SANDIA REPORT

SAND2014-0451
Unlimited Release
Printed January, 2014

Validating Agent Models Through Virtual Worlds

Kiran Lakkaraju, Jonathan H. Whetzel, Jina Lee, Asmeret Bier, Rogelio Cardona-Rivera, Jeremy Bernstein

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

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.

Sandia National Laboratories

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

Printed in the United States of America. This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865) 576-8401
Facsimile: (865) 576-5728
E-Mail: reports@adonis.osti.gov
Online ordering: http://www.osti.gov/bridge

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Rd
Springfield, VA 22161

Telephone: (800) 553-6847
Facsimile: (703) 605-6900
E-Mail: orders@ntis.fedworld.gov
Online ordering: http://www.ntis.gov/help/ordermethods.asp?loc=7-4-0#online

SAND2014-0451
Unlimited Release
Printed January, 2014

Validating Agent Models Through Virtual Worlds

Kiran S. Lakkaraju
Cognitive Systems
Sandia National Labs
P.O. Box 5800
Albuquerque, NM 87185-1188
klakkar@sandia.gov

Jonathan H. Whetzel
Interactive Simulations, Systems, and Analysis
Sandia National Labs
P.O. Box 5800
Albuquerque, NM 87185-1004
jhwhetz@sandia.gov

Jina Lee
Systems Research & Analysis
Sandia National Labs
P.O. Box 969
Livermore, CA 94551-0969
jlee3@sandia.gov

Asmeret Bier
Cognitive Modeling
Sandia National Labs
P.O. Box 5800
Albuquerque, NM 87185-1188
abier@sandia.gov

Rogelio E. Cardona-Rivera
Department of Computer Science
North Carolina State University
Campus Box 8206
Raleigh, NC 27965-8206
recardon@ncsu.edu

Jeremy R. Bernstein
Gaikai, Inc.
65 Enterprise, Ste 200
Aliso Viejo, CA 92656
utilityjerm@gmail.com

Abstract

As the US continues its vigilance against distributed, embedded threats, understanding the political and social structure of these groups becomes paramount for predicting and disrupting their attacks. Agent-based models (ABMs) serve as a powerful tool to study these groups. While the popularity of social network tools (e.g., Facebook, Twitter) has provided extensive communication data, there is a lack of fine-grained behavioral data with which to inform and validate existing ABMs. Virtual worlds, in particular massively multiplayer online games (MMOG), where large numbers of people interact within a complex environment for long periods of time provide an alternative source of data. These environments provide a rich social environment where players engage in a variety of activities observed between real-world groups: collaborating and/or competing with other groups, conducting battles for scarce resources, and trading in a market economy. Strategies employed by player groups surprisingly reflect those seen in present-day conflicts, where players use diplomacy or espionage as their means for accomplishing their goals. In this project, we propose to address the need for fine-grained behavioral data by acquiring and analyzing game data a commercial MMOG, referred to within this report as Game X.

The goals of this research were: (1) devising toolsets for analyzing virtual world data to better inform the rules that govern a social ABM and (2) exploring how virtual worlds could serve as a source of data to validate ABMs established for analogous real-world phenomena. During this research, we studied certain patterns of group behavior to compliment social modeling efforts where a significant lack of detailed examples of observed phenomena exists. This report outlines our work examining group behaviors that underly what we have termed the Expression-To-Action (E2A) problem: determining the changes in social contact that lead individuals/groups to engage in a particular behavior. Results from our work indicate that virtual worlds have the potential for serving as a proxy in allocating and populating behaviors that would be used within further agent-based modeling studies.

Acknowledgment

The authors would like to thank the following students who assisted in performing this research: Rogelio Cardona-Rivera (North Carolina State University), Jeremy Bernstein (University of New Mexico), Kurt Hollowell (University of New Mexico), Rachel Skocypec (University of Texas), Alisa Rogers (University of New Mexico), and Hannah Gillis (University of New Mexico). This work was funded as a Laboratory Directed Research & Development project at Sandia National Laboratories.

Contents

1 Introduction **13**

Overview . . . 13

Why MMOGs? . . . 15

Data gathering methodologies . . . 15

Lab Experiments . . . 16

Social Media . . . 16

Reality Mining . . . 17

MMOGs as a source of data . . . 17

2 MMOGs, Validation and Expression to Action **19**

MMOG Examples of National Security Issues . . . 19

Corrupted Blood Spell and the *World of Warcraft* . . . 19

Espionage in *EVE Online* . . . 20

Expression To Action (E2A) . . . 20

3 Description of Game X **23**

Introduction . . . 23

Overview . . . 23

Economic Activities . . . 24

Social Activities . . . 25

Combat Activities . . . 25

Potential for Large Scale Conflict . . . 26

A Player's Death . . . 26

4 Validation of Agent Based Models **27**

Introduction 27

Behavior Influence Assessment (BIA) Model 29

How BIA Models Are Typically Built 30

Requirements for Application of BIA 31

A Path Forward 31

5 Motivations to Join Conflicts **33**

Introduction 33

Large Scale Conflicts in Game X 34

Analogues to real world Large Scale Conflict 34

Theoretical Framework for LSC Analysis 35

Behavior of Interest: Dimensions of Combat Behavior as a Function of Community Networks 36

Experimental Methodology 36

Operationalizing the Hypotheses 36

Participants 37

Participants of Interest 37

Hypothesis Tests 38

Results and Discussion 39

Conclusion 40

6 Guilds Organization **41**

Roles in Game X 41

Roles in Guilds 43

Do all guilds have the same roles? 44

Discussion 48

Conclusions & Future Work 51

7 Predicting combat from public communication **53**

Related Work . . . 53

Method . . . 55

The Data Set . . . 56

Communication in Game X . . . 56

Forum based measures . . . 56

Cleaning forum data . . . 57

Vector Space Representation . . . 57

Feature Selection . . . 58

Results . . . 58

Sentiment Analysis . . . 62

Dictionary Based Sentiment Analysis . . . 62

8 Public and Private Interaction **67**

Introduction . . . 67

Why Massively Multiplalyer Online Games (MMOG)? . . . 68

Experiments . . . 68

Methods . . . 69

Results . . . 70

Conclusions and Future Work . . . 71

9 Conclusions **73**

References **74**

List of Figures

2.1 Comparing Application of the Expression to Action (E2A) Problem to Real World and Virtual World Data 21

2.2 Diagram showing virtual world data used toward informing and validating an established Agent Based Modeling framework 22

3.1 A diagram of the game world. 24

4.1 Conceptual drawing of the Behavior Influence Assessment (BIA) Model 30

5.1 Combat and messaging patterns throughout the entire Game X data set. The number of combat attacks and the number of messages sent both spike during the periods of war, consistent with real world accounts of conflict. 34

5.2 Distribution of Number of Attacks v. Number of Attackers during the first Game X war. Consistent with other accounts of participation in armed conflict, our virtual world LSC data is well-modeled by a power law. 35

5.3 Box plot of log-in percentage. Log-in percentage is determined by the number of days that players logged in over the total number of days in our data set. We filtered the data for all players whose log-in percentage was less than 10%. 38

6.1 Histogram of guild sizes 42

6.2 Boxplot of silhouette values over multiple values of k. 45

6.3 Plot of the clusters against the first two principal components for Guild 1. . . . 46

6.4 The Medoid of Medoid Method 47

6.5 Plot of all medoids of all guilds on the first two principal components for $k = 3$ 49

6.6 Plot of mean silhouette value for varying k. 50

7.1 Example data set. 54

7.2 Assigning classes to the data. 54

7.3 Feature vector for terms. 55

7.4 Illustration of window size and delay. Posts are aggregated over the space of the window, and a classification is given based on whether an attack occurred at the current time + delay. 55

7.5 Area under the Curve (AUC) as a function of window size. 63

7.6 Area under the Curve (AUC) as a function of delay. 64

7.7 Forum structure in Game X. Each forum can have multiple topics, and each topic can have multiple posts. 65

7.8 New posts per day from day 5 to day 739. Highlighted time periods indicate the two wars and the evaluation period. 66

8.1 Density distribution plot of model resampling with area under ROC curves as a metric. 71

List of Tables

5.1 Hypothesis Test Results 39

6.1 Medoids of Guild 1 44

6.2 Medoids of the MoM clustering. Cmb = Combat, Ec= Economy, Mvm=Movement 48

7.1 Overview of post/authors/and topics for each forum. Bold entries are the max values for the column. These are calculated over the entire 737 day period. . 57

7.2 Window size 1, Delay 0,1,2 59

7.3 Window size 2, Delay 0,1,2 60

7.4 Window size 3, Delay 0,1,2 61

8.1 Paired-samples t-tests for resampling results 70

8.2 T-tests for the resampling results (LM: Linear Model, BT: Boosted Tree, SVM: Support Vector Machine). 72

Chapter 1

Introduction

Overview

Agent Based Social Systems modeling is a growing field [39]. Initially based on simple models of agent behavior, models now have complex internal dynamics and interaction between agents ([55]).

An important part of modeling, though, is identifying how well the model connotes to reality, i.e., the "validation" problem. We focus specifically on "Empirical Validity" (EV), by which we mean the following (from [21]) "... empirical validation involves examining the extent to which the output traces generated by a particular model approximates reality..". Another term for EV include pattern, point and distributional validity defined in [14]. EV is just one of a set of validity measures to apply to a model (see [14] for an overview of others); however we find it the most intriguing as it speaks to the relevance of the ABM for use in analysis and decision making.

Calibration ("Calibrating is the process of tuning a model to fit detailed real data..." [14]) is a step towards empirical validation and focuses on tuning the set of parameters of a model to fit real data.

While several techniques have been developed to establish EV (see the special issue in computational economics, 2007 30), underlying these techniques is the issue of appropriate real world data. The question that become important for a practitioner is often not *how* to validate, but rather *with what* to validate data against.

This problem becomes even more difficult with the rise of complex, multi-scale models that capture micro and macro behavior. For instance, models that capture the internal dynamics of cognition that lead to external interaction and behavior changes. To properly calibrate such models, one needs data at the individual *and* group level [46]. Specifically, a good data set should:

- Contain individual characteristics: How do individuals behave?
- Provide temporally extended data: To observer long term dynamics.
- Contain data from a large and diverse set of individuals: To ascertain general patterns.

Validation techniques require data, and without it we are lead to creating simpler models that can be analytically understood. Simpler models (”intellective” models, [14]) may not capture the factors needed to accurately anticipate behavior.

The goal of this two year LDRD was to investigate the use of Massively Multiplayer Online Games (MMOGs) in the creation and validation of Agent Based Social Systems at Sandia. More generally, we also strived to understand where this dataset could be used in other Sandia domains.

Our conclusions are as follows:

First, game data can be extremely useful to Sandia. Apart from the creation and validation of ABMS, there are other domains (such as testing graph analysis algorithms) where temporally extended, large scale data can provide a useful data set.

In terms of creation and validation, game data provide a rich panoply of events and relationships that can be mined to inform/validate existing ABMS and motivate new ABMS.

Secondly, we show that game data can be useful for the creation and validation of ABMS. This is an argument in two parts, first: do games exhibit complex dynamics that reflect the complexities of the real world, second: how can we use the data to inform/validate ABMS?

We provide evidence from the scientific community that games, in general, reflect the complexities of the real world (see Chapter 2). We also show, in Chapter 6, the groups within Game X exhibit a non-simple structure, potentially making them a great source of data on organizational dynamics.

We show in Chapter 5 that real-world motivations (such as peer pressure) to join violent conflicts can help predict players joining violent conflicts within the game. This provides more evidence that we we can use game data as a proxy for real-world data on violent conflicts.

In Chapter 4 we discuss the issues and problems with using MMOG data for validation of ABMS, in particular we look at the Behavioral Influence Assessment (BIA) model.

Thirdly, the unique ability to have fine grained, extensive data about all player actions within the game allows the exploring the links between public “expressions” and private actions/relationships. We call this the “Expression to Action” problem, and games are uniquely suited to help address this problem.

In Chapter 8 we highlight work on using forum data to predict guild memberships within Game X. In Chapter 7 we show how forum data can be used to predict kinetic actions, specifically combat actions, of players.

In the rest of this chapter we provide more information on why we believe games can provide a new, unique data source.

Why MMOGs?

Initially it may come as a surprise to think of games as a valid source of data for ABMS and other models of national security interest. However, the fact that players are spending immense amounts of time within the game world executing long term complex actions highlights the potential usefulness. Just as social media can be used to understand facets of human behavior, games also highlight other facets.

In this section we will place MMOGs within the context of other data.

Data gathering methodologies

Before we begin describing MMOGs, we identify a set of criteria to help compare different data gathering methodologies (see also [33]). Each method comes with it's pro's and cons.

We will use the term "subjects" to indicate the individual we are gathering data from. We use the term "context" to denote the environment in which we gather data. For instance, the context of a lab experiment would be a lab where individuals must come and be physically present. In contrast, the context of social media is an online, virtual space.

Number of Subjects People influence each other; the behavior of a single individual may be quite different from behaviors of groups of individuals. Striking examples abound from the social psychology literature [6].

Diversity How diverse is the set of subjects? Facebook for instance, attracts people from a wide range of ages and socio-economic statuses, whereas lab experiments predominantly attract college students.

Realism How realistic is the context? We can draw from the social psychology literature to identify different types of realism: "experimental realism" – how involved were the participants in the experiment, "mundane realism" – how likely the events occurring in the experiment would be to occur in the "real world", and "psychological realism" – are the psychological processes involved the same that occur in everyday life? [62]. The Stanford prison experiment would be an example of a situation which had experimental realism, but not mundane realism. This can be important, as the engagement of an individual can determine how well their behaviors correspond to what they would do in real life.

Dynamics Does the method capture behavior over time? In studying innovation diffusion, we would like data on what individual behavior over time [33].

Level of manipulation Does the method passively observe behavior in a fixed context, or can the observer manipulate the context? For instance, we can observe Twitter traffic easily, this is a passive method of observing behavior. In contrast, with lab experiments (and to a certain extent social media [8]), we can manipulate the context of the subjects in order to ascertain causal relationships.

Privacy Does observing the behavior violate the privacy expectations of the individuals?

Domain What can be observed? Is it just behaviors within the context, or can we get information about the cognitive processes underlying these behaviors, communications, or biological processes (EEG/fMRI)? Communications is especially important in terms of social influence.

Commitment How important is behavior in this context? For instance, "Facebook friends" are friends who one wouldn't interact with in real life, but are "friends" in the facebook context. This is generally an idea of how closely the behavior in the context relates to the "real" behavior of an individual. There will be a correlation between commitment and realism.

