GTPA: A Generative Model For Online Mentor-Apprentice Networks

Muhammad Aurangzeb Ahmad, David Huffaker, Jing Wang, Jeff Treem, Marshall Scott Poole, Jaideep Srivastava

mahmad@cs.umn.edu, davidhuf@umich.edu, jingwang2008@u.northwestern.edu, jtreem@u.northwestern.edu, mspoole@uiuc.edu, srivasta@cs.umn.edu

[1] Department of Computer Science, University of Minnesota

[2] University of Michigan

[3] University of Illinois Urbana-Champaign

[4] Northwestern University

Abstract

There is a large body of work on the evolution of graphs in various domains, which shows that many real graphs evolve in a similar manner. In this paper we study a novel type of network formed by mentor-apprentice relationships in a massively multiplayer online role playing game. We observe that some of the static and dynamic laws which have been observed in many other real world networks are not observed in this network. Consequently well known graph generators like Preferential Attachment, Forest Fire, Butterfly, RTM, etc., cannot be applied to such mentoring networks. We propose a novel generative model to generate networks with the characteristics of mentoring networks.

Introduction

There is a large body of literature on analysis of complex networks in the real world (Wasserman 1994). Empirical work suggests that there are many commonalities among these networks such as a shrinking diameter (Albert 1999) or power law distributions (Barabasi et al 2002). Given such common characteristics researchers have proposed several graph generating mechanisms for these networks (McGlohon 2008, Akoglu 2008, Akoglu 2009). While a wide range of networks including blogs, patents, and scientific citations have been studied, rarely—if ever—have scholars examined networks consisting of mentor-apprentice dyads. In this paper, we empirically analyze a mentoring network and show that it does not share many of the characteristics of "regular" networks. We frame our conception of this network in terms of exchange theory and then develop a generative model that best simulates it. We rely on data from EverQuest II (EQ2), a fantasy-based massively multiplayer online role playing game (MMO) where tens of thousands of players can simultaneously interact with one another while engaging in activities such as completing quests and battling monsters. Many of these game activities require players to collaborate and team up in order to be successful. More information on the game is available on EverQuest II's official website.[1]

We inquire as to the nature of mentoring in large-scale virtual worlds. Is it primarily a one-on-one phenomenon, in which mentor and apprentice form a strong mutual relationship? This is how it is portrayed in much of the literature. However, recent research on networks suggests that many phenomena previously regarded as primarily individual-level exchanges are in fact more complex. Rather than being a one-to-one relationship, mentoring may be more communal in nature. With this view mentoring is conducted by a larger community which gives the apprentice coaching, and the apprentice is embedded in a mentoring community rather than connected to a single mentor. In order to answer this question, we rely on a temporal data-set of a social network of mentoring links between all players over an eight-month period.

This analysis enables us to gain insights into mentoring in online games and, it can be argued, more generally. The analysis points to some key problems with widely accepted network models for complex relationships such as mentoring. These models have been developed primarily on relatively simple relationships, such as internet connectivity or small world phenomena. Mentoring is a more complex relationship than these graphs represent and thus represents an excellent context for inquiry into fundamental properties of networks.

We present a generative model GTPA (Generative Temporal Preferential Attachment) which can recreate a set of desired features that are observed in mentoring networks, which can not be explained by other models such as Preferential Attachment, Forest Fire, Butterfly, RTM. The models we will employ in this analysis are centered on exchange relationships, which may be multi-tiered and multi-level. Prior to developing the models we will consider how mentoring can be conceptualized as 'exchange'. We will then consider two fundamentally

[1] http://everquest2.station.sony.com/

different models of exchange that provide basic frameworks for dynamic modeling of mentoring networks.

Mentoring in Virtual Worlds as Exchange

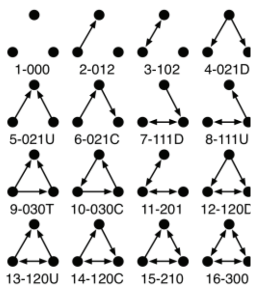


Figure 1. Structure of 16 possible triads

Definitions of mentoring range from basic aide to formalized organizational arrangements (Chao, Walz & Gardner, 1992), and instances of mentoring have been found in a variety of educational, organizational, and social settings. Research on mentoring has predominantly focused on the respective costs and benefits for both mentors and apprentices. For example, Hunt & Michael (1983) discuss how mentoring in organizational settings can increase work competence -individual salary and job satisfaction, while Kram (1983) finds that mentoring also serves psychosocial functions including providing friendship and counseling.

