

# Predicting Guild Membership in Massively Multiplayer Online Games

Hamidreza Alvari<sup>1</sup>, Kiran Lakkara<sup>2</sup>, Gita Sukthankar<sup>1</sup>, and Jon Whetzel<sup>2</sup>

<sup>1</sup> University of Central Florida, Orlando, Florida, USA

<sup>2</sup> Sandia National Labs, Albuquerque, New Mexico, USA

halvari@eecs.ucf.edu, {klakkar,jhwhetz}@sandia.gov, gitars@eecs.ucf.edu

**Abstract.** Massively multiplayer online games (MMOGs) offer a unique laboratory for examining large-scale patterns of human behavior. In particular, the study of guilds in MMOGs has yielded insights about the forces driving the formation of human groups. In this paper, we present a computational model for predicting guild membership in MMOGs and evaluate the relative contribution of 1) social ties, 2) attribute homophily, and 3) existing guild membership toward the accuracy of the predictive model. Our results indicate that existing guild membership is the best predictor of future membership; moreover knowing the identity of a few influential members, as measured by network centrality, is a more powerful predictor than a larger number of less influential members. Based on these results, we propose that community detection algorithms for virtual worlds should exploit publicly available knowledge of guild membership from sources such as profiles, bulletin boards, and chat groups.

**Keywords:** group formation, MMOGs, community detection, homophily.

## 1 Introduction

Guilds in massively multiplayer online games (MMOGs) have been shown to parallel real-world social structures such as work teams and friendship networks [1, 2]. Previous work has leveraged data from MMOGs and virtual worlds to conduct large-scale studies of group formation [3–5]. In the real-world, group membership can be a gateway to increased social capital. However, membership benefits are often more tangible in the virtual world, where guilds can confer direct social, economic, and military benefits to the players in the form of privileged communication channels, shared assets, control of physical territories, and tacit mutual defense agreements [6]. Note that in many MMOGs guild membership is an exclusive relationship, in which players can only belong to a single guild. Since the players are forced to choose, an examination of guild membership in these games can be highly revealing of the players’ internal assessment of the relative advantages of different social situations.

In this paper, we analyze the composition of guilds in Game X, a browser-based exploration MMOG [7]. Game X features player-led guilds who vie for physical and economic control of the fictional game world. Although nation-level conflict exists in Game X, unlike games such as Everquest (EQ) and World

of Warcraft (WoW), guilds are not primarily an enabler for player vs. player raiding activities. Our aim is to develop a computational model of the processes driving group formation and evolution within Game X and their interaction with in-game conflicts.

In this paper, we present results from the first part of our research agenda—the creation of an agent-based simulation for modeling group formation in Game X. The simulated players make guild membership decisions in a game-theoretic way by calculating the relative utility of joining vs. switching guilds. To evaluate our model, we compare the output of the simulation to the ground truth guild membership from a stable time period prior to the first nation-level conflict. Seeding our algorithm with a small number of known guild members provides the largest performance boost, particularly when the players have a high centrality. Since the leadership of guild is highly predictive of guild membership in Game X, we suggest that semi-supervised community detection approaches to community detection are likely to be particularly fruitful in this domain for a static analysis of network structure.

## 2 Related Work

To use virtual worlds for social science research, it is necessary to validate the “mapping” of group behavior in virtual worlds to real-world analogs [2, 8]; a topic of key importance is understanding how models of MMOG guild membership relate to group formation in the real world. A second question is whether these models generalize to guilds in a different MMOG setting. In this paper, we use Game X, a turn-based massively multiplayer online game, as our research testbed; Game X requires players to strategize how to make effective use of limited actions rather than encouraging “grinding” gameplay in which the players perform repetitive activities to gain wealth and experience in a low-risk way. Guilds in Game X provide one way of overcoming the action limitations, since multiple players can coordinate their action budgets toward the same mission.