These dimensions can be used to gain a qualitative understanding of a data set. We outline a few positive and negatives about various means of gathering data below. We do not intend this list to be complete, but we feel it captures most of the prevalent, current, ways of gathering data.

Lab Experiments

Lab experiments are often the "gold standard" in data gathering for a very good reason: experimenters have the ability to manipulate the context of the subjects. Physical and psychological stimuli can be added or subtracted for each subjects. In addition, behavior, communication and even biological processes can be easily captured.

There are several drawbacks however. The expense of having human subjects come into the lab often makes the number of subjects small. Dynamics are hard to track since long term monitoring of subjects is cost-prohibitive. Subject diversity is a major concern with experiments; many of the social psychology experiments take place in western colleges, leading to a lack of diversity among subjects [51]. Since all lab experiments must go through an Institutional Review Board approval process, the privacy concern is often negligible.

Realism and commitment will depend on the experiment and cannot be generalized.

Social Media

Social media, which we define as media sources such as Facebook, Twitter, newsgroups. are a rich source of data. There are millions of diverse subjects. Dynamics can be assessed by gathering data over a long period of time (which is relatively easy to do).

The drawbacks are primarily in realism, manipulation and domain. We can only passively observe the data coming from social media, and cannot manipulate it. In addition, the level of commitment is not understood. Consider the issue of ”radical chic”, in which users post comments on line in order to gain reputation, but do not actually mean them [38]. In this situation, an individual may not be committed to their behavior in this context, and are free to behave in ways opposite their character.

Reality Mining

Reality mining is a relatively new data gathering method. In this method, subjects are passively recorded throughout the day for a lengthy period of time, often through their cell phones [42, 30].

Reality mining has several positives: one can gather data from relatively large pools of subjects, subject diversity can be high (although often limited to subjects with a certain type of cell phone); the data is realistic and falls into the ”mundane realism” category. Dynamics can be captured through lengthy measurement periods. Commitment is also high as the behaviors are done by subjects as part of their daily lives.

The drawbacks are in level of manipulation, privacy and domain. This is a passive monitoring method, and thus we cannot manipulate the context. The domain is often very low level (GPS position), but can capture communications. An important drawback is privacy – subjects may object to passive, 24 hr monitoring.

The mundane realism of this method can also be a drawback. One of the benefits of lab experiments is placing subjects in situations out of the ordinary, in order to reveal basic characteristics of humanity. The Stanford Prison Experiment showed the influence of power on normal individuals – thus highlighting the underlying psychology of all humans; the experiment would have been much less compelling if done using actual prison guards who may have had training and environmental conditioning.

MMOGs as a source of data

Massively multiplayer online games are online games that attract millions of players to a shared, virtual world. Many varieties of online games exist, some familiar, such as *Farmville*, others less so, like Second Life, World of Warcraft, or Eve Online. While the term MMOG may encompass a variety of genres, we are interested in the large portion of games that are often labeled ”role playing games”. In these, players create an avatar that represents them in the virtual world[1].It is this genre that we specifically discuss in the following.

Some games have objectives and quests in order for the player to gain experience and skills (e.g., World of Warcraft), whereas others have an open-ended world in which all content

[1] In the following we use the term ”players” to refer to the avatars within the game.

is created by the players (e.g., Second Life). Still others are in the middle of this spectrum, providing means to gain wealth, power and experience, but allowing for open ended play within the universe (e.g., Eve Online).

MMOGs are appealing for their complex economies and social structures. Many of the games contains player created and controlled "guilds" or "corporations" which players can join. These groups regularly have conflicts and interactions in the world. In some instances, long term (approximately a year of real world time) espionage has been conducted [25]!

MMOGs have several advantages as a method of gathering data.

Number of subjects MMOGs have thousands to millions of players. Recent data indicates that more than 21 million active accounts on various MMOG games [2].

Diversity Contrary to popular belief, MMOGs have a wide array of player types. A study conducted with 30,000 players [64] indicated a mean age of 26.57 with a range of 11-68. In addition, both genders were represented.

Realism MMOGs are high in experimental realism, as players willingly spend hours playing (on average 22 hours per week [64]), however they clearly lack in mundane realism. It's still an open question as to whether they have psychological realism.

Dynamics MMOG data can be captured over years; events in game often occur faster than in the real world so one can see the rise and fall of organizations within the game.

Manipulation MMOGs can be observed passively, since they are naturally instrumented. Platforms are being built with the intent of manipulating the environment though [1].

Privacy All data is generated in a virtual world, so most privacy concerns are minimal.s

Domain MMOGs have a unique capability to observe communication and behavior of players. In-game forums and messaging data can be gathered along with behavior. This gives us one way of addressing the "radical chic" problem, by explicitly studying the correlation between communication and behavior.

Commitment MMOGs are still games, and player decisions do not affect their real lives. However, players do invest much time into their avatars and have strong emotions regarding said avatars. We believe this leads to players wanting to protect their avatars, making their level of commitment to behaviors stronger.

The main criticism against MMOG data is that player behavior in-game is not the same as real-world behavior – a question which is laid out in the "mapping principle" [61]. In Chapter 2, we describe some attempts to address this question. Some of the outcomes of our studies also address this point.

Chapter 2

MMOGs, Validation and Expression to Action

In this chapter we briefly outline the two main problem domains we are concerned with in this LDRD.

MMOG Examples of National Security Issues

Though a fairly new genre, MMOGs have their roots from the early days of Multi-User Dungeons (MUDs) made popular in the late 1970s. Advancements in computing technology have not only moved games into a more visual medium, but allow for thousands of players to simultaneously interact within these rich environments. MMOG designers have also evolved these worlds in-step with the technology, creating more socially realistic environments. Modern MMOGs now include constructs such as governing bodies and market-based economics, with players having self-ascribed goals to grow and preserve influence, either their own or select player group, amongst the entire population. These social constructs and participants willing to follow them have led to behaviors in-game mirroring those seen in populations during pandemics or engaging in espionage to undermine a particular group. We cover two events that inspired the authors in this research, the *Corrupted Blood* spell within *World of Warcraft* and the insider threat which took down a prominent player group within the MMOG *EVE Online*.

Corrupted Blood Spell and the *World of Warcraft*

World of Warcraft, or colloquially known as *WoW*, was released by Blizzard Entertainment in 2004 and currently remains one of the most popular games of the genre, recording nearly seven million paid subscribers by 2013 [34]. In 2007, Blizzard released an expansion to *WoW* where players would face against virtual opponent that fought with what the developers named the *Corrupted Blood* spell. If the virtual enemy successfully attacked a player with the spell, the player would see their character's health slowly deplete until their character died. Players under the *Corrupted Blood* spell would infect others with it if they got within a close proximity. A player would only have the spell lifted if the virtual enemy was slain or their character died, allowing the player to resurrect and try again with full health.

Developers for *WoW* did not take in account that players would retreat from the enemy if they felt overmatched. Retreating players, some of which under the *Corrupted Blood* spell, would return back to the highly-populated community player areas and begin infecting others with the spell. Epidemiologists at Tufts University, who were fellow *WoW* players, documented their qualitative observations on the event [40], noting how this episode related closely to actual pandemics. For instance, Blizzard attempted to institute player quarantines to have all infected characters die in a controlled area to prevent further spread. This act was met with resistance from the player population, causing some infected players to flee the area, or non-infected to enter the area out of their curiosity. This episode lasted for several days until Blizzard took stronger measures by shutting down their servers temporarily to remove the *Corrupted Blood* spell entirely from the game.

Espionage in *EVE Online*

EVE Online is another popular MMOG released in 2003 with a current player base of over 500,000 people worldwide [41] In the *EVE Online* world, player-formed groups are called *corporations* which compete for control over physical space within the virtual world. Players can contest physical spaces through combat, or negotiate sales of physical territory and assets to other corporations using marketplaces and internal currency of *EVE Online*. In 2009 [13], the largest corporation at the time, *Band of Brothers*, had their Treasurer sell a large percentage of the corporation's assets, significantly reducing their economic presence within the game. It was revealed the Treasurer had been a spy for rival corporation *Goonswarm*, and ordered to take such actions at the behest of *Goonswarm* leadership.

From player accounts, some *Goonswarm* members had been adept at contacting and turning other players against their previous allegiances. Over the course of several months, *Goonswarm* met with and finally convinced the *Band of Brothers* Treasurer to defect, while tearing down the corporation in the process. Several *Band of Brothers* players filed protests to *EVE Online* developers, insisting on retributions for the assets that were stolen from them via this espionage attack. *EVE Online* developers refused citing that their former Treasurer had not conducted any actions that violated their game rules. During informal conversations with other MMOG developers, we have learned that espionage has now become a popular tactic across many MMOGs for undermining player groups.

Expression To Action (E2A)

In both examples, we can qualitatively observe parallels between these events in virtual and real worlds. Yet, for better understanding and predicting when conditions will yield these kinds of social behaviors, we require quantitative data providing the state of each person,

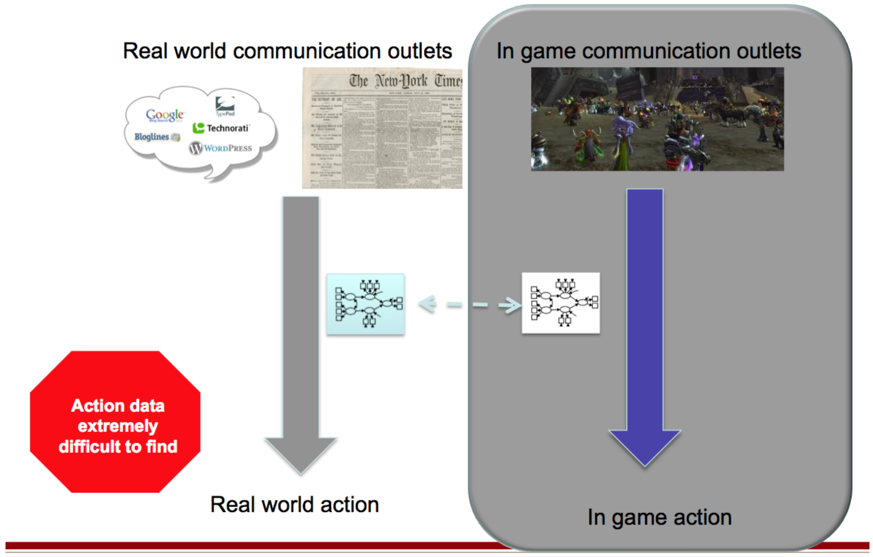


Figure 2.1. Comparing Application of the Expression to Action (E2A) Problem to Real World and Virtual World Data

communication patterns within the group of interest, and environmental factors that lead toward these events. Agent Based Modeling (ABMs) provide a mechanism where one may establish parameters that define these quantitative measures, and simulate how the agents will behave given some initial condition.

As noted in at the beginning of this chapter, using data from the real world as a means for informing or validating an ABM can fall short given its imperfect nature. In figure 2.1, we view the differences between how publicly-known information alters group behavior toward some particular action. In the real-world, print and on-line media provide accounts on the state of the world, with social media allowing for the expression of opinion and commentary on this information in real time. What we cannot passively observe are the social connections made in private between people as an effect of the publicly disseminated information. We hypothesize that knowing information on group membership becomes vital for understanding and predicting the likelihood for our events of interest. Virtual worlds allow for capturing all data encompassing the world in which people inhabit. Obtaining such a rich data set allows for the possibility at conducting empirical validation of ABMs, testing theories on how a person's social networks will shift and lead to some action by the person due to these changes. This gives rise to researchers better understanding Expression-To-Action (E2A) inherit in studying these kind of phenomena: how does the shift in communication, both public and private, lead to changes in a person's social network and ultimately their observable behavior?

Figure 2.2 provides a method illustrating how virtual world data can aid in understanding E2A for the benefit of agent-based modeling. In this illustration, we can perform data mining techniques to extract information on shifts in player(s) communication and behavior

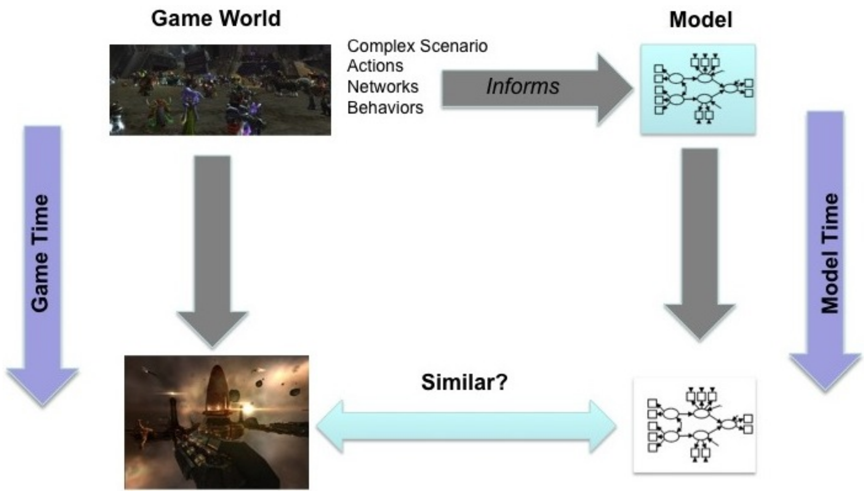


Figure 2.2. Diagram showing virtual world data used toward informing and validating an established Agent Based Modeling framework

surrounding a given phenomena. From this exercise, we can diffuse our findings into the quantitative parameters for an established ABM framework for this same phenomena. By providing information on events from the virtual world to the ABM, we can generate a simulation in an attempt to predict when our targeted phenomena would transpire in the virtual world. Simulations that correctly predicted the event in virtual world serves as an empirical validation of the underlying social theories used in governing the interactions between the agents in the ABM. Discrepancies in the results of the ABM prediction and virtual world outcomes would raise questions: does the chosen virtual world have the ability to accurately replicate our given phenomena?, or does the ABM make some false assumptions on social dynamics that requires further refinement. In either case, the virtual world data provides a tremendous asset by providing researchers a complete record of a person's behavior and communications to use in testing social theories. Our remaining chapters explore how we used a MMOG dataset for mining social behaviors, and how this could be applied to known ABM frameworks.

Chapter 3

Description of Game X

Introduction

Accomplishing our research goals for this project required finding a game development company that would allow for an external party to examine their dataset, an uncommon practice at the time of this research. Besides not only finding a willing participant, these company needed an existing MMOG whose data we could immediately access given the short duration of this project. As well, this MMOG needed to have a large enough player population and known behaviors of interest characterized in the last chapter. We acquired data from an existing MMOG entitled *Game X* to conduct our research for this project.