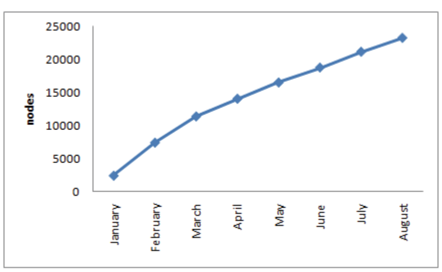
There is reason to believe that many real-world phenomena such as mentoring may occur in much the same way in virtual worlds. Studies of socializing, trust, and expertise in virtual worlds suggest that causality in virtual worlds is similar to that in the real world (Williams et al., 2006; Wang et al., 2009). This is coupled with the fact that EQ2 has an explicit design feature which encourages mentoring relationships.

In EQ2, character levels range from 1–70 and higher-level players can select a lower-level player and enter in a mentoring relation, in which their level is lowered to match their apprentice. This allows apprentices to benefit from the experience and abilities of their mentors when fighting monsters or completing quests. It also allows friends to play together regardless of level differences, or players in the same guild to help guild-mates complete difficult encounters or level-up in order to tackle high-level raid encounters. In addition, mentoring offers bonus points for both mentors and apprentices, which expands their overall achievement in the game. This suggests that mentoring in EQ2 also serves both social and performance-enhancing functions. Though mentoring can be established, maintained, or dissolved for a variety of reasons, at the foundation of these interactions is some type of exchange among individuals (Young and Perrewé, 2000). Treating mentoring as an exchange relationship allows us to consider the dynamic and interdependent nature of mentoring in organizational settings.

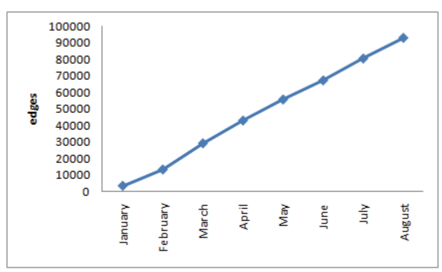
Exchange as a Basis for Network Generation

Monge and Contractor (2003) review much of the literature on exchange theory as a theory of networking. They note that while a great deal of work has been done on exchange relationships between individuals and among groups, larger networks are generally assumed to simply be the sum of dyadic relationships. This assumes that exchange operates primarily at the micro-level and that resulting networks will be extensions and complexities of micro-level relationships. As such, this approach relies heavily on discrete exchanges and does not reflect all of the ways exchange networks may evolve over time

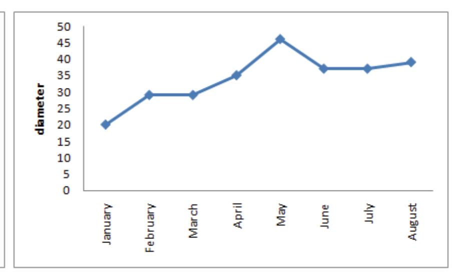
Ekeh (1974) distinguishes two versions of social exchange models that trace back to individualistic and collectivistic traditions in social theory. *Restricted exchange* is organized around exchanges between two parties, each of whom benefits directly from interactions



