Analyzing Human Behavior from Multiplayer Online Game Logs
- A Knowledge Discovery Approach -

Kyong Jin Shim, Nishith Pathak, Muhammad Aurangzeb Ahmad, Colin DeLong,
Zoheb Borbora, Amogh Mahapatra, Jaideep Srivastava

{kjshim, npathak, mahmad, delong, zborbor, amahapa, srivasta}@cs.umn.edu
Computer Science & Engineering, University of Minnesota

1. Introduction

Observation and analysis of a phenomenon at unprecedented levels of granularity not only furthers our understanding of it, but also transforms the way it is studied. For instance, invention of gene-sequencing and computational analysis transformed the life sciences, creating fields of inquiry such as genomics, proteomics, etc., and the Hubble Space Telescope has furthered the ability of humanity to look much farther beyond what we could otherwise.

With the mass adoption of the Internet in our daily lives, and the ability to capture high resolution data on its use, we are at the threshold of a fundamental shift not only in our understanding of the social and behavioral sciences, but also the ways in which we study them. Massively Multiplayer Online Games (MMOGs)[1] have become increasingly popular and have communities comprising tens of millions. They serve as unprecedented tools to theorize and empirically model the social and behavioral dynamics of individuals, groups, and networks within large communities.

MMOGs are online spaces, providing users with comprehensive virtual universes, each with its own unique context and mechanics. They can range from the fantastical world of elves, dwarfs and humans, to space faring corporations, to a mirror of our physical world. MMOGs operate by having a large number of users populating them, interacting and role-playing within its context via the mechanics and tools provided. As MMOGs become increasingly popular, researchers are realizing there potential as a means to fully observe an entire isolated universe. Each action in a MMOG can be logged, and the level of granularity and completeness with which information is collected is unmatched by any real life experimental setup. The opportunity this offers for social scientists to test their theories empirically is unparalleled.

We use knowledge discovery as a core mechanism to analyze MMOG logs and build models of social interactions in games, and understand how users' relationships are affected by the variety of environmental factors present in each world. MMOGs provide mechanisms to foster social activities among users, allowing them to form groups, guilds, corporations, etc. and tackle collaboration-oriented tasks, e.g., raiding a dungeon for gold. Data collected from these mechanisms provides an excellent means of studying human social behavior with respect to the complete context of the environment. Such research is not feasible for real life activities as it is impossible to track and record complete information on a large population.

Since MMOGs usually have a thin-client architecture, practically all player actions are captured in the click-stream logged at the server. This data contains a comprehensive record of every player's in-game activities, accomplishments, interactions, economic status, etc. A brief record of the player's real life profile is also stored. It is common for popular virtual worlds to have hundreds of thousands of players generating copious amounts of data based on the many different activities they are participating in at any given time. One key research challenge is developing analysis methods that can analyze relationships while scaling to terabytes of data. The data also has a temporal component which is often an integral part

---

[1] Massively Multiplayer Online Games (MMOGs) and Virtual Words (VWs) are similar in that they both have complex, persistent environment, in which large number of players interact with each other. They differ in that activities in VWs are directed by the players themselves, while in MMOGs their actions are directed by the game towards achieving specific goals.

of the analysis and introduces further relationships that must be accounted for. Thus, while providing an exciting new tool for the social sciences, the virtual worlds also present a set of difficult and novel computational challenges. Figure 1 provides an overview of the overall approach to our project.
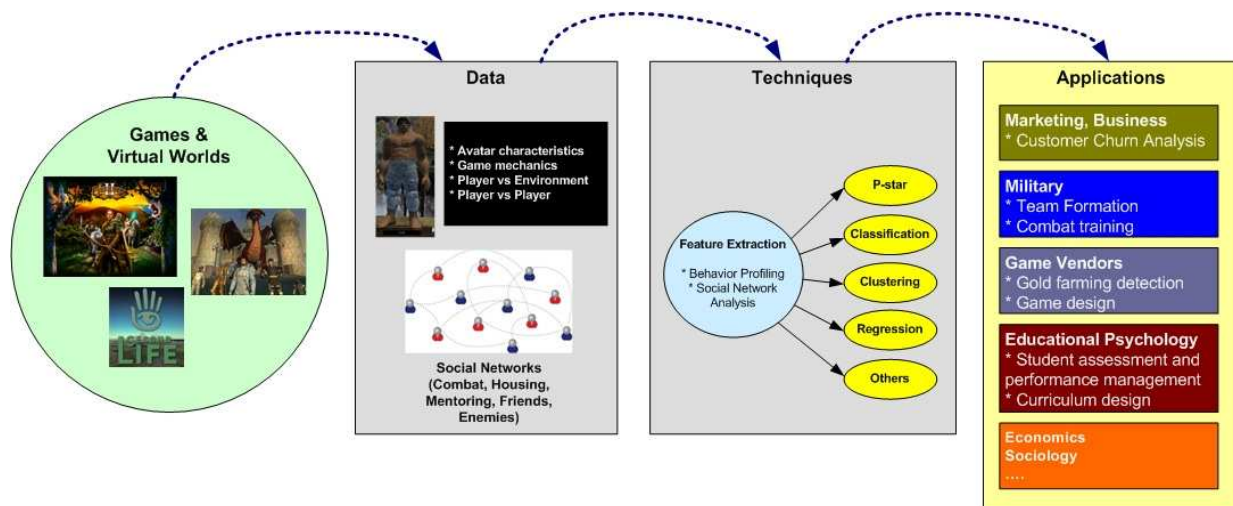