A clear dichotomy between research approaches is whether they seek to recover the static guild membership structure [9, 10] or create a process-oriented model of group formation [3, 5]. For instance, Shah and Sukthankar showed that even transient groups from different regions in Second Life possess a group “fingerprint” that can be recognized by examining a combination of network and topic features [9]. Static models of community detection based on network structure have been employed within virtual worlds without attempting to understand the process by which these communities were created [10]. Our approach is based on a modification of a stochastic community-detection algorithm, NGGAME [11]; since NGGAME separately models the decision-making actions of individual agents rather than optimizing a measure of network partitions, it is well suited for modeling the dynamics of group formation and evolution. Similar to our work, the group formation model proposed by Ahmad et al. [5] also uses a stochastic

optimization process by which agents join and leave guilds according to a combination of network and player attributes. The aim of their work is to model changes in the number and size of guilds rather than the membership. Chen et al. [12] note that WoW guilds follow a life cycle in which guilds are formed from a small group of low level members and transition between different stages such as small, elite, unstable, newbie, and big. Only a few guilds succeed—most disband after failing to solve guild management problems and some become unstable because of the strong level disparity between the founding members and a revolving membership of lower-level members.

Game X does not have a concept of character level, but more experienced players have more wealth and higher average skills. Previous work has examined the role of avatar skills in guild composition [5]. Skills are often correlated with player behavior since to advance in a skill the player either has to spend time developing the skill or has to allocate limited character design resources towards acquiring it. This would indicate that guilds may exhibit skill homophily. However, in cases where guilds must operate like teams to accomplish missions, cultivating *skill diversity* can be advantageous. In this paper, we evaluate the use of both skill homophily and diversity for predicting guild membership, in combination with network structure and existing membership.

### 3 Method

To predict the guild membership, our agent-based simulation seeks to optimize each players' utility through a stochastic search process. Following work described in [13], we treat the process of guild formation as an iterative game performed in a multi-agent environment in which each node of the underlying network graph is a rational agent who decides to maximize its total utility. In this paper, we examine the contribution of two factors, network similarity ( $C_{ij}$ ) and skill diversity ( $D_{ij}$ ), toward predicting guild membership. Our code is freely available at: <https://github.com/hamidalvari/GuildDetection>.

Suppose that we have a graph  $G = (V, E)$ , with  $n = |V|$  vertices and  $m = |E|$  edges representing the network data. During the simulation, each agent can select an action (*join*, *switch* and *no operation*) to modify or retain its guild membership. In our method, each node of  $G$  represents a selfish agent which has a utility function denoted by  $u_i$  which is a linear function of two parts:

$$u_i(S) = \frac{1}{m} \sum_{j \in s_i} (\alpha C_{ij} + (1 - \alpha) D_{ij}), \quad (1)$$

where  $s_i$  is the label of the guild that agent  $i$  belongs to and  $\alpha \in [0, 1]$ .

An agent periodically decides to choose one of the operators according to the player's current utility. The set of all such guilds is denoted by  $[k] = 1, 2, \dots, n$ . We also define a strategy profile  $S = (s_1, s_2, \dots, s_n)$  which represents the set of all strategies of all agents, where  $s_i \subseteq [k]$  denotes the strategy of agent  $i$ . In our framework, the best response strategy of an agent  $i$  with respect to strategies  $S_{-i}$

of other agents is calculated as:  $\arg \max_{s'_i \subseteq [k]} u_i(S_{-i}, s'_i)$ . The strategy profile  $S$  forms a pure Nash equilibrium of the community formation game if all agents play their best strategies.

One can compute the similarities between each pair of vertices in  $G$  with respect to some local or global properties, regardless of whether the nodes are directly connected or not. When it is possible to embed the graph vertices in the Euclidean space, the most commonly used measures are Euclidean distance, Manhattan distance and cosine similarity, but when a graph cannot be embedded in space, adjacency relationships between vertices are used [14]. In this work we use separate similarity measures for the two halves of the utility function. For the first half, we use neighborhood similarity [11] to quantify structural similarity between users:

$$C_{ij} = \begin{cases} w_{ij}(1 - \frac{d_i d_j}{2m}) & A_{ij} = 1, w_{ij} \geq 1 \\ \frac{w_{ij}}{n} & A_{ij} = 0, w_{ij} \geq 1 \\ \frac{d_i d_j}{2m} & A_{ij} = 1, w_{ij} = 0 \\ -\frac{d_i d_j}{2m} & A_{ij} = 0, w_{ij} = 0 \end{cases} \quad (2)$$

where  $w_{ij}$  is the number of common neighbors node  $i$  and  $j$  have.  $C_{ij}$  assumes its highest value when two nodes have at least one common neighbor and are also directly connected, i.e.  $A_{ij} = 1$ .

