network betweenness measure to decrease after the twentieth time step (after all individuals have fixed their affiliation preference), much like experiments 1 and 3, (although not to the same degree) where the high proportion of type B or P vertices caused a decrement to the network betweenness measure.

The organizational success of the population under the closeness metric exhibited correlations between the number of vertices of each type and the network measure which were all larger in magnitude (either positively or negatively) than those observed for the betweenness measure discussed above (*see* Table 3). The number of type C vertices was again seen to be positively correlated with the network closeness measure and, as a result, experiment 2 had the highest level of organizational success (*see* Fig. 1B). As predicted by the negative correlation between the number of type B or P vertices and the network measure, experiments 1 and 3 achieved the lowest level of success. Finally, the network popularity measure was observed to behave in exactly the opposite manner: experiments 1 and 3 achieved the highest level of organizational success while experiment 2 had the lowest network measure.

Interestingly, although experiment 5H eventually converged to a network popularity measure higher than that attained by experiment 4, it did not achieve this until about the 120th time step (*see* Fig. 2C). This was different from what was seen when network success was measured in terms of closeness (*see* Fig. 2B), where experiment 5H began with a higher organizational success and consistently maintained this advantage over experiment 4 throughout the 200 time steps. When network success was measured in terms of betweenness, experiment 5H also generally achieved higher organizational success than experiment, with the only exceptions occurring during the initial 20th time step and the 160th time step (*see* Fig. 2A).

Interestingly, the incorporation of a learning period, and the concomitant flexibility of individuals being allowed to make their own choice of affiliation strategy (experiment 5), produced no substantially different results in the organizational success of the network measured under either the Betweenness or Closeness metrics, and in fact caused a substantial decrease in the organizational success as measured by Popularity as compared to those levels of success attained in an evenly heterogeneous population with no learning period (experiment 4) (*see* Fig. 1).

Since each independent realization of experiment 5 yielded a different distribution of individual types, the overall organizational success may have been the average of very separate and different results of particular distributions of preference types (unlike the results of any of the other experiments). Therefore, to examine the relative success of populations with individual learning as compared with that in which affiliation preferences were determined from the beginning, the average success of 5H and 5L were also examined. Under all three measures, the network success achieved after convergence by experiment 4 was always smaller than that attained by experiment 5H but was larger than experiment 5L (*see* Fig. 2). However, the benefits of the learning period were seen to be most advantageous when considering the network closeness measure, resulting in an 8.14% increment in total network centrality. This was in contrast to the 0.85% and 3.46% attained under the network betweenness and popularity measures respectively.

Discussion

Experiments 1 to 3 examined populations in which individuals all employed the same affiliation strategy. In experiment 4, we investigated the relative successes of heterogeneous populations, in which a single population included some individuals with each of the three affiliation strategies, determined before the initiation of computation and held constant throughout the experiment. These populations and their subsequent organizational successes under each of the three measures therefore provided bases for comparison against which we were able to compare the results from experiment 5, in which individuals were allowed the learning period before selecting their own affiliation strategy.

Although no single affiliation strategy was chosen by all individuals within the entire population after the learning period, there were significantly more individuals choosing closeness as their affiliation strategy than there were choosing either popularity or betweenness. With the appropriate mixture of individuals with different affiliation strategies (experiment 5H), the population was able to attain a higher level of organizational success than was achieved by the neutral model of randomly assigning equal proportions of individuals to each of the three strategies (experiment 4). Given a choice of which affiliation preference to adopt, individuals within the same population not only made different decisions, but if the organizational success of the population was measured after the network structure converged, it was most successful when the population self-distributed into a mix of different individual types (experiment 5H, *see* Fig. 2). Interestingly, for network organizational success measured using betweenness, although the number of C vertices was the only one positively correlated with the network measure (*see* Table 3), a homogeneous population of C vertices did not attain the highest organizational success (experiment 2, *see* Fig. 1A). This leads to the intriguing notion that there really could be "too much of a good thing": to achieve the highest possible levels of population-level success, a population would require a sufficient diversity of preference types to complement the C vertices before achieving maximal population-level success.