Overview

Game X is an open-ended free-to-play Massively Multiplayer Online Game. Players can pursue a variety of different roles and interact with other players (real and artificial) in the virtual world. In Game X, every player commands one vehicle, with a set cargo capacity as well as defensive and offensive capabilities. Players use this vehicle to explore an open, persistent game world.

In Game X, players explore a 2-D world using vehicles. They are able to mine resources and transform resources to other products. Figure 3.1 is a schematic of the Game X world.

Game X is unique in that the game does not impose an explicit goal structure and a player cannot win the game. Instead, the game encourages players to make up their own goals, role-play and to acquire wealth, fame, and power, in an environment driven by several societal factors, such as friendship, cooperation, competition, and conflict.

Players have a limited number of ”turns” per day. Nearly all actions cost some number of turns to execute. Turns are replenished per hour. The basic categories of actions players can undertake are: economic, social, and combat.

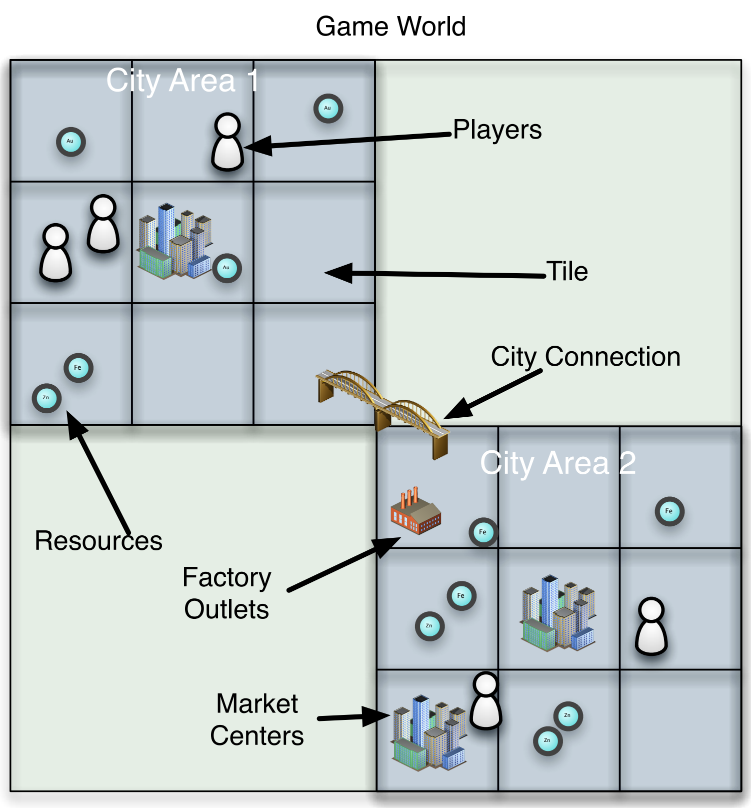


Figure 3.1. A diagram of the game world.

Economic Activities

Players can mine goods integral to improving their economic performance in game. To obtain other goods, they may engage in trade, which can be a form of barter or via the use of *marks* an in-game currency. With enough game experience and the proper amount of marks, a player can construct a factory outlet, which can manufacture sellable goods. In addition, a player can eventually earn enough resources to build and maintain a market center, which can facilitate bartering and selling with passerby players.

Players can also be strategic about the location of their factories and market centers, by exploring critical city areas that are ripe for developing profitable trade routes. By the same token, players must take care to avoid establishing business in areas that are targeted by pirates; other players can plunder factory outlets, market centers, as well as vehicles.

In addition to piracy, other "illegal" (in terms of society, not in terms of what is permissible in game) options for economic activity also exist: a player may elect to bootleg illegal goods. If caught, a player can be subject to social sanctions and be barred from bartering and (if it was a repeat offense) from being able to engage socially with other players. Other sanctions include: not being able to repair your vehicle, and not being able to approach/visit city areas.

Social Activities

Players in Game X can also socialize and associate with other players through a number of different ways. In fact, to be successful in the game world, it behooves players to forge social partnerships.

To socialize, players can post on public forums, send personal messages to other players as well as broadcast messages top specific groups of players. An important feature of communication methods in Game X is that they do not cost players any amount of available turns; communication is “free of charge.”

To associate, players can join one of three pre-defined nations; players can only affiliate with one nation at a time and doing so yields certain benefits, such as the possibility to gain access to nation-specific technology. Players may also elect to not join a nation, which allows them to be exempt from nation-centric wars. In addition, players can create or join a guild, which are entities designed to combine groups of players and allow them to operate for (possibly) common aims. Guild membership is independent of nation membership. For both nation and guild memberships, if a player accumulates sufficient in-game experience, he or she can be promoted to a senior level, which commands a higher influence in the respective nation or guild. In fact, senior level members of nations command considerable political power in the decision to go to war (see Section 3).

Finally, players can designate other players as friends or hostiles, which facilitates or hinders communication and other game activities with those players. Friend/hostile tables are completely private, meaning that no one except the labeling and labeled players has information about ties between them (i.e. it is not possible to see second degree neighbors, such as friends of friends).

Combat Activities

Players can engage in combat with other players (real and artificial), as well as with factory outlets and market centers. Players can outfit their vehicles with a variety of different weapons and defensive armors that (alongside a player’s skill) can be used to give certain advantages in battle.

Players have an array of skills they can improve based upon their successful in game battles. Higher skill values increase the probability of successful combat in the future.

Potential for Large Scale Conflict

Large Scale Conflicts (i.e. wars) are socially centric and very related to combat activities. Wars are only possible between the three pre-defined nations. Each nation can have one of the following diplomatic relations to all others: Benign, Neutral, Strained, or Hostile.

The senior members of a nation constitute the nation's governing body. Every day, each nation's governing body convenes and each of the senior members chooses a disposition with regards to diplomatic relations with the other nations. Non-senior members cannot vote, but can exert influence by lobbying senior members to vote a certain way. If enough members of a governing body select hostile diplomatic relations against another nation, a war is declared between the respective nations.

When a war has broken out, additional combat actions are available for the warring nations. In particular, war quests are available, which provide medals of valor to the players that wish to undertake and complete the quests. Any attack against the opposing nation (be it in the form of a war quest or not) results in accumulating a set number of war points. When the war ends, these war points determine the "winner" of the large-scale conflict. A war situation will (via the game's design) gravitate towards a state of peace. Each of the respective governing bodies must maintain a majority vote to continue the war effort. Over time, the amount of votes required to continue is increased by the game itself. Eventually, no amount of votes will suffice and the nations return to a state of peace.

A Player's Death

A player cannot permanently die in Game X. If an enemy destroys a player's vehicle, then the player loses a fixed amount of skill points, as well as all the cargo on his or her vehicle and in addition loses some available actions for the play session.

Chapter 4

Validation of Agent Based Models

Introduction

Model validation is a process of assessing a model and its results. Validation is used to determine how well a model represents reality and whether this representation is adequate for the intended purpose of the model. Validation can build understanding of a models capabilities and limitations, and can strengthen confidence in the model and its usefulness.

The validation approach that is widely used in the engineering field is comparison of a models results to data collected from experiments or real-world systems [48, 36, 45]. Because social systems are fundamentally complex, some researchers have suggested that models of these systems should use an expanded concept of model validation that includes a variety of confidence-building exercises [28, 22]. Implementation of this more general concept of model validation may require significant resources, but is important for building confidence in social systems models [11].

Many different methods for validating models of social systems have been proposed. A few of these methods are briefly discussed here, including face validation, structural analysis, boundary adequacy, extreme condition tests, and parameter range assessment. For more in-depth descriptions of these and other methods, see [22, 9, 15, 45].

Face validation is a subjective but valuable assessment method. Face validation involves sharing a model with experts who are asked to determine whether the model is a reasonable representation of reality, particularly with respect to its outputs [9]. The experts who participate in face validation might include subject matter experts, modelers, stakeholders, or others. Since experts are sometimes the best sources of information on social systems, face validation is commonly used for validating these models.

Experts might also be asked to evaluate the structure of a model. Science and engineering disciplines often base models on first principles, information that is not always available for social systems. Instead, expert opinion may be the best source of information about how a social system model should be designed. Structural analysis involves sharing diagrams or other representations of model structure with experts, and asking those experts to assess the adequacy of the model [22, 9]. This may include an assessment of the boundaries of the model [22], which can help to determine whether the models scope and detail are sufficient to portray the systems behavior.

Extreme condition tests and parameter range assessment can also be used to evaluate the structure of a model. Extreme condition tests assess a model by setting chosen inputs at extreme values [22]. The associated output is then assessed for reasonableness, considering expert opinion about what the systems behavior would look like under these extreme conditions. For example, a model might be simulated with food availability set to zero, with the expectation that the simulated population would not survive. Parameter range assessment can also be used to assess model structure, by simulating the model with different sets of inputs. The results of this process can increase understanding of model robustness and can be used evaluate the reasonableness of results generated by various sets of inputs.

To bolster confidence in a model, it is best to use a variety of validation methods. However, comparison of model results to real-world data is likely the most important technique for enhancing credibility of a model. Comparative validation techniques such as behavior reproduction tests [22], (which can be either qualitative or quantitative) can be used to assess how well a models results align with data. For example, is the model capable of replicating the different behavior modes that the system has historically produced? Are the patterns and events that have been seen in the system generated sufficiently by the model? Are important characteristics of the results, such as frequencies of significant events and relevant outputs generated under different scenarios, captured by the model output? How well does the model replicate historical data, and how well does it predict future events? To answer these questions, various methods of calibrating and empirically validating a model are available [63]. One common method is cross-validation, in which models are calibrated to a subset of the data and then validated using the remaining data [23]. Comparison of model results to real-world or experimental data can be done using subjective methods (such as Turing tests) or quantitative metrics.

Because comparative validation methods are so important, data on social systems can greatly enhance the confidence in the model attained by the validation process. Such data are often difficult to attain, and in many cases creativity must be used to identify data sources that can contribute the desired information. Since socio-cognitive agent based models necessitate the use multiple disciplines, each with different levels of granularity, they require several types of validation data. These models often include representations of cognition, individual behavior, interactions between agents, resulting emergent behaviors, non-social characteristics of the system (including economics, environmental conditions, etc.), and system-level outputs. An ideal validation dataset would include information about all of these factors. In many cases, however, data on the system being modeled is sparse or unavailable. Creativity must often be used to find applicable datasets that cover cognitive, behavioral, physical, and system-level data for the system and situation of interest.

Previous modeling efforts have used many different data sources to validate socio-cognitive models along cognitive, behavioral, physical, and system-level dimensions. System-level and physical data are often available, usually as historical data, and can be used to assess the overall model. For example, [19] used historical data about economics, social dynamics, and religion to validate a model of the Anasazi. [47] used historical data on water consumption to validate a behavioral model. Behavioral model components can be compared to historical, experimental, or other behavioral data. For example, [32] validated a model using empirical data on human behavior from economic experiments, and [3] used recorded observations of astronauts as validation data.

Behavioral and cognitive data can be elicited from experts or stakeholders [47, 3]. Some studies have used role-playing games to explain model structure to stakeholders, which then allowed those stakeholders to asses the model [29, 10] and to clarify and communicate decision-making strategies and behavioral patterns. Cognitive data are often difficult to attain, so these components are often validated subjectively. Some studies have considered cognitive model components to be validated if they are based on psychological theories that have themselves been validated against empirical data [47, 32].

A variety of validation methods should be used to attain the greatest possible confidence in a socio-cognitive model. Possibly the most valuable class of validation methods for these models involve comparing model results to real-world data. Data on cognitive, behavioral, physical, and system-level components are needed for comprehensive analysis of socio-cognitive agent based models. These data are often difficult to attain, and researchers have used varied and creative ways of attaining or finding substitutes for these data. A dataset that includes cognitive, behavioral, physical, and system-level data would be highly valuable for validating socio-cognitive agent based models.

Behavior Influence Assessment (BIA) Model

In our studies, we utilized the Behavior Influence Assessment (BIA) modeling framework as our targeted environment in studying how MMOG datasets could inform and validate ABMs. BIA [7] uses a system-dynamics framework for simulating systems that involve human decision-making. BIA constructs an internal model of decision-making for individuals/groups, governed by a system dynamics model representing known theories of psychological and social behavior. These models interact within a larger system dynamics model that reflects physical, social, or economic activity between each of the individual/group models.

Figure 4.1 illustrates the inter-workings of the BIA framework. Each simulated iteration begins with a BIA model presented with stimuli, or cues, founded on some defined parameter space. These stimuli form and/or manipulate the cognitive perceptions used within the model, an internal representation of the world or situation the model embodies. If certain perceptions persist during the simulation, these perceptions solidify into expectations by the model of the world/situation. Any differences between the model's current perceptions and formed expectations create the model's discordance with the world/situation. Both discordance and cognitive perceptions feed into the model's belief structure, which encompasses many cognitive traits, such as theory of planned behavior [4], which govern the decision-making based how the model handles factors such as social norms, attitudes, and perceived control. Intentions form based the model's updated beliefs that determine realized behavior

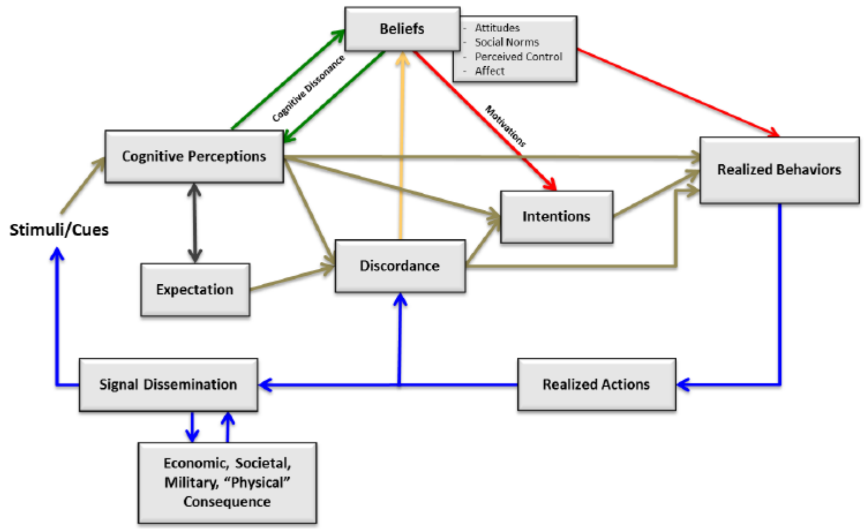


Figure 4.1. Conceptual drawing of the Behavior Influence Assessment (BIA) Model

and actions taken by the model. These actions feed into a signal dissemination to feed the model's resulting output (actions) as stimuli into other individual/group models within the system.