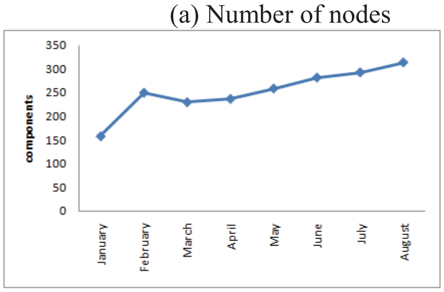
(a) Number of nodes



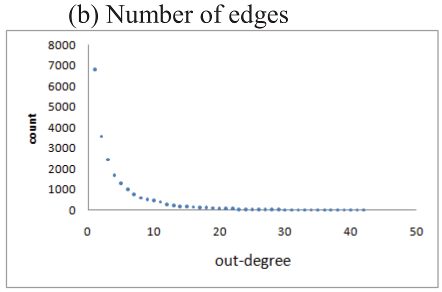
(b) Number of edges



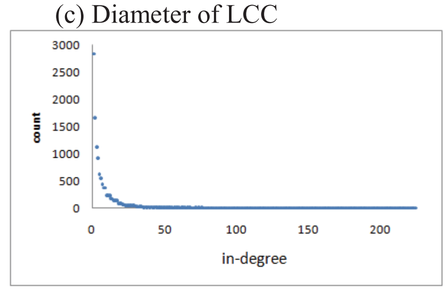
(c) Diameter of LCC



(d) Number of components



(e) Out-degree



(f) In-degree

Figure 3: Various Network Characteristics of the Mentoring Network over time

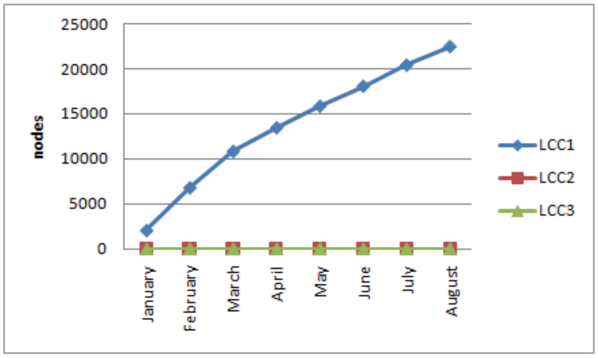


Figure 2: Components of the Mentoring Netwok over time

and transactions with the other. In restricted exchange, there is a high degree of accountability on the part of both parties. Each knows what he or she is getting from the other and can call the other to account if the relationship is not satisfactory. Second, they tend to involve quid pro quo relationships between the parties that become very specialized.

Generalized exchange, on the other hand, is organized around a community whose members are linked "in an integrated transaction in which reciprocations are indirect, not mutual" (Ekeh 1974). Exchange occurs among members of a community rather than between two individuals. Ekeh notes that this might occur in a chain of exchange, where A gives to B who gives to C who gives to D, etc. It may also occur when a group joins together to give an individual value that no single member could, that when A, B, C, and D jointly give to E (a bridal shower where a group of friends give gifts to the bride and convey community approval on her marriage is one example of this). Finally, generalized exchange may occur when individuals "successively give to the group as a unit and then gain back as part of the group from each of the unit members" (Ekeh 1974). Each of these patterns represents exchange across a more complex network.

Considering mentoring, both types of exchange seem possible. Restricted exchange would occur when friends or regular partners mentor one another. Chat sites for EQ2, exhibit numerous stories about mentoring that reflect this. Friends mentor friends to help them advance, and in return receive thanks and the satisfaction of helping those close to them. Generalized exchanges of at least two types seem likely to occur. First, some members may seek to build the community in the game by mentoring others, helping them "learn the ropes" and advance. Second, multiple mentors may help a single individual to gain by helping them, which represents the final type of generalized exchange discussed by Ekeh.

If restricted exchange holds, then the primary generative mechanism behind the network will be reciprocation of ties, once a single tie is formed. This will tend to generate particular triadic structures such 3, 7, 11, 12, 13, 14, 15, and 16 seen in Figure 1, and these should be more common than expected by chance in the network. On the other hand, if generalized exchange holds, then chains and lengthy cycles of links might hold, as well as it should favor triads 4, 5, 6, and 10, which should be more common than expected by chance. Triad 8 is likely to occur when both types of exchange occur (Ekeh 1974).

Data Description and Observations

Although we have data available from multiple servers, in this paper we report the results of experiments from only one of the servers. However we note that the results are generalizable to other servers as well (similar results were obtained on those servers). The network data is available at the granularity level of seconds. We analyzed the data at various levels of temporal granularity and observed that the network behaves in a similar manner at various though not all levels of granularity. Figure 2 gives the visualization of the mentoring network at hourly, daily, weekly and monthly levels of granularity.