To evaluate the predictive value of skill diversity between users, we calculate the  $L_2$  norm over the  $t$  skills using:

$$D_{ij} = \frac{1}{t} \sqrt{\sum_{d=0}^t \frac{1}{l_{max}^d} (l_i^d - l_j^d)^2} \quad (3)$$

where vector  $l_i = \{l_i^1, l_i^2, \dots, l_i^t\}$  and  $l_j = \{l_j^1, l_j^2, \dots, l_j^t\}$  are skill vectors for player  $i$  and player  $j$ , respectively. Using this measure, we can evaluate whether skill diversity or homophily (the inverse) are more predictive. Figure 1 shows our proposed framework. After calculating network similarities between each pair of agents (Equation 2) and skill diversity (Equation 3), the multi-agent game commences and the agents have the option of switching guilds. The community structure of the network emerges after agents reach the local equilibrium. One research question of interest is to examine how the proposed method benefits from including partial knowledge of the guild membership. To do so, we select  $M$  members from each guild to seed our agent-based simulation using one of the following centrality measures calculated with the JUNG package<sup>1</sup>: degree, betweenness, and eigenvector centrality.

## 4 Game X Guild Dataset

In Game X, players can communicate with each other through in-game personal messages, public forum posts and in chat rooms. To create our dataset, we

<sup>1</sup> <http://jung.sourceforge.net/>

1. Calculate network similarity between each pair of users
2. Calculate homophily between each pair of skill vectors
3. Select a subset of the initial guild structure
4. Repeat until convergence in the agents' utilities
  - (a) Iterate over agents
  - (b) Iterate over actions (join, switch, no action)
    - i. Calculate the change in agent utility resulting from the action
    - ii. If the change exceeds a threshold, execute action.
    - iii. Update guild membership

**Fig. 1.** Agent-based simulation

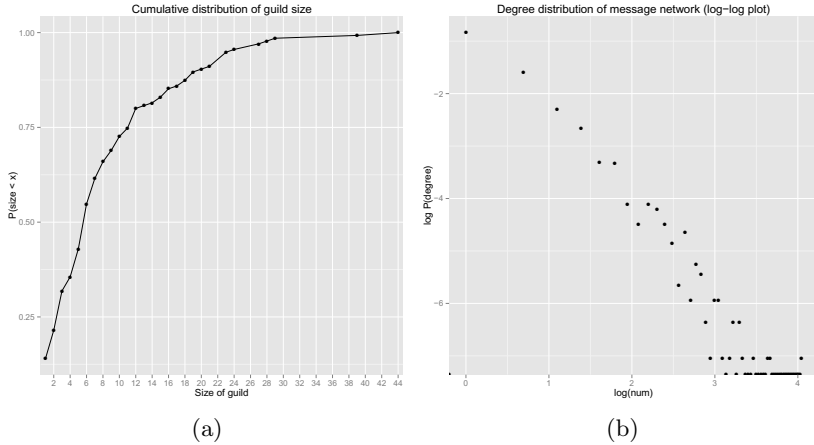
selected a relative stable period before the first nation-level conflict, during which time relatively few new users joined the game. The network structure was created by examining all personal messages sent during the time period from day 350 to 400; edges were added if the players exchanged 25 or more messages during the time period. Players who were unaffiliated with a guild were removed from the time period, but singleton guilds (with one member) were left in the dataset. The Guild X dataset contains 1150 nodes, 1936 edges, and 135 guilds; it has a clustering coefficient of 0.135. Figure 2 shows a histogram of guild sizes and node degree distribution.