Constructing a BIA model requires establishing not only the input and output parameter space, but also defining the equations that govern the specified cognitive processes within the model. Subject matter experts, historical data, and known psychological and social theories are used in modeling a particular situation. This generates a highly extensible framework for modeling individual/groups and their interactions at varying levels of fidelity. However, this flexibility requires having substantial knowledge on the BIA framework, increasing time and effort needed for building models of a given situation.

How BIA Models Are Typically Built

BIA models typically begin with a question of interest, often posed quite generally, e.g. *What are the prospects for relations between country A and country B?* Because the scope of such questions is too broad to indicate a starting point for modeling, the question must then be refined via dialogue between customers and the modeling team. As an example, a sufficiently refined question might be, *How might an increasingly protectionist trade policy on the part of country A affect upcoming elections in country B?* A question of this form contains the seed of a model by having two quantities of interest explicitly called out: degree of protectionism and electoral results, along with an interest in their relationship.

Modeling then begins with a comprehensive review a variety of sources of information and concurrent distillation to yield one or more *dynamic hypotheses*. Typical sources of information include interviews with subject matter experts (SMEs), publicly available articles or accounts, and intelligence reports. The resulting dynamic hypotheses concern causal relationships between quantities of interest, e.g. increased protectionism on the part of country A might be seen as decreasing the GDP of country B, which could in turn reduce support for the ruling party in country B in the upcoming elections.

Requirements for Application of BIA

Applying BIA to the game data is markedly distinct from past applications of BIA. One difference is simply that we do not start with a particular question or extrinsic motivating interest. More striking, however, is the great differences in the kind of information available concerning the target domain. For the game, we have essentially three sources of data: (1) folk psychology (i.e., common sense wisdom concerning human behavior), (2) the official rules of the game, and (3) the extensive, detailed relational database.

Only one of these sources, folk psychology, is common to both past applications and the gaming data. While folk psychology helps us to calibrate our basic expectations of behavior, it only goes so far. While the remaining game-associated information is marvelously detailed, what is only minimally available in comparison to past BIA applications is any substantial narrative about specific individuals or groups in the game.

This narrative generally indicates who the relevant actors might be, what motivations they might have, and what relevant actions they might take, all of which is summarized by modelers in the form of a dynamic hypothesis. It is dynamic hypotheses expressed in the form of equations that ultimately make up the model.

At a more granular level, BIA models are about choice of action. Associated with each BIA-modeled action must be one or more *intentions*, each prompted by one or more *cues*. Normally, the actions to be modeled follow from the question of interest and related dynamic hypothesis. For validation purposes, what we instead require are actions along with plausibly associated cues (intentions may be inferred) that we can observe in the data. Validation could then be explored by supplying the model with cues observed in the data and comparing model output to subsequently observed actions.

A Path Forward

An aspect of the data that would be of particular interest to model would be trading behavior. Such behavior would fit well with the motivation for McFaddens Qualitative Choice Theory [44], which underlies the equations modeling choice of intention in BIA.

While individual trades are observable in the data, the cues associated with such actions present more of a challenge. While the likely most important cue for trades, prices, is not directly available, a sampling approach seems plausible for at least roughly determining which prices individuals are exposed to (at specific locations) before deciding whether or not to trade.

As for the narrative requirement, there is an implicit (albeit extremely thin) narrative associated with prices and trades: In general, individuals seek to buy low and sell high. Beyond this core dynamic, it might be possible to infer additional cues associated with specific commodities and/or individuals, e.g. low fuel reserves likely prompt the purchase of fuel, etc.

Interestingly, however, the volume and completeness of the data may at least partially offset the need for a richer narrative. In typical BIA applications, some fraction of the narrative serves as the basis for a *world model* upon which cognitive models may act and which may in turn supply cognitive models with many of their cues. Here, we intend instead to use the game data as a continuous source of cues (such as prices of specific commodities at specific locations).

Chapter 5

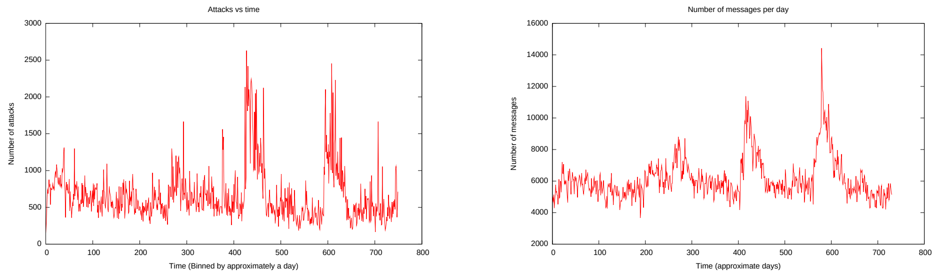
Motivations to Join Conflicts

Introduction

Understanding and anticipating changes in complex social systems, such as those relating to economies, financial institutions, and conflict is a problem of importance for national security. The complexity of these systems, such as the large number of factors and the human element, makes gathering data and running controlled experiments difficult. Promising methods such as modeling and simulation have made headway, however they are also subject to additional complexity issues and may face limited applicability.

MMOGs afford and promote complex social interactions amongst hundreds to thousands of players in online fictional worlds, attracting players from a wide variety of backgrounds, age groups, and genders. MMOGs serve as a tractable way of analyzing complex social interactions, due to two important features. Firstly, they serve as environments with a high-degree of expressivity, i.e., they allow the participants (also known as "players") to pursue a wide variety of complex social actions, in broad categories such as peer-to-peer and group communication, economic trading, and congregating with other players. Secondly, due to the virtual nature of the environments, MMOG's are able to capture a great amount of data and at high-fidelity, often simultaneously tracking the actions of all individuals in near-real time. MMOG's have been used to study phenomena such as education [52, 53], social networks [20, 56], and financial systems [5, 49]; our focus is to identify behaviors related to Large-Scale Conflict (LSC).

This chapter covers our review on how large-scale complex behavior emerges within Game X and draw parallels between virtual-world LSC's and real-world LSC's. Our data set covers the actions of players in this MMOG for a period greater than one year, which we analyzed to identify the following LSC behavior of interest: under what conditions does a game player participate in a LSC? We analyzed this information with the goal of informing an agent-based model that predicts when any one person is likely to engage in conflict. Our methodology involved identifying virtual world behavioral analogues to real- world behavior of interest (i.e. insurgent behavior), and analyzing the virtual behavior with real-world predictive models of participation in LSC. To that effect, we employ a (previously derived) theoretical framework that analyzes the determinants of participation in the civil war of Sierra Leone. This framework identifies three general theories that collectively predict participation in civil war



(a) Number of attacks over Time (b) Number of messages sent over Time

Figure 5.1. Combat and messaging patterns throughout the entire Game X data set. The number of combat attacks and the number of messages sent both spike during the periods of war, consistent with real world accounts of conflict.

(a type of LSC); we operationalize one of the theories (the *Theory of Social Sanctions*), and look at how virtual insurgent behavior can occur as a function of community networks, which are assumed to impose social sanctions for non-participation in a LSC. These communities are defined by communication patterns, as well as virtual group co-memberships. Generally, our hypotheses predict that the more members of a players community are involved in a LSC, the more likely the player will engage and be active in the LSC. Our results apply to a virtual setting, and we discuss how they might generalize to real-world setting, which is of primary concern.

Large Scale Conflicts in Game X

Analogues to real world Large Scale Conflict

Despite some of the issues highlighted in the previous section regarding mapping virtual world to real world phenomena, we noticed very clear analogues between virtual- and real-world LSC. In particular, during both war periods of Game X, we noticed a significant spike in both the number of combat attacks and the number of messages sent, as seen in Figure 5.1. This is consistent with real-world accounts of conflict, in which participants of armed conflict see an increase in mobilization and coordination for combat.

In addition, our virtual LSC is consistent with real LSC as it relates to the number of participants that participate in combat actions. Specifically, our virtual world LSC exhibits a pattern over the number of attackers that pursue different levels of attacks, which is consistent with the power-law distribution [18] as exhibited in other accounts of participation in armed conflict [12]. This distribution is illustrated in Figure 5.2.

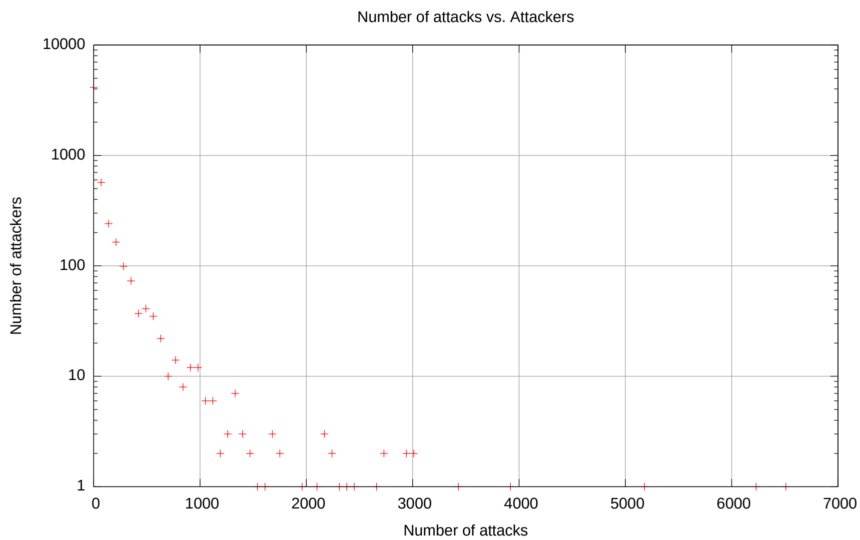


Figure 5.2. Distribution of Number of Attacks v. Number of Attackers during the first Game X war. Consistent with other accounts of participation in armed conflict, our virtual world LSC data is well-modeled by a power law.

Theoretical Framework for LSC Analysis

Given that we identified some parallels to real-world large-scale conflict, we were interested in understanding *a priori* what were the theoretical potential reasons participants would engage in armed conflict. One very influential framework in trying to identify potential reasons, developed by Humphreys and Weinstein [31], analyzed the determinants of participation and non-participation in the civil war of Sierra Leone. Humphreys and Weinstein originally sought to identify which of three "competing" theories better explained insurgent and counter-insurgent participation and non-participation in Sierra Leone. Through their analysis, they found support for all three theories, suggesting that the theories should not be taken in contrast to each other, but rather as an ensemble, capable of identifying multiple influencing factors that affect participation in an armed conflict. One of the three theories was particularly interesting due to its applicability to our game environment: *The Theory of Social Sanctions*. This theory predicts that an individual's participation in large-scale conflict is a function of the community that the individual is a part of. If the community is *strong*, then it can bring to bear a social pressure that will prompt individuals to fight in the conflict on behalf of their respective community. A strong community is (for instance) defined by (1) shared core beliefs and values, (2) close and many-sided relationships between the community's constituents, and (3) activities of reciprocity between the community's constituents [59].

Behavior of Interest: Dimensions of Combat Behavior as a Function of Community Networks

We operationalized the Theory of Social Sanctions in the context of our game, and were interested in answering the following questions, solely on the basis of an individual's community:

- Will the person engage or not? (i.e. participation)
- Will the person be an active agent in the engagement or not? (i.e. activeness)
- How fast will the person engage? (i.e. time to first response)

Thus, we developed the following set of hypotheses, that are explored in the remainder of this paper.

- ♢ H1: The greater the amount of community participation for a player, the more likely the player will participate in conflict.
- ♢ H2: The greater the amount of community participation for a player, the more the player will participate in conflict.
- ♢ H3: The greater the amount of community participation for a player, the faster the player will participate in conflict.

Experimental Methodology

Operationalizing the Hypotheses

To make some of the hypotheses more precise, we introduced operational definitions for the terms "participation," "community," and "community participation" To be considered a participant of the virtual LSC, an individual had to commit at least one combat action during the war period under study. An individual's community could be defined in three ways; each of them represents a different dimension of interaction between the players of Game X and tried to capture the spirit of the definitions used by Humphreys and Weinstein [31].

Definition 1 (Friendship Community Definition)**.** *Let p and pc be players. For any $pc \neq p$, pc is in p's community if p and pc are bidirectional friends and pc is* not *in p's hostile table. A bidirectional friendship between two players p and pc exists when p is on pc's friend table and vice-versa.*

Definition 2 (Communication Community Definition)**.** *Let p and pc be players. For any pc $\neq$ p, pc is in p's community if pc actively communicates with p. Specifically, p and pc are in each other's community if they send and receive at least 4 messages between them during the war period under study.*

The threshold of 4 messages is arbitrary, and was chosen on the basis of the trend of sent and received messages across all players during the war period under study. Specifically, during the first war period, approximately 50% of players had less than 4 messages sent and received.

Definition 3 (Guild Co-Membership Community Definition)**.** *Let p and pc be players. For any pc $\neq$ p, pc is in p's community if p and pc belong to the same guild for a majority of the war period under study.*

A natural inclination is to use the intersection of all three communities as a definitive measure of community. However, such a combination did not yield statistically significant results. In addition, we also felt it better to study different types of communities to see whether or not the expected behaviors appeared throughout. Therefore, each hypothesis has three variants, one for each definition of community. Finally, to define community participation, we chose to represent it as the proportion of players within a community that were active during the war period under study. This ensured that the numbers weren't too biased for large communities, by ensuring that all the community participation statistics varied within a common range ($[0.0 - 1.0]$).

Participants

Our data set includes data for over 50,000 players across ¿700 days. The data set is historical, beginning in 2007. All participant data has been anonymized and all players agreed as part of Game X's sign-up process to have their data collected for purposes of scientific research. Despite the magnitude of the data, only a small percent of players were actually considered as part of the analysis; several players did in fact sign up, but did not participate enough in Game X to consider their presence meaningful. A great majority of players did not sign-in to play for more than 10% of the entire data set time period. After excluding these players, our participant pool was reduced to 6,156 players. Approximately 13.56% of players were female and 86.44% were male.