Finally, the introduction of a learning period yielded only a barely observable benefit under the closeness measure, and in fact effected a small decrease in success under betweenness (*see* Fig. 1), The magnitude of these alterations in success was seen to depend upon the measure used, and was also clearly dependent on the distribution of preference types within the population (*see* Fig. 2).

While diversity in individual behavior in a population does, itself, have a significant impact on the organizational success of the population, under the closeness metric of success, we see that a learning period improved the overall success of the population over the benefits conferred merely from diversity in preference types. Similarly, both predetermined heterogeneous preference populations and populations with learning achieved similar levels of success under the betweenness measure. Building on the investigations of Fefferman and Ng (2007), these results imply that the evolution of behavioral plasticity, rather than the genetic predetermination of social behaviors, could have been actively selected for if the operative pressure for population-level success was betweenness or closeness. Under selective pressure to maximize popularity, we can infer that individual behavioral phenotypes would have been actively selected against on a population level.

These measures are only three potential metrics for both individual- and population-level success, but the implications for the evolution of individual choice and social complexity are far broader. Together, these results provide a first glimpse into the possible role of behavioral plasticity, individual choice and the evolution of social complexity.

The need for simulation methods in dynamic network modeling

The study of individual behaviors and how these behaviors can affect the evolving and emergent network structure of a population is very complex, even for a population of small size. Analytically tractable models would necessarily involve the making of simplifying assumptions involving 'average behaviors' which would render the models unrealistic. The assumption of these averages would make impossible investigations into experimentally learned behaviors based upon stochastic events, such as those investigated within this paper. Since it is usually equally impossible for laboratory based biologists to empirically manipulate the behaviors of real-life populations, simulation methods such as those described in this study provide a unique opportunity for empirical investigation into the effects of learning and behavior on social network evolution.

Acknowledgements

The author would like to thank DIMACS where this work is completed while he was visiting and also Nina H. Fefferman for her help and encouragement in the process of preparation of this manuscript.

References

Altizer, S., Harvell, D. & Friedle, E. 2003: Rapid evolutionary dynamics and disease threats to biodiversity — *Trends Ecol. Evol.* 18: 589–596.
Doreian, P. 2006: Actor network utilities and network evolution. — *Soc. Networks* 28: 137–164.
Dorogovtsev, S. N. & Mendes, J. F. F. 2003: *Evolution of networks: From biological nets to the Internet and WWW* — Oxford Univ. Press, Oxford, UK.
Freeman, L. C. 1979: Centrality in social networks: concept clarification — *Soc. Networks* 1: 215–239.
Fefferman, N. H. & Ng, K. L. 2007: The role of individual choice in the evolution of social complexity — *Ann. Zool. Fennici* 44: 58–69.
Huberman, B. A. & Adamic, L. A. 1999: Growth dynamics of the world-wide web — *Nature* 401: 131.
Pastor-Satorras, R. & Vespignani, A. 2004: *Evolution and structure of the internet: A statistical physics approach* — Cambridge Univ. Press, Cambridge, UK.
Réka, A. & Barabási, A. L. 2002: Statistical mechanics of complex networks — *Rev. Mod. Phys.* 74: 47-97.
Snijders, T. A. B. 1997: Stochastic actor-oriented models for network change — In: Doreian, P. & Stokman, F. (eds.), *Evolution of social networks*: 185–208. Gordon and Breach Science Publishers, The Netherlands.
Vázquez, A., Pastor-Satorras, R. & Vespignani, A. 2002: Large scale topological and dynamical properties of the internet — *Phys. Rev. E* 65:066130.
Wasserman, S. & Faust, K. 1994: *Social network analysis: methods and applications* — Cambridge Univ. Press, Cambridge, UK.