Players skills are attributes of a player that can impact their ability to collect resources and successfully attack and defend. Skills can increase and decrease based on the activities of the player. Players who focus on gathering resources will have high resource gathering skills, while players who focus on combat will have high combat skills. Since players have a limited amount of turns, they may not be able to excel in both. There are 10 skills that all players can have, divided into three groups: 1) *Combat* skills, which help player attack and defend; 2) *Gathering* skills, which help players gather resources more effectively; and 3) *Movement* and *Hiding* skills, which allow a player to move and hide better. There is an additional skill that is oriented toward repair.

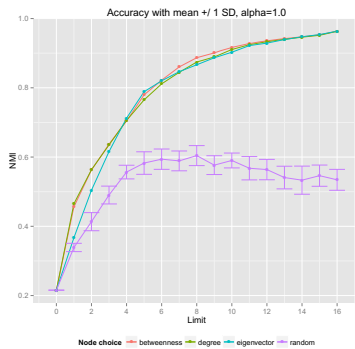
## 5 Results

In this paper, we evaluate the predictive power of our method for identifying guild membership from the data. First we must define criteria to measure how similar the discovered partition is to the partition we wish to recover (the ground truth guild membership). Here, we evaluated our results with respect to a well-defined metric, normalized mutual information (NMI). In the best condition, our proposed method for predicting guild composition scores an NMI of 0.9, assuming that we use the best utility measure and initialize the model with the ten players with the highest centrality measures in each guild. If we initialize the model with 4 players per guild, the NMI is 0.7.

To understand effect of the initialization options, we compare the performance of selecting the seed group using different centrality measures (degree,



**Fig. 2.** (a) Histogram of guild sizes (b) Degree distribution for the dataset



**Fig. 3.** Effects of exploiting different node choice options (network-based utility only). There is a clear separation between random (pink) and centrality-based (other) selection criteria on asymptotic NMI performance.

betweenness, eigenvector) vs. a random benchmark. There is a clear performance distinction between random selection of members, which reaches a maximum NMI of 0.6 even with a large number of seed members, and centrality-based which can achieve a near perfect NMI.

To tease apart the relative contribution of the network-based and skill-based utility measures, we vary the parameter  $\alpha$ . When  $\alpha$  is 1, only the network-based utility is used and when  $\alpha$  is 0, the utility measure relies entirely on the player skill attributes. If we look at the unsupervised case where we provide no seed information to the algorithm,  $\alpha=1$  (network only) is the clear winner. Skill diversity (higher  $L_2$  distance) slightly outperforms skill homophily ( $1/L_2$ ) in the unsupervised case. Interestingly, when a seed set is provided, the best performing

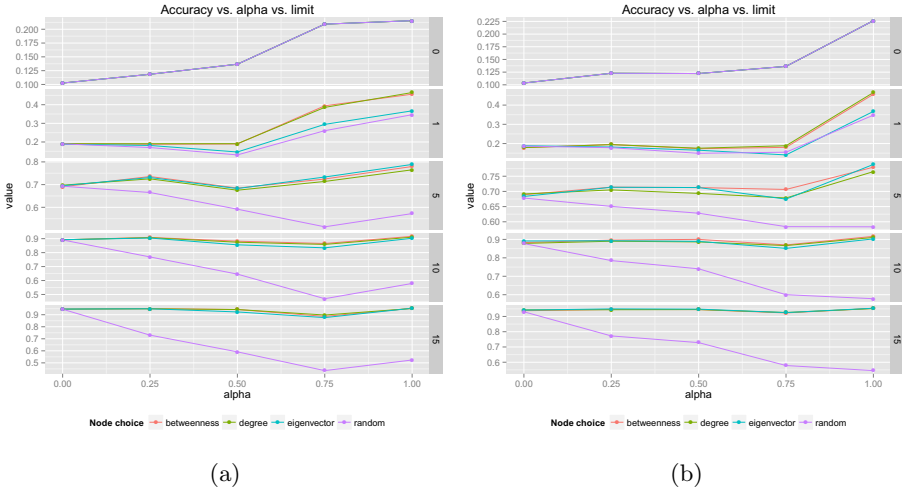


Fig. 4. Accuracy vs.  $\alpha$  vs. different subset sizes. (a) Skill diversity (b) Skill homophily.