Participants of Interest

Our focus was on the first Game X war. After having filtered the data once to remove inactive players, we filtered the data again, this time filtering by two measures: "cumulative actions taken prior to war period" and "log-in percentage". In our description of Game X in Section 3, we discussed how players had a limited number of turns per day; "cumulative actions taken prior to war period" is a measure of how many turns they have taken per day

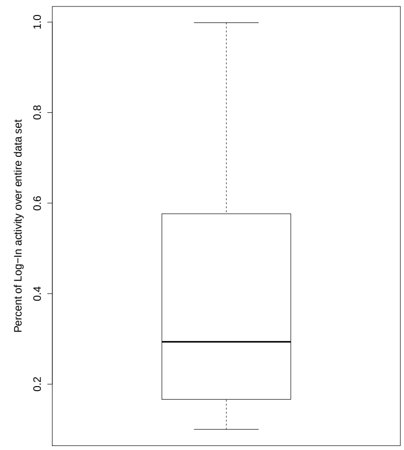


Figure 5.3. Box plot of log-in percentage. Log-in percentage is determined by the number of days that players logged in over the total number of days in our data set. We filtered the data for all players whose log-in percentage was less than 10%.

across all days prior to the start of the first Game X war. The threshold for consideration was 500,000 turns taken prior to the start of the first Game X war. The reason for filtering by actions taken was because we wanted to control for players who were signing-in and not doing anything, which does not represent the behavior we were interested in studying. The measure "log-in percentage" is defined similarly to how it was defined previously. However, the threshold for consideration was 80% as opposed to the original 10%. These combined filters reduced our participant pool from 6,156 players to 981 players. Of these, approximately 11.62% were female and 88.38% were male.

Hypothesis Tests

Our hypothesis testing was restricted to the first Game X war. For all cases of the hypothesis tests, the variable `communityParticipation` is the proportion of the player's community that was active during the first Game X war. Given the operational definitions, and our hypothesis set, we tested each hypothesis as follows:

- ◇ H1: Likelihood → Logistic Regression of the variable `hadWarAction`, which took value 1 if the player had at least one combat action during the first Game X war and 0 otherwise, over the variable `communityParticipation`. H1 thus predicts that the greater the proportion of `communityParticipation`, the more likely the response variable `hadWarAction` will be 1.

Table 5.1. Hypothesis Test Results

Hypotheses	Community Definitions		
	Guild Community	Friend Community	Communication Community
H1: Likelihood (Logit)	1.633 [0.604]**	-4.703 [0.826]***	0.404 [0.247]
H2: Amount (OLS)	0.669 [1.697]	-9.748 [2.394]***	-1.952 [1.124]

Notes: Standard errors are in brackets. **Significant at 5%; ***Significant at 1%.

- ♢ H2: Amount → Ordinary Least Squares Regression of the variable `numWarAction`, which is a count of the number of combat actions taken during the first Game X war, over the variable `communityParticipation`. H2 thus predicts that the greater the proportion of `communityParticipation`, the higher the variable `numWarAction` will be.

- ♢ H3: Time to First Attack → Survival Analysis of the variable `timeOfFirstAttack`, which is a number indicating the day the player's first combat action during the first Game X war was registered, over the variable `communityParticipation`. H3 thus predicts that the variable `communityParticipation` will lower the survival function of the variable `timeOfFirstAttack`; in other words, `communityParticipation` will predict how quickly a player commits his or her first combat action as measured by `timeOfFirstAction`;

Results and Discussion

The results of our hypothesis tests are shown in Table 5.1. Survival analysis for H3 did not yield statistical significance for any case, thus we omit the results from the table. We achieved different results under different community definitions, with the communication community not yielding statistical significance for either hypothesis.

The results found indicate a surprising interplay – while friend community participation does *not* affect one's likelihood to participate, your guild community does. This seems to indicate the importance of guild membership and community over other relationships. The non-significant result for H2 with guild community is odd, since from H1 we would assume a greater number of attacks. We posit that a "free rider" effect may be occurring, where individuals in a community may participate, but leave the bulk of combat to others who are better suited for combat. In further work we are looking at the nature of guilds and whether they have a heterogeneous set of players (in terms of skills).

Conclusion

In this chapter we have compared one particular type of complex behavior, large scale conflict (LSC), in a MMOG (Game X) and in the real world. A high level similarity was seen – a power law distribution of the number of attacks vs. the number of attackers – which corresponds to known patterns in real world conflict. Assessment of community influence on player participation was surprising, as individual friendships did not have a positive influence; however guild communities did.

This lack of correspondence is interesting, and there may be several causes for it. Firstly, the hypotheses we tested were developed (and evaluated) with data from a civil war. This type of conflict may not be as relevant for Game X conflicts. Secondly, there could be a strong "free rider" effect – the more my community is willing to assume the cost of war, the more likely I am going to abstain and "free-ride" on their participation. This could explain the negative interaction for the friend community case. Thirdly, there may be a division of labor within guilds; with some individual being more combat oriented and others more economically oriented. This could explain why there was a positive effect for H1+guild community, but a non-significant effect for H2+guild community – others may be taking on the combat roles.

Further work will focus on addressing these issues to identify the reasons for why players participate in LSCs.

Chapter 6

Guilds Organization

Roles in Game X

An interesting feature of groups is the emergence of roles, ”coherent sets of behaviors expected of people in specific positions (or statuses) within a group or social setting” [24, pg174]. We believe that guilds in Game Xexhibit roles as well. In this section we describe an experiment in identifying roles through a clustering analysis of player ”skills”

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. The tight correlation between skill level and behavior allows us to use ”skills” as a proxy for behavior.

There are 10 skills that players can have (veteran player have access to more advanced skills; we just consider these 10). The skills can be divided into 3 groups: *Combat skills*, which help player attack and defend; *Gathering skills*, which help players gather resource more effectively; and *Movement and hiding skills*, which allow a player to move and hide better. Finally there is one skill that is oriented toward repair. We labeled these skills Combat1 . . . 4, Gathering1 . . . 4, and Movement1, Movement2, and Other (for the repair skill).

Our sample set of players was drawn from the first day of the first war. We chose players who were ”experienced and dedicated”– they have player for more than 500,000 turns, and they participated in the war.

Figure 6.1 shows the variety in guild sizes. Guilds vary greatly in size, although they only need 1 member to exist. There are, however, many guilds that have more than 10 members. Organizations of this size may need some structure (i.e., roles) to pursue goals.

We filtered out players from guilds that were less than size 10, these would be at too early of a stage to have the potential for exhibiting organization. This resulted in data from 3185 players who represented 89 guilds.

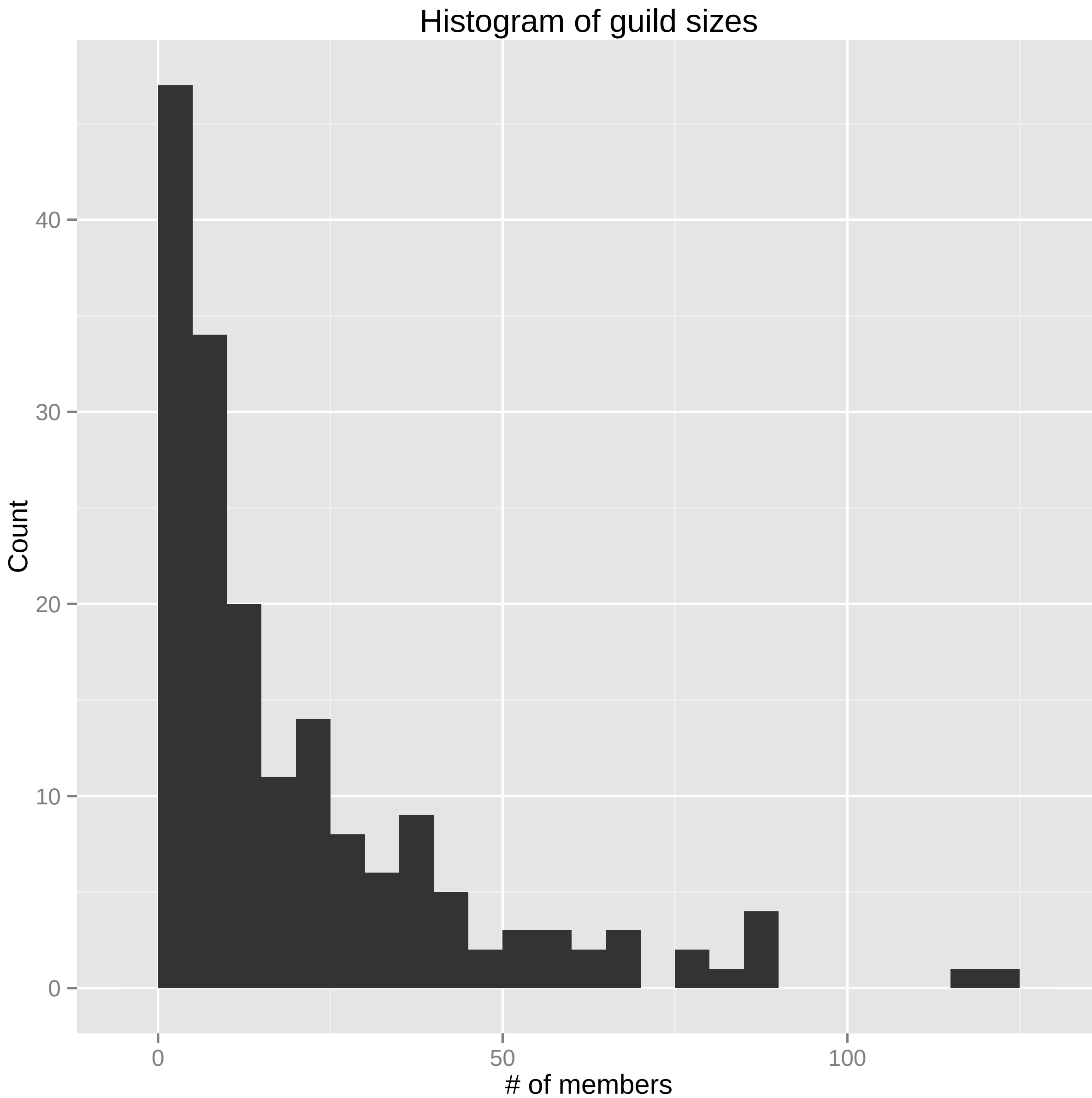


Figure 6.1. Histogram of guild sizes

We have two research questions:

RQ1: Do guilds contain different roles?

Our second question, if **RQ1** is true, is:

RQ2: Do all guilds have the *same* roles?

We took a unsupervised clustering approach to address these research questions. All clustering and computation was done using the cluster package [43] in the R statistical computing language [60].

Roles in Guilds

We first clustered the players of each guild using the *Partioning Around Medoids* (pam) algorithm provided in the cluster package [43, 35]. pam is a $k - medoid$ clustering algorithm where k is the number of clusters the data should be divided into.

We iterated over 10 values of k, for each k we stored the average silhouette value of each guild. The silhouette coefficient combines a measure of cohesion and separation into a single value. The coefficient is calculated in the following way: (from [57, 541] and [35]):

1. For a point i, calculate it's average distance to all other points in it's cluster – call this a_i. This is a measure of the cluster cohesion.
2. For each cluster *not* containing i, calculate the average distance from i to all the points in that cluster. Find the minimum (over all the clusters) of this value and call it b_i. This is the closest other cluster to point i – thus a measure of the cluster separation.
3. The silhouette coefficient for i is:

$$s_i = \frac{b_i - a_i}{max(a_i, b_i}$$

s_i ranges from -1 to $+1$. It is $+1$ when $a_i = 0$ – that is all points in the cluster are the same. It is -1 when $b_i = 0$, which indicates that i is closer to all other clusters. In general, a positive value indicates that the distance from i to other clusters is larger than the distance from i to members of it's own cluster. A negative value indicates the opposite.

The silhouette value can be used to find the "correct" value of k. The process is to look at the silhouette value over multiple values of k and identify an "elbow" in the graph where a large decrease in the silhouette coefficient occurs [54].

Figure 6.2 shows a box plot of the silhouette values for all guilds per k value. We can notice an "elbow" between $k = 2$ and 3. There is a significant difference in the mean silhouette value over these two ($t = 13.1973$, $df = 175.83$) In addition, the silhouette value is

	Medoid 1	Medoid 2
Combat 1	33.05	21.28
Combat 2	43.99	25.18
Combat 3	51.88	28.53
Combat 4	10	10.32
Economy 1	17.3	14.76
Economy 2	10	13.29
Economy 3	13.56	11.38
Economy 4	19.24	16.9
Movement 1	11.3	10
Movement 2	46.4	24.99
Other	28.13	16.57

Table 6.1. Medoids of Guild 1

high for $k = 2$. This provides evidence that $k = 2$ makes the best sense in terms of number of clusters.

Clusters for a specific guild can help shed light on the clusterings. Figure 6.3 is a plot of the points in guild 1 along the first two principal component axes. We can see that the clusters, for this guild, seem to differentiate two different types of individuals. The cluster on the left looks more tightly packed than the other cluster however.

Table 6.1 shows the medoids for a particular guild (labelled "1"). Observe that one of the medoids is heavily combat based (high values for Combat 1, Combat 2, Combat 3) as compared to Medoid 2. This indicates a cluster that contains players that are more combat oriented than the others.

The other cluster does not seem to have much of a pattern. This could mean the guild has specific fighters, but everyone else shares the remaining tasks.

Do all guilds have the same roles?

The next question we ask is whether all guilds have the same roles. We address this question by doing a clustering of the medoids generated for each guild with $k = 2$. We take the medoids to represent the canonical values of the guilds, they represent the clustering. We call this "Medoid of Medoid" (MoM) method.

The basic idea is to take the 2 medoids from each guild and create a new data set with all the medoids from all the guilds. We then cluster over all the medoids. If the two medoids of each guild are distributed among the clusters, this means that the guilds had similar clusterings. Figure 6.4 depicts this method.