Figure 3 summarizes many commonly used graph characteristics. Figure 3(a) through 3(d) illustrates that the number of nodes, number of edges, number of components and the diameter of the mentoring network increases over time. Power law distributions of both in-degree and out-degree are observed here as in many real world networks, along with a long tail. Figure 4 gives the size of the Largest Connected Component (LCC1), the second and the third

Triads	January	February	March	April	May	June	July	August
1 - 003	0	0	0	0	0	0	0	0
2 - 012	-0.01	-0.01	-0.01	-0.01	-0.01	-0.02	-0.02	-0.02
3 - 102	11.34	32	42.08	51.73	63.4	76.46	87.31	98.78
4 - 021D	2.33	5.06	5.1	4.96	5.08	5.16	5.21	5.26
5 - 021U	0.26	0.4	0.64	0.77	0.82	0.84	0.86	0.87
6 - 021C	-0.19	0.03	0.37	0.53	0.66	0.73	0.79	0.83
7 - 111D	6.6	45.61	65.26	89.33	114.5	136.38	159.48	177.07
8 - 111U	15.88	88.28	146.56	183.86	228.72	283.53	332.94	389.6
9 - 030T	139.12	320.61	371.7	375.47	392.8	406.99	419.73	431.53
10 - 030C	-1	1.15	2.24	2.99	3.28	4.52	5.44	5.74
11 - 201	1523.24	2989.24	4789.53	9944.71	12660.17	18275.82	24179.46	24567.75
12 - 120D	35056.47	101667.2	102729.2	131859.9	164847.4	193823.6	237207.36	270255.2
13 - 120U	7620.19	113628.17	159150.9	205650.6	277531.8	335976.6	383489.18	461066.5
14 - 120C	761.12	5979.48	6918.65	10265.54	13419.84	18388.64	22580.76	25733.42
15 - 210	-1	6234638.77	23645474	32352174	47536672	72387840	99573627.01	1.31E+08
16 - 300	14910406189	2.33986E+11	2.21E+11	3.03E+11	3.61E+11	6.45E+11	1.06579E+12	1.46E+12

Table 2: Triadic census of the mentoring network over time

largest connected components (LCC2, LCC3) over time. From Figure 3(a) and Figure 4, it is apparent that the overwhelming majority of the nodes belong to the largest connected component.

It should be noted that Pearson's Correlation cannot be used to study how much overlap there is between two successive iterations in the network. This is because if the graph is sparse and thus most entries are zero, would create a very large—and misleading—correlation value. Instead, we use the Adjacency Correlation γ_j as defined by Clauset and Eagle (2007):

$$\gamma_j = \frac{\sum_{i \in N(j)} A_{i,j}^{(x)} A_{i,j}^{(y)}}{\sqrt{\left(\sum_{i \in N(j)} A_{i,j}^{(x)}\right)\left(\sum_{i \in N(j)} A_{i,j}^{(y)}\right)}} \qquad (1)$$

In (1), *A(x)* and *A(y)* are the adjacency matrices of the graph at Time *x* and at Time *y*. *N(j)* is the union of row elements which are non-zero in at least one of the two matrices, γ_j is the correlation for the row for the two graphs. The adjacency correlation for the network is defined as the average of the adjacency correlation for all the rows in the adjacency matrix. The results for adjacency correlation for the mentoring network for eight months are given in Table 1. It is interesting to note that the adjacency correlation between a month and the next month is often close to 0.12 and drops thereafter. This demonstrates that while there *is* overlap between the networks, the overlap in successive months is not very large, implying that between any two time slices only a certain subset of the network is active (i.e., participants in the growth of the network). We refer to this subgraph as the *Active Graph.*

Given the 16 types of possible triads as described above, the following quantity computed via Pajek (Nooy2005) is a standard measure of determining the relative importance of each type of triad in a network:

$$\tau = \frac{n_i - e_i}{e_i} \qquad (2)$$

In (2) n_i is the number of triads and e_i is the number of expected triads in a random network.

Table 2 gives the value of τ for each of the 16 types of triads. The results show that the types of triads that were most common were consistent with specialized exchange rather than generalized exchange (as defined in the previous section on exchange). These include Triads 11, 12, 13, 14, and 16.

Graph "Laws" in the Mentoring Network

Akoglu et al., (2009) observed that a number of laws or observed patterns are found in a large number of real world networks. Based on their observations, they develop a set of 11 laws and an RTG generator for realistic graphs. In the mentoring dataset we observe that several of these laws do not hold:

Small and shrinking diameter: the (effective) diameter of the graph should be small with a possible spike at the 'gelling point'. It should also shrink over time (Leskovec et al. 2007). However, our analysis shows that the diameter of the mentoring network increases over time but not in a manner predicted by scale-free networks (Albert et al 2002).