Figure 1 – Overview of Behavior Analysis from MMOG Data

2. Game Data and Social Networks

Our present research has focused on data from Sony's popular MMOG Everquest II, which primarily consists of three types, namely avatar characteristics, player demographics and in game behavior.

Avatar characteristics provide data on players' virtual avatars in the game. This includes the avatars' gender, race, faction, health and damage statistics, etc. Player demographics include real life age, gender, geographic location, etc. Survey results from a small sample of players are also available, and include response on multiple facets such as feelings towards the game and other users in the game, psychological fitness and real life behavior outside the game.

Finally, the most unique are the in-game behavior logs. This data is a record of each user's activity in the game. Anytime an event of significance occurs, such as, the user kills a monster, completes a quest, forms a group with other users, engages in trade and so on, a record of the same is written to the in game behavior log. The behavior logs are time stamped and can be parsed to uncover a wealth of information.

For our research, the most important information gleaned from the logs is about user relationships. Players can socially interact with each other using a variety of methods provided by the game mechanics. They can form groups to tackle monsters, go on quests together, trade with each other, express different levels of trust, create and join guilds, mentor other players, etc. Each of these can be represented as a social network among players, which are then studied to understand how they impacted by, and affect, both the players and the environment. Following are examples of the type of relationship networks that can be constructed from player logs.

- **Combat Network** - Understanding the formation of task-oriented groups is an important first step to study the dynamics of team collaboration. Group formation is highly influenced by players' common interests on challenging tasks. Players with less combat experience are more likely to participate in group events for difficult tasks and team performance is positively correlated to group size.
- **Mentoring Network** - Mentoring in MMOGs is the phenomenon where an experienced player helps a novice player gain skills in a particular domain. We identify four different types of mentoring in MMOs, i.e. mentoring because of altruistic reasons, helping friends, duty towards guildmates and miscellaneous reasons [9]. We have also studied mentor-apprentice networks in

MMOs and some of the graph 'laws' which have been observed in many other real world networks are not observed in these networks. Consequently well known graph generators like Preferential Attachment, Forest Fire, Butterfly, RTM, etc., cannot be applied to such mentoring networks and we have proposed a new generative model for mentoring networks [10].

- **Trust Network** - Trust networks are difficult to study, since mapping trust relationships in the real world is very difficult. Online interactions, however, provide a unique opportunity to study trust relationships. In some MMOGs, e.g. EQ2, players can buy virtual houses and grant access to other players to their houses with varying levels of access. Thus trust in these games is described with respect to a commodity. Analysis of these trust networks reveals that their properties are similar to trust networks in other domains.

Interestingly, it is observed that networks which are generated by in-game processes that have a close analog in the real world, e.g. money transactions, are very similar in the game world and the real world. In contrast, role-playing networks are quite different from any real-world networks, since this is a uniquely in-game phenomenon.

3. Applications

A shown in Figure 1, data obtained from MMOGs can be used for various purposes, both scientific and commercial. Scientific purposes include data-intensive empirical studies of various social science theories/models, including organization behavior, team dynamics, competition and cooperation, mentoring and teamwork, and performance and achievement. Commercial applications include predicting player churn, identifying undesirable behavior, social influence based recommendations, etc. In the following we provide brief details of some of these.