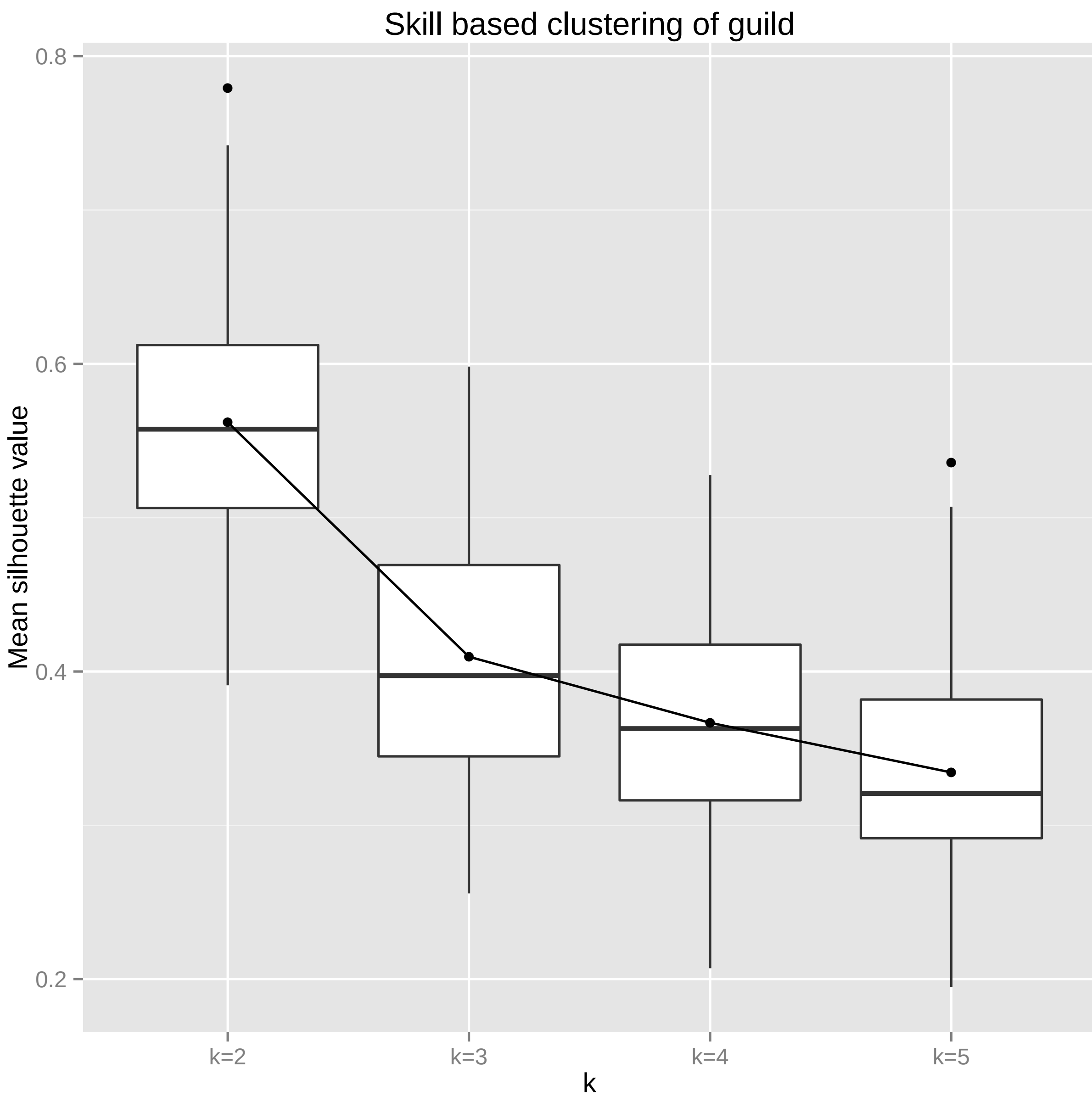


Figure 6.2. Boxplot of silhouette values over multiple values of k.

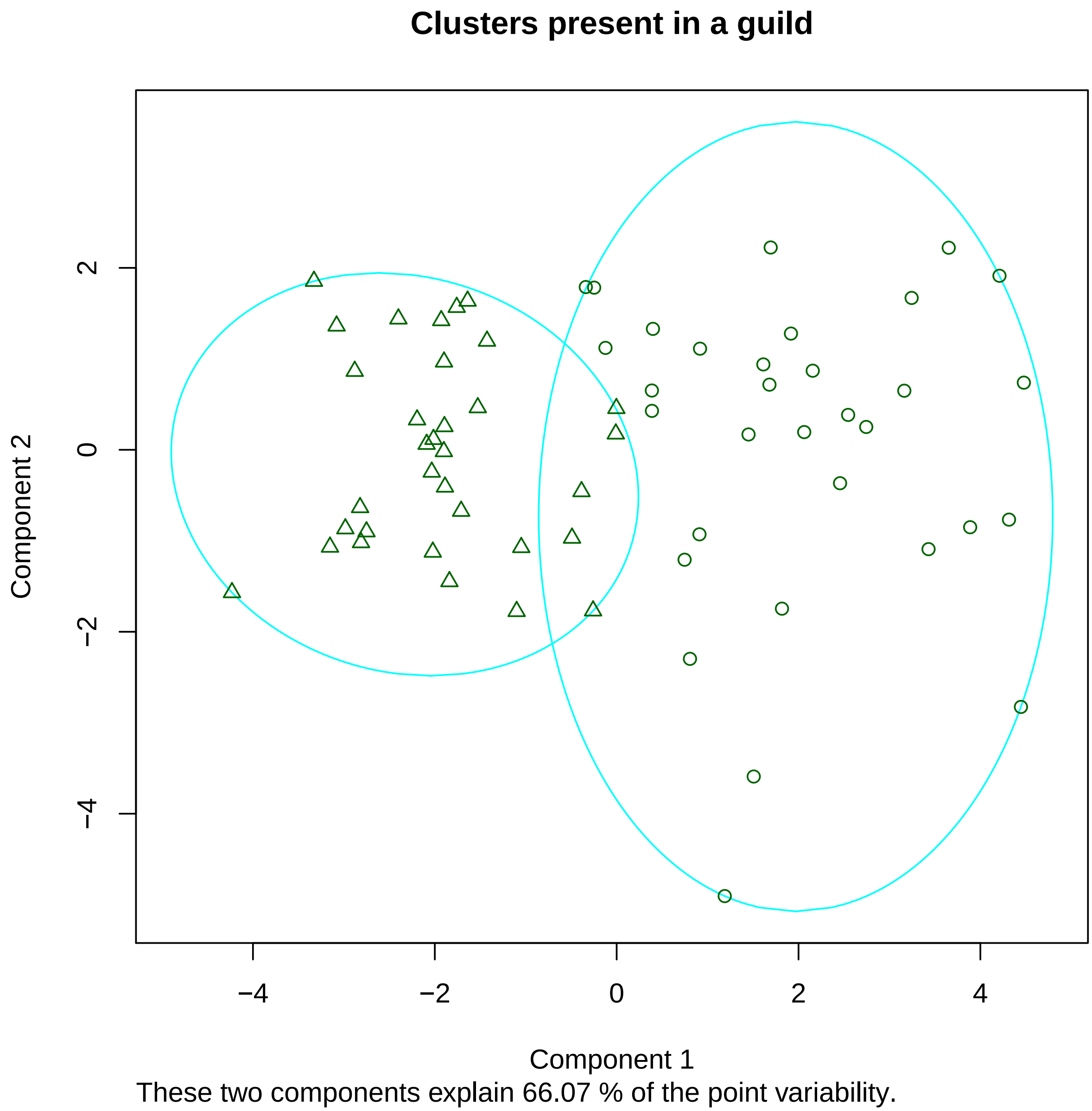


Figure 6.3. Plot of the clusters against the first two principal components for Guild 1.

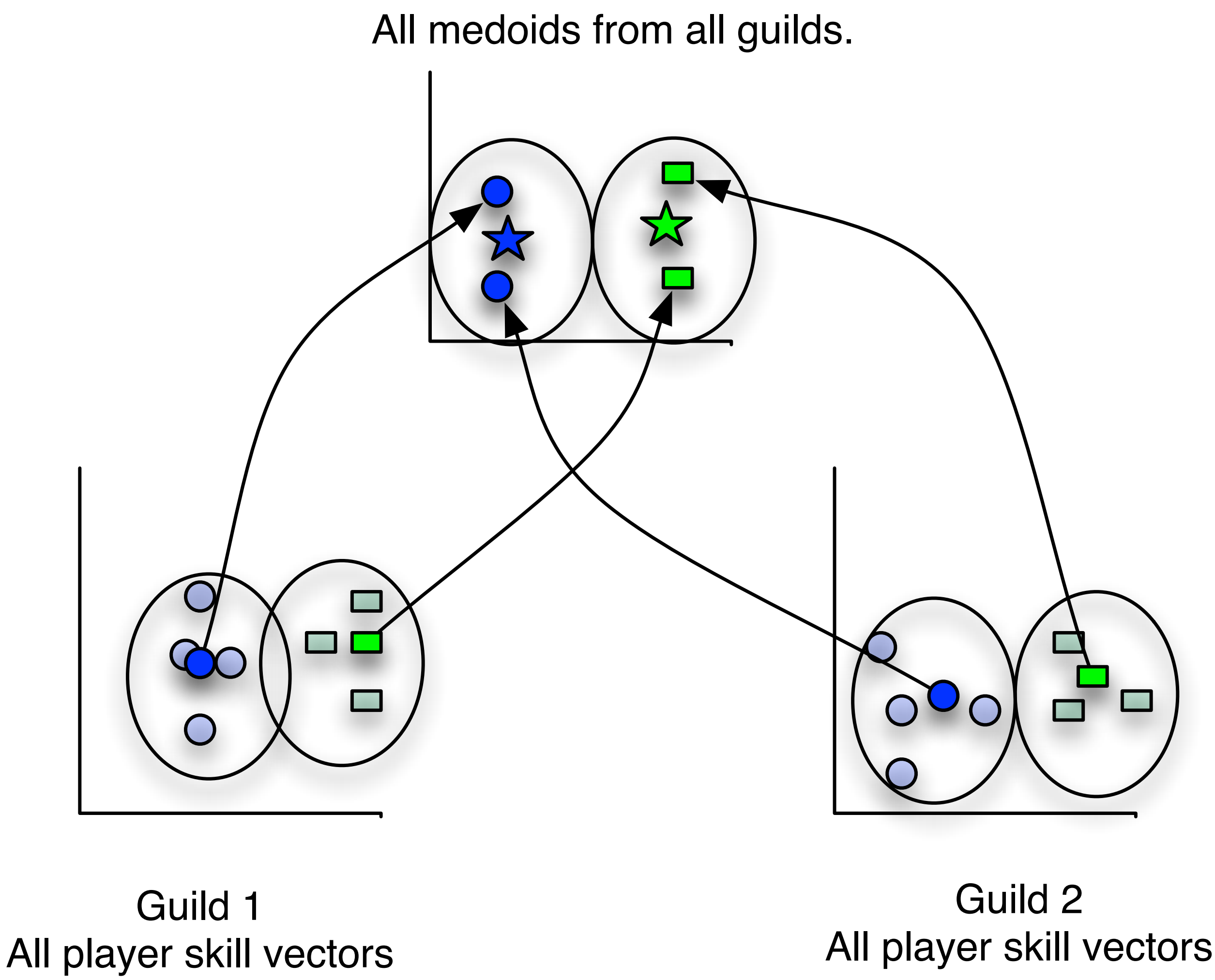


Figure 6.4. The Medoid of Medoid Method

	Cmb. 1	Cmb. 2	Cmb. 3	Cmb. 4	Ec. 1	Ec. 2	Ec 3	Ec. 4	Mvm. 1	Mv
Medoid 1	18.1432	20.9268	25.2066	10	17.8340	10.0000	11.9113	15.7311	10.0000	18.
Medoid 2	37.9142	58.5283	63.6472	10	20.3241	10.0000	13.1681	22.8533	15.2375	58.
Medoid 3	26.0851	37.6808	41.2855	10	17.8735	12.6567	12.5931	19.9724	11.4194	38.

Table 6.2. Medoids of the MoM clustering. Cmb = Combat, Ec= Economy, Mvm=Movement

Figure 6.6 shows the mean silhouette value over several different values of k. We (once again) used the *pam* method. While the silhouette value is still high, we see less of an elbow. The greatest gap is between $k = 3$ and $k = 4$, so unlike the previous clustering, the best clustering of the medoids is for $k = 3$.

Table 6.2 shows the medoids of the clustering. We can see that cluster 1 reflects skills that are *not* strong in combat related activities. Cluster 2 and 3 are focused on combat related skills.

Figure 6.5 provides a cluster plot of the medoids generated.

We can think of this clustering as only two clusters, cluster 1 being the non-combat, and cluster 2 and 3 being combat oriented. From this perspective, 78 unique guilds were represented in cluster 1 – indicating that 84% of the guilds had one medoid in cluster 1 and the other medoids in either cluster 2 or 3. This indicates that guilds had a similar clusterings – that is, most guilds had some people involved in combat, and the rest in non-combat.

Discussion

The first question we had was whether guilds can cluster. The clustering results indicate that guilds can be clustered based on skills, and the clusters may correspond to a subset of the guild being ”fighters”.

Our next question was whether all guilds have the same type of clustering. The results on this were mixed. The best clustering involved 3 different clusters, however 2 of the clusters were very similar (both combat oriented). The distribution of medoids over the clusters indicated that 84% of the guilds had 1 medoid in cluster 1 and another medoid in either cluster 2 or 3. This provides evidence that most of the guilds had the same division of medoids.

If guilds had different clusters, we would expect the MoM clustering to reveal several different clusters. Instead, it fit best with $k = 3$, and even then the clusters were really combat and not-combat.