Constant size secondary and tertiary connected components: Even though the 'largest connected component' continues to grow, the secondary and tertiary connected components tend to remain constant in size with small oscillations. In our data set, the majority of the nodes belong to LCC1 (Figure 4) even though there is more than one component. This contrasts with the preferential attachment model (Albert et al 1999).

Bursty/self-similar edge/weight additions: Edge (weight) additions to the graph over time should be self-similar and bursty rather than uniform with possible spikes. The last law is only partially violated as self-similar behavior is indeed observed at the monthly as well as the weekly level. The growth of the network is different on different days of the week because of differences in playing activity for different days (i.e., players tend to play more on weekdays). The same effect is observed on holidays.

GTPA – Graph Generative Model

Based on the observations described in the previous section we propose the following criteria that a generative model for mentor networks should satisfy:

1) The diameter of the network increases over time.
2) The number of components increases over time.
3) Bursty behavior is observed at certain levels in the network while periodic behavior at others levels of granularity.
4) The size of the active sub-graph remains more or less the same.
5) The overlap between the graphs between successive iterations is small.
6) Generate sub-structures that favor specialized exchange.

We describe this model by modifying the preferential attachment model in the following way:

(i) Consider a set of initially connected nodes n_0.

(ii) Consider another set of n_1 nodes ($|n_1|>2$) which have to be added to the network. We add these nodes one by one. When adding a new node we randomly select them and connect to one another. This ensures that there is more than one component.

(iii) From the second iteration onwards randomly select a set of n_s nodes from the graph from the previous iteration. These nodes and the edges between them form a new graph G_N. Connect all the new incoming nodes to one another according to the scheme described in (ii) and connect them to n_s according to (iv).

(iv) Temporal Preferential attachment: When choosing the nodes to which a new node connects, assume that the probability that an edge will be created from new node *j* to an existing node *i* is given as follows:

Mon	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug
Jan								
Feb	0.12							
Mar	0.09	0.12						
Apr	0.06	0.07	0.12					
May	0.05	0.08	0.09	0.13				
Jun	0.05	0.06	0.07	0.08	0.13			
Jul	0.04	0.05	0.05	0.06	0.08	0.14		
Aug	0.05	0.05	0.06	0.06	0.07	0.1	0.13	

Table 1: Adjacency Correlation for the Mentoring Network over the course of 8 months in 2006

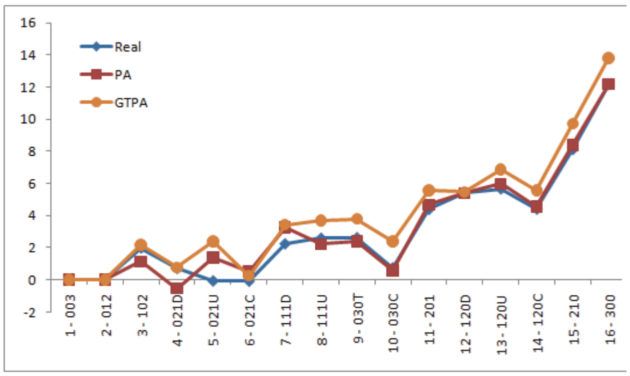


Figure 6: Triadic Census of the various Networks

$$p = \frac{k_i}{\sum_j k_j} \cdot \left(1 - \frac{t(i)}{\max(t(j))}\right) \quad (3)$$

In (3), k_i is the connectivity of the node and $\Sigma_j k_j$ is the total number of nodes in the network, *t(i)* is the age of the node and *max(t(j))* is the maximum age of any node in the network and thus gives the age the network.

In step (iv) the choice of having a new node more likely to connect to an already present node which is younger as compared to an older node seems to be counter-intuitive at first since one would expect people would prefer to be mentored by people who are more established. However we note that the number of player that a player knows is limited and it is usually in a small window of opportunity that a mentor mentors another player.

Properties

We assume that the graph and its subgraphs being considered are connected.

Lemma: The diameter of the network generated by GTPA will either remain constant or increase over time.