3.1 Performance and Learning

In recent years, researchers have found games and virtual worlds to be a sound venue for studying human behaviors. In particular, in learning sciences and educational psychology communities, systematic studies of player performance and social interactions as captured in game logs have led to a better understanding of learning, collaboration, social participation, literacy, and learning trajectory at the individual level as well as at the group level. Additionally, simulation-based training in virtual environments is widely in use by the military in recent years. Traditionally, human instructors are tasked to abstract from the behaviors of individual soldiers to the behavior and goals of teams. One issue with this approach is that the interpretation is subject to the instructor and it can vary from one instructor to another. Another difficulty is that information about rapidly evolving events or information too much low level cannot be captured by human instructors. Therefore, efficiently and effectively determining what teams are doing, how they are doing it, and why they are doing it remains a challenging task. A systematic study of individual player characteristics [2,3,11], team composition and characteristics [4,8], social interactions amongst the team members [8,11], and game environments [12] can reveal a great deal about what are the recipes for success in achieving various objectives in the game.

3.2 Player Churn Prediction

The estimated worth of the MMOG market currently stands at $6 billion, with little letdown in growth expected going forward [14]. Activision-Blizzard's game 'World of Warcraft', has the largest market share of any MMOG by far at roughly 60%, and the huge revenue potential has attracted several new game makers to this market segment. With increased competition, customer acquisition and retention is of major concern to game companies. Using web-mining techniques [1], we analyzed the in-game behavior of players and, when combined with subscription-related information, we gained significant insight into the mechanics of player engagement and their likelihood to churn [1]. This knowledge can be used by game developers to improve marketing and retention strategies, as well as to tweak game mechanics in order to enhance player engagement.

As part of churn analysis we try to understand the different motivations of players. Two key indicators of

player engagement in a MMORPG are their individual involvement and achievements as well as their level of in-game socialization with other players. Churn prediction models built on a combination of these factors have greater predictive power than those that are not. MMORPGs have different business models ranging from subscription-based to free-to-play. The churn analysis and prediction models also vary based on the nature of the underlying subscription model.

3.3 Identifying Undesirable Behavior

Gold farming refers to the illicit practice of gathering and selling virtual goods in online games for real-world currency. Gold farmers are especially frowned upon by the game developers because they disrupt the economic balance of a game's virtual economy, go against the ethos of fair play, and may make certain areas of the game difficult to access for legitimate players. Consequently, game administrators actively seek to ban gold farmers from their games. The problem of gold farmer detection can be posed as a binary classification task using in-game features such as the amount of time played, sequence of play, associations with other players, and offline characteristics such as language, location, and gender [5]. It is widely believed that a large number of gold farmers are Chinese and our analysis was the first study leveraging in-game information from more than 2.1 million characters to confirm that it is indeed the case. Gold farmers can be further divided into four sub-types: "Gatherers" accumulate gold (virtual currency), "Bankers" are low-activity accounts which hold gold in reserve, "Mules" (or "Dealers") are one-time characters that act as a link in the chain to distance the customer from the operation and, finally, "Marketers" (or "Spammers"), who advertise the company's services. Different in-game and network features must be used to detect each type.

Not all people who are gold farmers are labeled as gold farmers. Thus, labeled propagation techniques must be used in order to identify other players who may potentially be gold farmers. Network analysis of social networks of gold farmers in EverQuest 2 and the social networks of drug trafficking rings obtained from a Canadian law enforcement task force called "Project Caviar" revealed that these networks are very similar to one another in terms of connectivity, assortativity, and attack tolerance [6]. Further analysis revealed that gold farmers exhibit distinctive behavioral signatures which are masked by brokering affiliates. The similarities in both organizations' network structures reflect similar effects of secrecy, resilience, and efficiency.

4. Conclusion

With the mass adoption of the Internet in our daily lives, and the ability to capture high resolution data on its use, we are at the threshold of a fundamental shift not only in our understanding of the social and behavioral sciences, but also the ways in which we study them. Massively Multiplayer Online Games (MMOGs) have become increasingly popular and have communities comprising tens of millions. They serve as unprecedented tools to theorize and empirically model the social and behavioral dynamics of individuals, groups, and networks within large communities. Knowledge discovery is a key computational approach to realize this promise, and in the process creates opportunities to push the frontiers of knowledge discovery itself. The close collaboration between social scientists and computer scientists is creating an emerging area called 'computational social science', where computation is used as an integral mechanism to do social science research. Given current trends of user behavior, this area can only increase in importance.