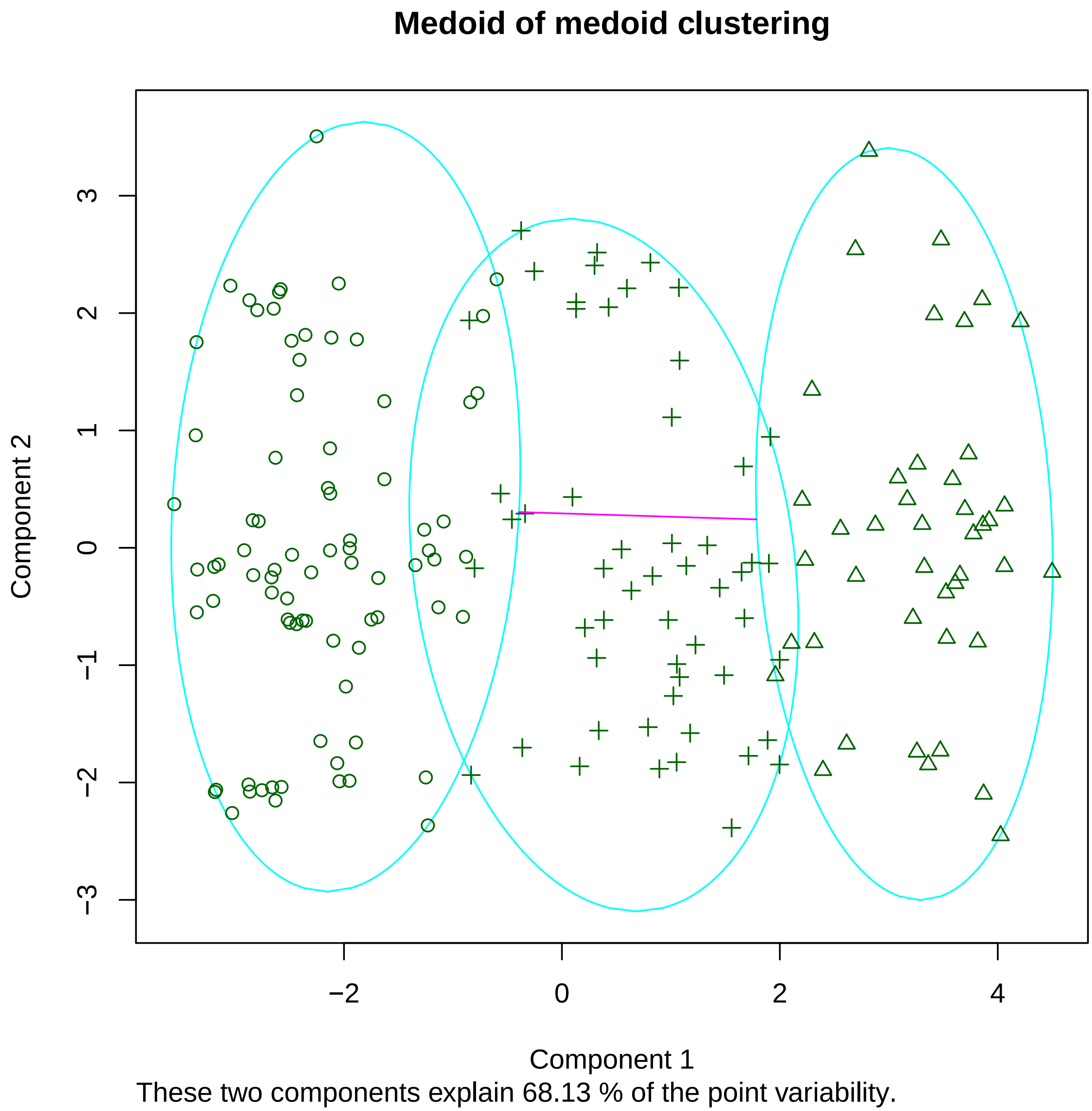


Figure 6.5. Plot of all medoids of all guilds on the first two principal components for $k = 3$

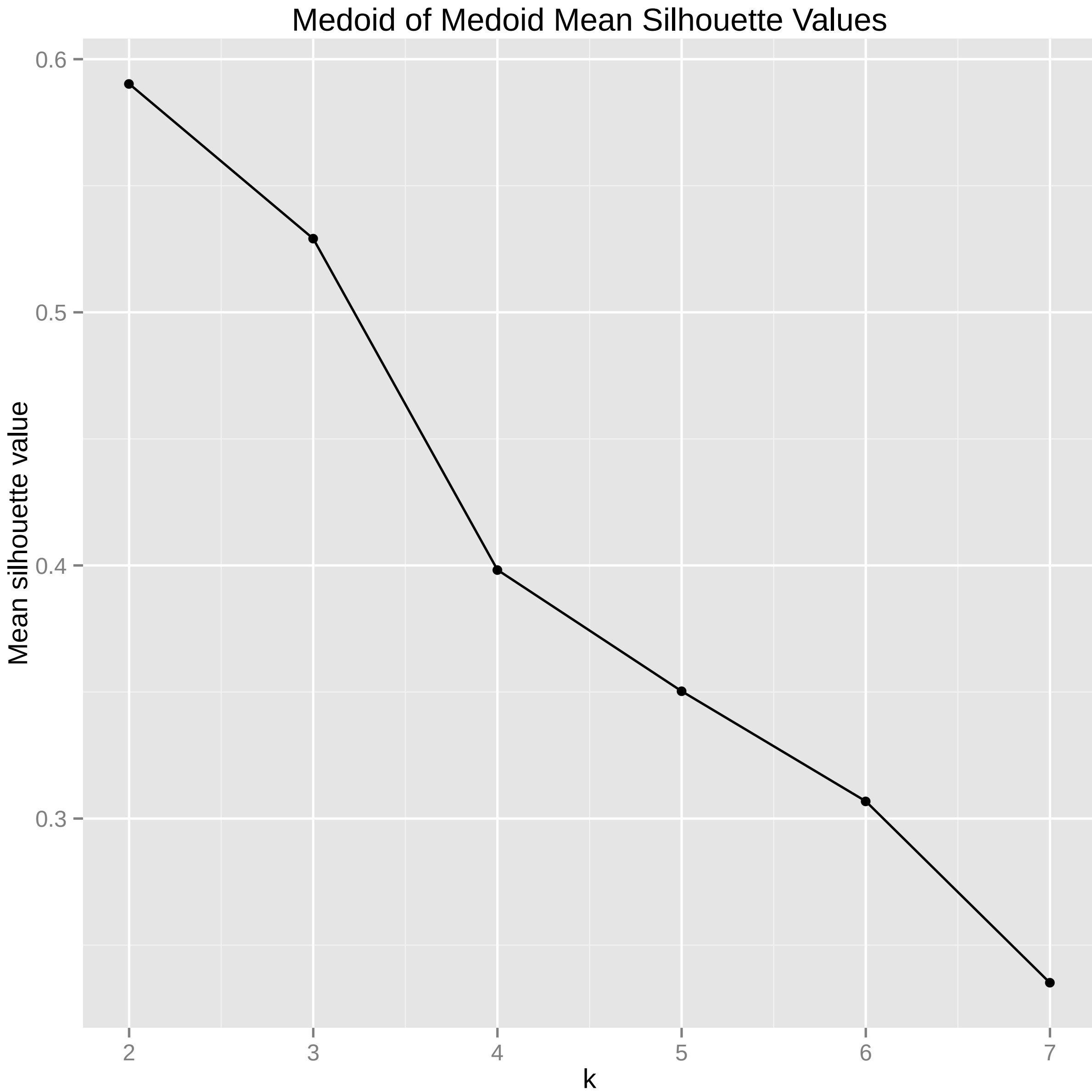


Figure 6.6. Plot of mean silhouette value for varying k.

Conclusions & Future Work

The goal of this work was to see if there is evidence supporting the concept of "roles" in guilds. Through an unsupervised clustering analysis method, we found evidence that:

1. Individual guilds seem to cluster well into two clusters.
2. The two clusters of each guild seem to correspond to the roles of a "fighter" and, "not-fighter".
3. These clusterings are prevalent over all the guilds.

There are several paths forward.

First, we are using a skill vector as a proxy for behavior. However, wed have access to behavior data, so we can create feature vectors for individuals that contain all the behavioral traces over a period of time. This may be better than skill vectors because skills may not reflect current behavior – I may have been a fighter before, but not now.

Secondly, we are using the intuitive "elbow" method to identify the best value of k. There are many other methods to address this problem that we plan on exploring [54].

Thirdly, we currently are looking at a single time point. Another important question is how do roles develop, i.e., the problem of *role differentiation* [24]. Our dataset contains several guilds that started, rose in membership, then declined and eventually disbanded. With this full set of data we can analyze the emergence of roles within the guilds as they grow in power.

Chapter 7

Predicting combat from public communication

The overall goal of this effort was to see if knowing what individuals say can help determine their actions. As such, this falls squarely within the ”expression-to-action” thrust of the project.

The words that people use can shed light on their thoughts, behaviors and emotional states. The study of history is in large parts a study of the textual record of leaders of the past to understand their underlying thoughts and emotions.

Our Game Xdata set provide a unique data set by providing extensive records of all actions of a player *and* all public communication of the players within the game. Can we computationally model how what people say predicts what they do?

Related Work

There are more and more computational models to predict behavior as publicly available data sets continue to be collected. Often the method is to survey individuals and then study their behavior within social media (for instance Twitter) and identify words that predict characteristics of people.

Munmun De Choudhary [17] has extensive work on predicting health related problems based on social media data (such as Facebook and Twitter).

Eric Gilbert has work on predicting status relationships among individuals based on language from email within the Enron corpus [27], [26].

Drawing from these results, our hypothesis is that using public communication can help predict probability of combat by players.

We are going to focus on the probability of a *single* player committing an attack against *any* other player. Future work can incorporate “cooperative combat”, where multiple players join together.

Our task is made slightly easier by the fact that we are not predicting who the player is attacking, just that a player is attacking. Thus, we hope to obtain more general phrases that could indicate propensity of combat among others.

We only consider words and phrases. Many studies have shown that pronouns can also encode emotional states, focus, social relationship etc ([58]). For the current study we do not consider pronouns as a separate class of words.

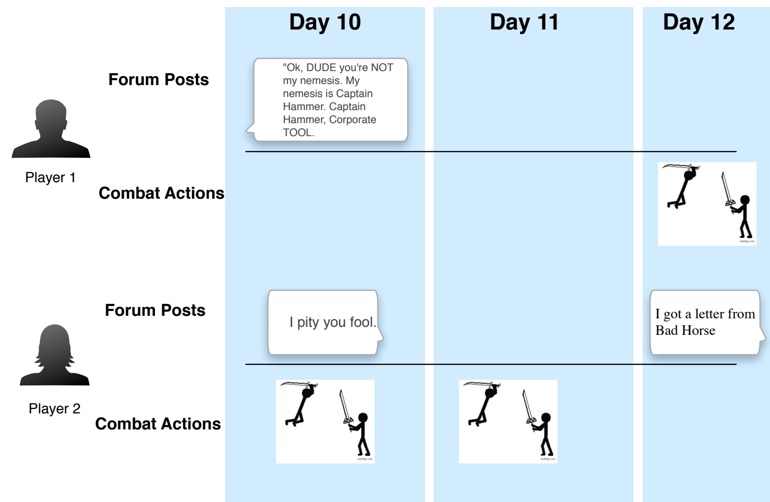


Figure 7.1. Example data set.

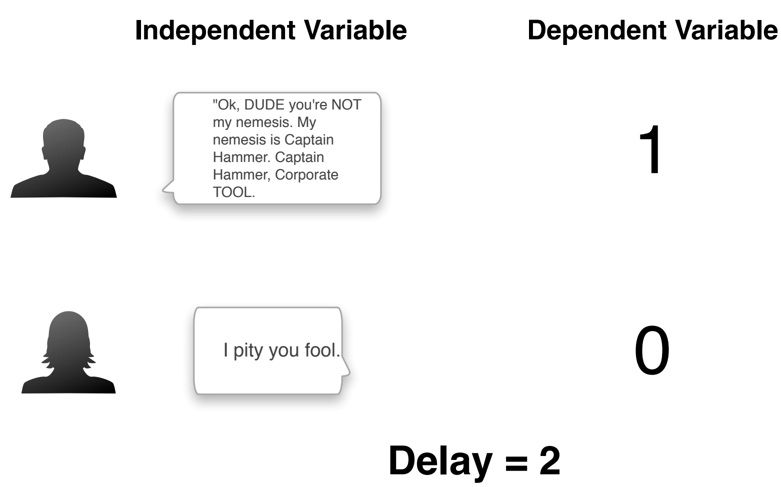


Figure 7.2. Assigning classes to the data.

	ok	dude	you're	not	my	nemesis	is	captain	hammer	corporate	tool	i	pity	you
"Ok, DUDE you're NOT my nemesis. My nemesis is Captain Hammer. Captain Hammer, Corporate TOOL.	1	1	1	1	2	2	1	2	2	1	1	0	0	0
I pity you fool.	0	0	0	0	0	0	0	0	0	0	1	1	1	1

Figure 7.3. Feature vector for terms.

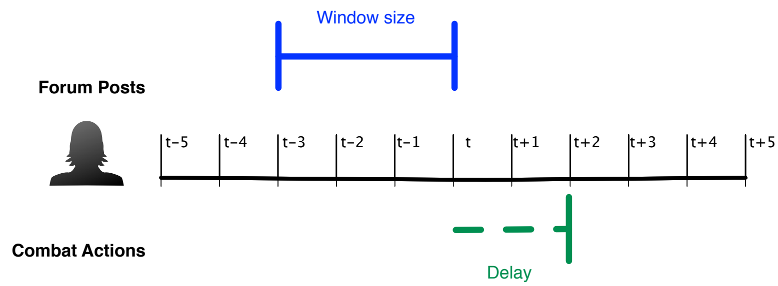


Figure 7.4. Illustration of window size and delay. Posts are aggregated over the space of the window, and a classification is given based on whether an attack occurred at the current time + delay.

Method

Our methodology draws from the one used in [26]:

Evaluation Period Instead of working with the whole dataset, which includes several special events such as wars etc, we choose a relatively innocuous evaluate period in the hopes that more general trends would be present here.

Data gathering and Cleaning We gathered all posts by all players during the evaluation period. We anonymized game specific terms to generic terms. For instance, names of particular vehicles were assigned to the same category "snlgamevehicle". Every post was aggregated with posts from w days before it. This reduces the data to a single feature vector per player per day. If player one posted anything on day 10, for instance, we would construct a feature vector with all the words from the posts on day 10, and the preceding days according to our window length.

Create and clean vector space representations Using 1,2, and 3-grams we constructed a vector space representation of each post. We removed phrases that were used by less than X number of people, and we removed phrases that were used less than Y times.

Assign classification For each feature vector we identify whether the player associated with the feature vector committed an attack d days after.

Identify words Using Elastic-Net regression we identify words that could predict combat.

Figures 7.1, 7.2, and 7.3 provide a graphical overview of process.

Things that we can try:

1. Predict cooperative combat instead of just individual combat.
2. Look at other aspects, such as economic status, etc.

The Data Set

Communication in Game X

Game X includes 3 methods by which players can communicate with each other:

1. Personal Messages: An email like system for communicating with other players, or in some cases groups of players.
2. Public Forum: A Usenet like system in which players can post topics and replies (see below).
3. Chat: An IM like system for players to chat with others in their guild.

The structure of the forum's are shown in Figure 7.7.

Each forum posts includes the name of the player who posted an image of their avatar in the game, and their guild affiliation.

Forum based measures

We have data on more than 700 days from the game. Included in this dataset are posts from 7 different forums within the game. One of the forums is meant for role playing (RP) discussion, that is all players must discuss in the role of their character. One forum is meant for Non-Role Playing discussion (NRP). The rest of the forums are both RP and NRP.

Table 7.1 provides some high level statistics of the forums. We can see that forum 2, which was only RP, was the most popular in terms of posts. However, the number of topics was low – indicating higher average topic lengths.