Proof: Suppose the diameter of the network G_0 initially is d_0 then the diameters of a subgraph (*Active Graph*) G_{0S} of G_0, is given by $d_{0S} \leq d_0$. At the end of the first iteration the diameter of the active graph and its union with the graph consisting of the new nodes is given by:

$d_{0U} = d_0 \pm \tau$, where $\tau << d_0$

This is so because G_1 is generated by the same mechanism that generated G_0 and has (roughly) the same number of nodes and edges. Here τ is the uncertainty in the diameter. The network at the end of the iteration is given by the G1 $= G_{0U} + G_{0L}$. The diameter of this graph is given by:

$d_1 = d_{0U} \pm \tau + r$

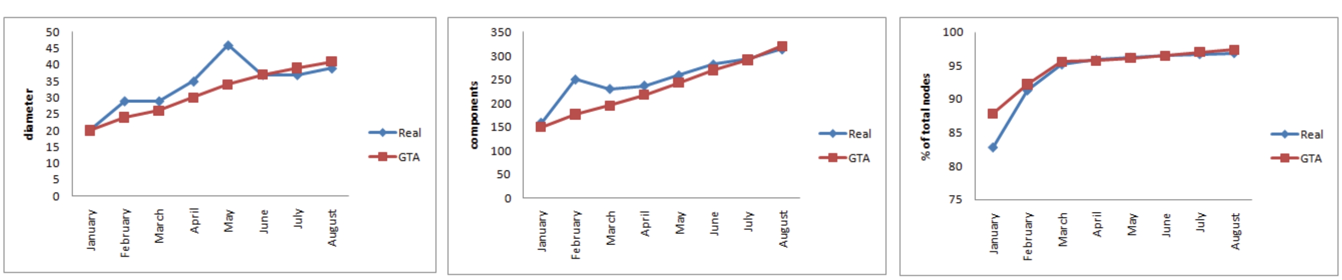
The maximum value of r is when d_{0L} is equal to (n–1) nodes and the minimum value is obtained when r is zero i.e., $d_1 = d_{0U} \pm \tau$. Thus when the value of r is minimum, the diameter remains constant while the diameter increases in all the other cases.

Periodicity: Periodicity in the model can be introduced by adding new nodes and edges to the graph based on a regular intervals such that the net effect of such an addition of a constant addition.

Experiments

The main question that we want to address here is to see if the proposed model can generate the desirable features of the mentoring network. Our model has three free parameters: τ, N_s and β. We used the grid search method (Hsu et al., 2003) to determine the most suitable set of values for these parameters. The main idea behind grid search is that given a parameter space it ties a whole range of values in geometric steps. If the model fit improves then the search moves to the next value, if not then it reduces the step size until the step size is smaller than a pre-specified threshold.

Although we have only given the results at the monthly level of analysis we ran the experiment for the monthly, weekly and the daily levels as well. The best results obtained through grid search are given in Figure 5, which are plotted alongside the observed characteristics of the mentoring dataset. Figures 5(a) and 5(b) show that the diameter and the number of components increases over time. Figure 5(c) shows that the size of the largest connected component for both the real network and the generated network. One noticeable difference between the two is that the diameter and the number of components



(a) diameter (b) number of components (c) size of LCC1 relative to the graph

Figure 5: Network characteristics of the real and the simulated network

from the generative model are monotonically increasing while in the observed network these quantities increases but with some oscillations.

Figure 6 gives the values of log of τ for all the triads in the mentoring network, the scale-free network and the GTPA network. It should be noted that the values log τ are very close to one another indicating that the triadic substructures have been recreated at the global level and thus similar types of exchanges are going on in the observed and the generated network.

Conclusion

In this paper we analyzed a special type of network formed by mentor-apprentice relationships. Many of the characteristics that are observed in this network are not observed in many other real world networks e.g., the diameter and the number of components of the network increase over time. We also explored the relationship of types of exchanges to mentoring networks (i.e., mentoring networks are characterized primarily, but not exclusively by restricted exchange). Thus, because any of the well-known graph generator models cannot be applied to this data, we presented a new model GTPA for generating networks which have characteristics similar to the mentoring networks.