5. Acknowledgments

are solely the authors', and do not represent the official positions, either stated or implied, of the sponsoring agencies. More information about this project can be obtained from http://vwobservatory.com/drupal, and all publications be accessed from http://www.vwobservatory.com/fileshare.php.

6. References

[1] Jaya Kawale, Aditya Pal, Jaideep Srivastava, "Churn Prediction in MMORPGs: A Social Influence Based Approach, " Proceedings of the 2009 IEEE Social Computing (SocialCom-09). Symposium on Social Intelligence and Networking (SIN-09). Vancouver, Canada, August 29-31, 2009.

[2] Kyong Jin Shim, Muhammad Aurangzeb Ahmad, Nishith Pathak, Jaideep Srivastava, "Inferring Player Rating from Performance Data in Massively Multiplayer Online Role-Playing Games (MMORPGs)," Proceedings of the 2009 IEEE Social Computing (SocialCom-09). Workshop on Social Intelligence in Applied Gaming. Vancouver, Canada, August 29-31, 2009.

[3] Kyong Jin Shim, Richa Sharan, Jaideep Srivastava: Player Performance Prediction in Massively Multiplayer Online Role-Playing Games (MMORPGs). PAKDD (2) 2010: 71-80.

[4] Kyong Jin Shim, Jaideep Srivastava, "Team Performance Prediction in Massively Multiplayer Online Role-Playing Games (MMORPGs)," Proceedings of the 2010 IEEE Social Computing (SocialCom-10). Minneapolis, Minnesota, USA, August 20-22, 2010.

[5] Muhammad Aurangzeb Ahmad, Brain Keegan, Jaideep Srivastava, Dmitri Williams, Noshir Contractor, "Mining for Gold Farmers: Automatic Detection of Deviant Players in MMOGS," Proceedings of the 2009 IEEE Social Computing (SocialCom-09). Symposium on Social Intelligence and Networking (SIN-09). Vancouver, Canada, August 29-31, 2009.

[6] B. Keegan, M. Ahmad, J. Srivastava, D. Williams, N. Contractor (2010). Dark Gold: Statistical Properties of Clandestine Networks in Massively Multiplayer Online Games. Proceedings of IEEE, SocialComm-10.

[7] Muhammad Aurangzeb Ahmad, Marshall Scott Poole, Jaideep Srivastava, "Network Exchange in Trust Networks," IEEE Social Computing (SocialCom-10). Workshop on Social Intelligence in Applied Gaming. Minneapolis, MN, USA, August 20-22, 2010.

[8] Huang, Y, Zhu M, Wang J, Pathak N, Shen C, Keegan B, Williams D, Contractor N. 2009. The Formation of Task-Oriented Groups: Exploring Combat Activities in Online Games. Computational Science and Engineering, IEEE International Conference on. Social Computing 4:122-127.

[9] Ahmad, M., Huffaker, D., Wang, A., Treem, J., Kumar, D., Poole, M.S. & Srivastava, J. (2010). The Many Faces of Mentoring in an MMORPG. Proceedings of 2010 International IEEE Computational Science and Engineering. Invited Paper to Symposium on Social Intelligence and Networking (SIN-10). Minneapolis, MN, August 20-22.

[10] Ahmad, M.A., Huffaker, D. , Wang, L., Treem, J., Poole, M.S., Srivastava, J. (2010). GTPA: A Generative Model for Online Mentor-Apprentice Networks. Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence. Atlanta, GA, July 11–15, 2010.

[11] Kyong Jin Shim, Jaideep Srivastava. An Analysis of the Effects of Mentoring on Player Performance in Massively Multiplayer Online Role-Playing Games (MMORPGs). University of Minnesota, Department of Computer Science and Engineering, Tech Report 10-xxx, 2010.

[12] Kyong Jin Shim, Jaideep Srivastava. Behavioral Profiles of Character Types in EverQuest II. Proceedings of IEEE Conference on Computational Intelligence and Games. Copenhagen, Denmark, August 18 - 21, 2010.

[13] Kyong Jin Shim, Jaideep Srivastava. Sequence Alignment Based Analysis of Player Behavior in Massively Multiplayer Online Role-Playing Games (MMORPGs). University of Minnesota, Department of Computer Science and Engineering, Tech Report 10-xxx, 2010.

[14] Bagga, Atul. The Emergence of Games As A Service. Industry Report, ThinkEquity LLC, Sn Francisco, May 4, 2009.