$\alpha$  value is dependent on the number of seed members provided. With a high number of seed members, the algorithm is relatively insensitive to the  $\alpha$  value and both skill diversity and homophily are useful; Figure 4 shows the relative performance of the algorithm with different seed set sizes.

## 6 Conclusion and Future Work

This paper presents a predictive model of guild membership in the Game X MMOG that incorporates social ties, skill attributes, and existing guild membership. Our results show that differences in gameplay between Game X and more raiding-centric MMOGs create an environment in which skill attributes are a less important consideration than network structure. Knowledge of existing guild members has the highest predictive power, particularly if the members have a high centrality measure. Currently, we are exploring other methods for selecting influential players as the seed, including the use of “privileged” guild members and members with high average skill attributes. One strength of our approach is that it is relatively simple to extend the utility measure to account for economic and military factors that affect player decisions. In future work, we plan to replicate this study using economic utility measures, combined with the trade network.

**Acknowledgments.** Research at University of Central Florida was supported by NSF award IIS-0845159. Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy’s National Nuclear Security Administration under contract DE-AC04-94AL85000.

## References

1. Huang, Y., Zhu, M., Wang, J., Pathak, N., Shen, C., Keegan, B., Williams, D., Contractor, N.: The formation of task-oriented groups: Exploring combat activities in online games. In: *IEEE International Conference on Social Computing* (2009)
2. Williams, D.: The mapping principle, and a research framework for virtual worlds. *Communication Theory* 20(4), 451–470 (2010)
3. Johnson, N., Xu, C., Zhao, Z., Ducheneaut, N., Yee, N., Tita, G., Hui, P.: Human group formation in online guilds and offline gangs driven by a common team dynamic. *Physical Review E* (2009)
4. Thureau, C., Bauckhage, C.: Analyzing the evolution of social groups in World of Warcraft. In: *IEEE International Conference on Computational Intelligence in Games*, pp. 170–177 (2010)
5. Ahmad, M.A., Borbora, Z., Shen, C., Srivastava, J., Williams, D.: Guild play in mMOGs: Rethinking common group dynamics models. In: Datta, A., Shulman, S., Zheng, B., Lin, S.-D., Sun, A., Lim, E.-P. (eds.) *SocInfo 2011*. LNCS, vol. 6984, pp. 145–152. Springer, Heidelberg (2011)
6. Williams, D., Ducheneaut, N., Xiong, L., Zhang, Y., Yee, N., Nickell, E.: From tree house to barracks: The social life of guilds in World of Warcraft. *Games and Culture* 1(4), 338–361 (2006)
7. Lakkaraju, K., Whetzel, J.: Group roles in massively multiplayer online games. In: *Proceedings of the Workshop on Collaborative Online Organizations at the 14th International Conference on Autonomous Agents and Multi-Agent Systems* (2013)
8. Castronova, E., Williams, D., Shen, C., Ratan, R., Xiong, L., Huang, Y., Keegan, B.: As real as real? macroeconomic behavior in a large-scale virtual world. *New Media and Society* 11(5), 685–707 (2009)
9. Shah, F., Sukthankar, G.: Using network structure to identify groups in virtual worlds. In: *Proceedings of the International AAAI Conference on Weblogs and Social Media*, Barcelona, Spain, pp. 614–617 (July 2011)
10. Pang, S., Chen, C.: Community analysis of social networks in MMOG. *Communications, Network and System Sciences* 3, 133–139 (2010)
11. Alvari, H., Hashemi, S., Hamzeh, A.: Detecting overlapping communities in social networks by game theory and structural equivalence concept. In: Deng, H., Miao, D., Lei, J., Wang, F.L. (eds.) *AICI 2011, Part II*. LNCS, vol. 7003, pp. 620–630. Springer, Heidelberg (2011)
12. Chen, C., Sun, C., Hsieh, J.: Player guild dynamics and evolution in massively multiplayer online games. *Cyber Psychology and Behavior* 11(3) (2008)
13. Alvari, H., Hashemi, S., Hamzeh, A.: Discovering overlapping communities in social networks: a novel game-theoretic approach. *AI Communications* 36(2), 161–177 (2013)
14. Wasserman, S.: *Social Network Analysis: Methods and Applications*. Cambridge University Press (1994)