Our paper also demonstrates how mentoring exchange emerges in MMORPG environment, and provides some insight into how mentoring mechanisms might emerge when introduced in other virtual worlds. For example, our finding that specialized exchange occurs more often contrasts with our intuitions that a virtual community would demonstrate complete egalitarian or equitable behavior. Instead one observes preferential attachment, reciprocity and stronger ties between specific dyads. Admittedly, these network structures may be in part a function of the specific features of this game. However, several of the known motivations for mentoring discussed earlier – playing with friends, replaying levels – appear in line with our findings. This suggests that mentoring may be difficult to coordinate among multiple people over time. We might expect to find more generalized exchange patterns in more simplistic help-giving interactions online such as discussion groups or message boards. Still, our results are quite relevant for other game developers and virtual world creators where players are both permitted and encouraged to interact and collaborate. Our results point to the value of treating mentoring as an exchange relationship that is interdependent with ones goals, and the affordances of the network at the time. Therefore, designers hoping to use mentoring to create a more communal environment will likely need to alter the incentive scheme to support this behavior.

Acknowledgement

The research reported herein was supported by the National Science Foundation via award number IIS-0729505, and the Army Research Institute via award number W91WAW-08-C-0106. The data used for this research was provided by the SONY Corporation. We gratefully acknowledge all our sponsors. The findings presented do not in any way represent, either directly or through implication, the policies of these organizations.

References

L. Akoglu, C. Faloutsos. (2009) RTG: A Recursive Realistic Graph Generator using Random Typing. ECML PKDD, Bled, Slovenia.

L. Akoglu, M. McGlohon, C. Faloutsos. RTM : Laws and a Recursive Generator for Weighted Time-Evolving Graphs. ICDM Pisa. 2008.

R. Albert; A.-L. Barabási (2002). *Statistical mechanics of complex networks*. Reviews of Modern Physics 74: 47–97.

R. Albert, H. Jeong, and A.-L. Barabasi. Diameter of the World Wide Web. Nature, 401:130–131, 1999.

Chao, G.T., Walz, P.M. & Gardner, P.D. (1992). Formal and informal mentorships: A comparison of mentor function and contrast with nonmentored counterparts. *Personnel Psychology*, 45, pp. 619-636.

A. Clauset and N. Eagle (2007), "Persistence and periodicity in a dynamic proximity network", DIMACS.

Monge, P. R., & Contractor, N. S. (2003). Theories of communication networks. New York: Oxford Uni. Press.

P. P. Ekeh (1974). Social exchange theory: The two traditions. Cambridge, MA: Harvard University Press.

C.W. Hsu, C.C. Chang, C.J. Lin. A practical guide to support vector classification. Tech. report, Dept of Comp. Science, National Taiwan University. July, 2003.

Hunt, D. M., & Michael, C. (1983). Mentorship: A Career Training and Development Tool. *The Academy of Management Review, 8*(3), 475-485.

Kram, K. E. (1983). Phases of the Mentor Relationship. *The Academy of Management Journal, 26*(4), 608-625.

J. Leskovec, J. M. Kleinberg, and C. Faloutsos. Graph evolution: Densification and shrinking diameters. ACM TKDD, 1(1):2, 2007.

M. McGlohon, L. Akoglu, C. Faloutsos. Weighted Graphs and Disconnected Components: Patterns and a Generator. *ACM SIGKDD*, Las Vegas, NV, Aug. 2008.

W. de Nooy, A. Mrvar, V. Batagelj: Exploratory Social Network Analysiswith Pajek, *Structural Analysis in the Social Sciences* 27, Cambridge University Press, 2005.

Wang, J., Huffaker, D. A., Treem, J. W., Fullerton, L., Ahmad, M. A., Poole, M. S., & Contractor, N. (2009, May). Focusing on the prize: Characteristics of experts in virtual worlds. ICA Convention, Chicago, IL

S. Wasserman, K. Faust, (1994). Social Networks Analysis. Cambridge University Press.

Williams, D., Ducheneaut, N., Xiong, L., Zhange, Y., Yee, N., & Nickell, E. (2006). From tree house to barracks: The social life of guilds in world of warcraft. Games and Culture, 1, 338-363.

Young, A. M., & Perrewé, P. L. (2000). The exchange relationships between mentors and proteges: The development of a framework. *Human Resource Management Review, 1-0*(2), 177-209.