Forum	# Posts	# Authors	# Topics	Posts/Topic
1 (NRP)	16847	1494	**2468**	6.8
2 (RP)	**62669**	**2244**	1813	34.6
3 (NRP/RP)	35069	1909	1240	28.3
4 (NRP/RP)	9544	1391	223	**42.8**
5 (NRP/RP)	11047	1424	2091	5.3
16 (NRP/RP)	13326	1497	875	15.2
7 (NRP/RP)	1778	341	286	6.2

Table 7.1. Overview of post/authors/and topics for each forum. Bold entries are the max values for the column. These are calculated over the entire 737 day period.

There are three ways of communicating with others on the forum:

1. Create a new topic.
2. Post on a topic created by another player.
3. Post on a topic and quote another player.

Cleaning forum data

Forum's are a place to communicate in natural, informal ways. Thus, the data is often filled with slang, misspellings and grammatical errors. This makes parsing the data somewhat difficult, as multiple spellings of the same word can cause feature vectors with multiple features for the same word. In addition, the science fiction esque names of planets/star systems etc. was ripe for misspelling.

Fully addressing this problem is of great interest for the community as we try to parse "dirty" forum data.

For our purposes, we constructed a mapping between common misspellings of key game-related terms (such as the names of planets and races within the game). This reduced some of the variance for this terms. However, much more needs to be done to make this a much cleaner data set.

Vector Space Representation

A vector space representation of a document constructs a (usually long and sparse) vector of positive integers to represent a document. Each dimension is associated with a single word, and the value at that dimension is the number of times the word was used in the document.

Figure 7.3 shows a vector space representation for two documents. The columns represent dimensions, with the associated word. The rows are each document, and the document is shown on the left.

Two things to observe:

1. The dimensions should be the unique set of words over all documents. You can see that the word "pity" appears as a dimension, even though it only appears in the second document.
2. The vectors are often extremely sparse as the number of documents increases, since no document contains any significant portion of all the unique words.

In our situation we had feature vectors with several million dimensions.

Feature Selection

Elastic-net regression is a method to choose features as well as identify relationships between features and the dependent variables. It works on millions of features and prunes them down to as many as one would like.

Results

These results, while preliminary, illustrate some interesting aspects. Primarily, some of the high ranked words actually make sense in terms of predicting attacks. For instance in row 12 of Delay 0, Window 1 we see the term "nice hunting". This could be in reference to a person hunting another person. Within Delay 0, Window 1, you can see a few more "aggressive" words, such as “standard defense”, “were threatened” and "hunting comments head money".

The delay/window parameters that resulted in the highest AUC was Delay 0, Window size =3 (although these differences may not be statistically significant). Looking at the words in that list you can see more aggressive words than the other parameter settings. For instance, "delete snlgamevehicle" is an anonymized phrase that indicates deleting a specific vehicle from the game. Other phrases such as "navy claiming", "opened hostilities" etc. are obviously more aggressive.

Given the overall limited variance in the AUC, it may not be saying much to claim that delay 0 and window 3 has the maximum AUC. Future work will explore the statistical significance of the differences between different parameter values.

What is not captured here is the generality of these phrases. Are these words general over multiple people, or appropriate for single individuals?

row	D0 W1	D1 W1	D2 W1
1	head money bunny	want free	want free
2	macros	sank	tobe
3	want free	level	open snlguild
4	were threatened	silenceer	darksteel
5	jabberwock	tobe	level
6	insert name here	experience	experience
7	bounty snlmarks which	(repeated a's)	turns_used
8	standard defense	(repeated a's)	(repeated a's)
9	level	(repeated a's)	(repeated a's)
10	affiliation trip	above level	(repeated a's)
11	which placed	adari davonrai	above level
12	nice hunting	adari davonrai playerref	adari davonrai
13	snlmarks which placed	adds playerref	adari davonrai playerref
14	subject you collected	affiliation trip	adds playerref
15	experience	after day after	affiliation trip
16	turns_used	aligned base	after day after
17	nice hunting comments	aligned members	aligned base
18	hunting comments	aligned snlguild may	aligned members
19	hunting comments head money	all work play	aligned snlguild may
20	(repeated a's)	allso	all work play

Table 7.2. Window size 1, Delay 0,1,2

row	D0 W2	D1 W2	D2 W2
1	event have been	your computers	event have been
2	nugz	criterion	enlighten you
3	someones snlfactoryoutlet	defeated playerref full	someones snlfactoryoutlet
4	zumon zumon	someones snlfactoryoutlet	woof
5	defeated playerref full	because he don't	defeated playerref full
6	woof	you're liar	because he don't
7	guild uta	we claimed snlspecificcityarea	drooling over
8	enter red zone	everything sank	everything sank
9	anything else from	drooling over	bar us
10	either way though	from keiran	rothstein
11	aid pirates	computers energy	central authority
12	caught owain	factory outlets fellow protectors	we claimed snlspecificcityarea
13	cose i	bar us	matches would
14	beeee	male morning	computers energy
15	affiliation trip attacked	merchants association	tobe
16	bad cop	deleted his acount	sage
17	stop hostilities	tobe	have code
18	you ran out	chaka	captain macaroni
19	powergaming bastards	silenceer	chaka
20	fuel usually available	aligned base color	oxidate

Table 7.3. Window size 2, Delay 0,1,2

row	D0 W3	D1 W3	D2 W3
1	nugz	visit from playerref	visit from playerref
2	delete snlgamevehicle	can consumed	loosing my
3	acceptable roleplay	loosing my	christmas hope
4	visit from playerref	consumption giving	take down two
5	can consumed	magic list	navy claiming
6	party party d	drooling over other	list me too
7	tense music	navy claiming	opened hostilities
8	consumption giving	opened hostilities	back economy
9	navy claiming	assume anything	three steps
10	dead energizer	back economy	assume anything
11	attention seeking pirate	built some	holds again
12	smart comment	your computers	syhon tried
13	opened hostilities	get snlmarketcenter s	would end ambush
14	would end ambush	three steps	samurai sword
15	looky	have little sense	sits out
16	back economy	repurchase account upgrade	main snltraderoute
17	have little sense	back you go	act mediator each
18	three steps	act mediator each	have little sense
19	all messages i	acceptable roleplay	acceptable roleplay
20	happy when people	main snltraderoute	christmas d

Table 7.4. Window size 3, Delay 0,1,2

Sentiment Analysis

The sentiment of a word, phrase, sentence or document refers to the emotional content. Is the author happy, or angry, or sad?

Sentiment analysis refers to computational methods to identify the sentiment of text. There is much previous work on sentiment especially within the review space. Oftentimes sentiment is distinguished between two categories, positive and negative sentiment.

The sentiment of posts could provide additional features the could predict combat. There are two general ways of doing sentiment analysis:

1. Have coders identify the sentiment of several hundred posts, then use machine learning techniques to learn a model of sentiment.
2. Use existing dictionaries of terms that have been established to predict positive/negative sentiment.

In this project we pursued both avenues.

Dictionary Based Sentiment Analysis

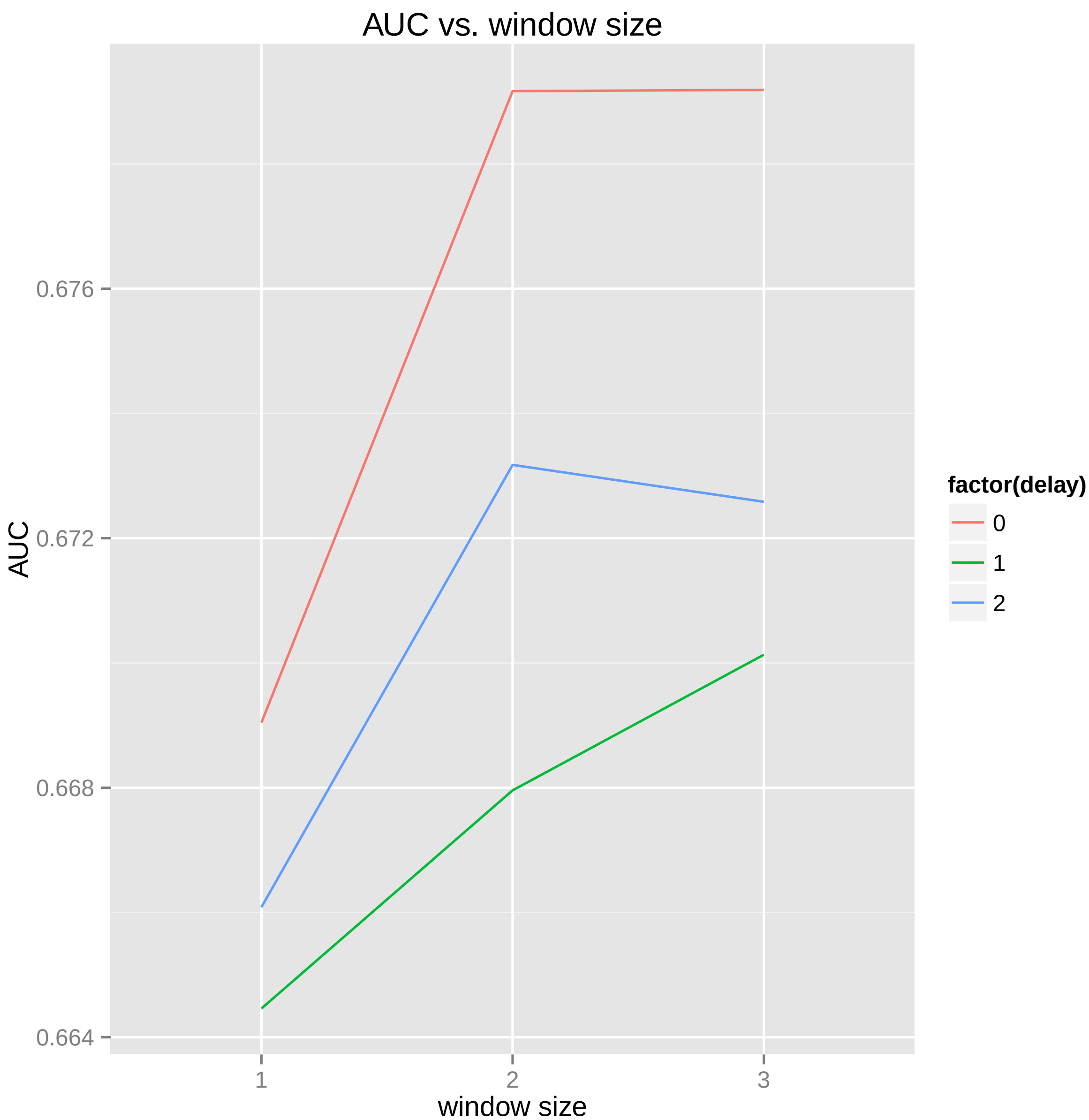


Figure 7.5. Area under the Curve (AUC) as a function of window size.

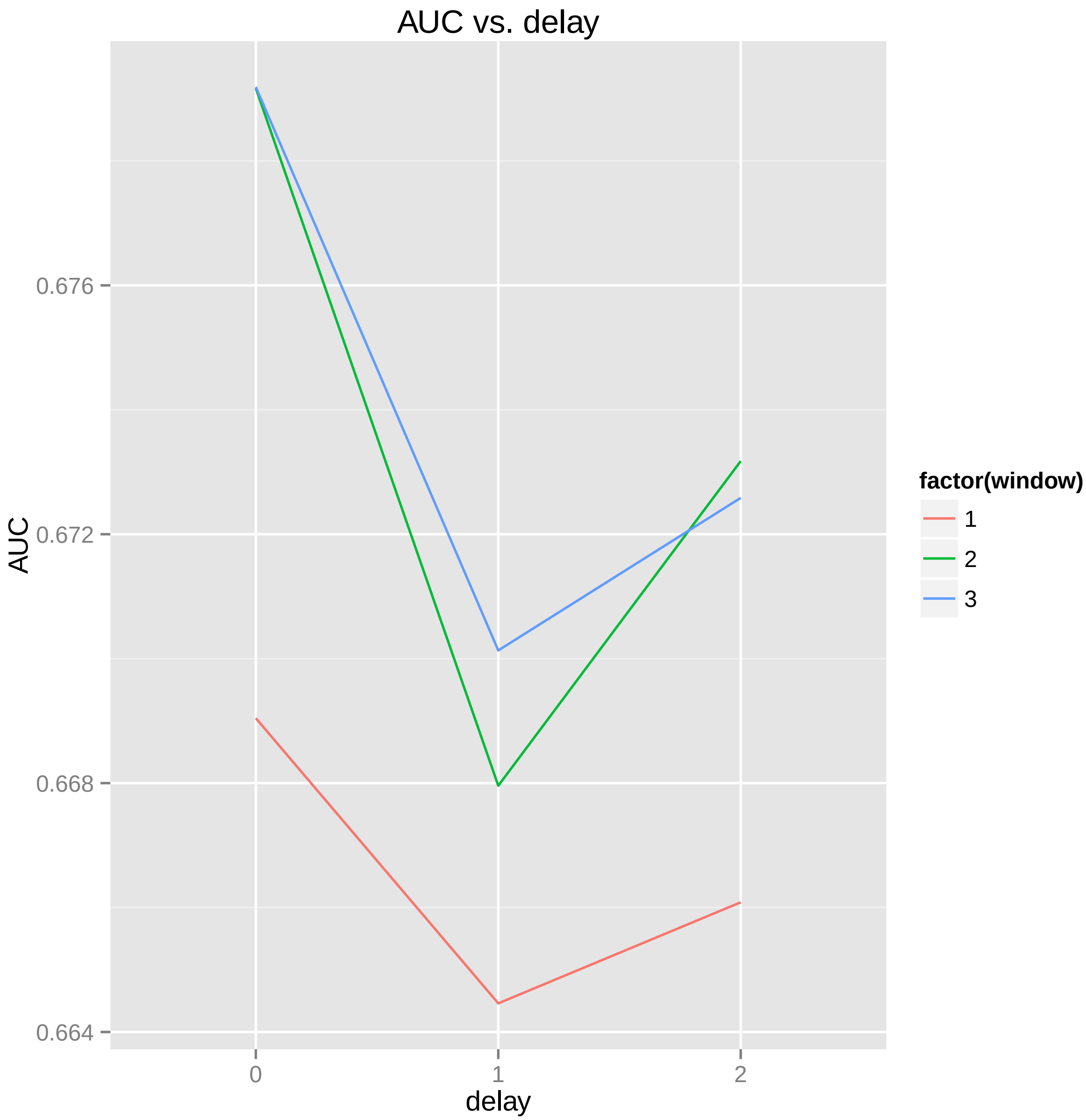


Figure 7.6. Area under the Curve (AUC) as a function of delay.

Figure 7.7. Forum structure in Game X. Each forum can have multiple topics, and each topic can have multiple posts.

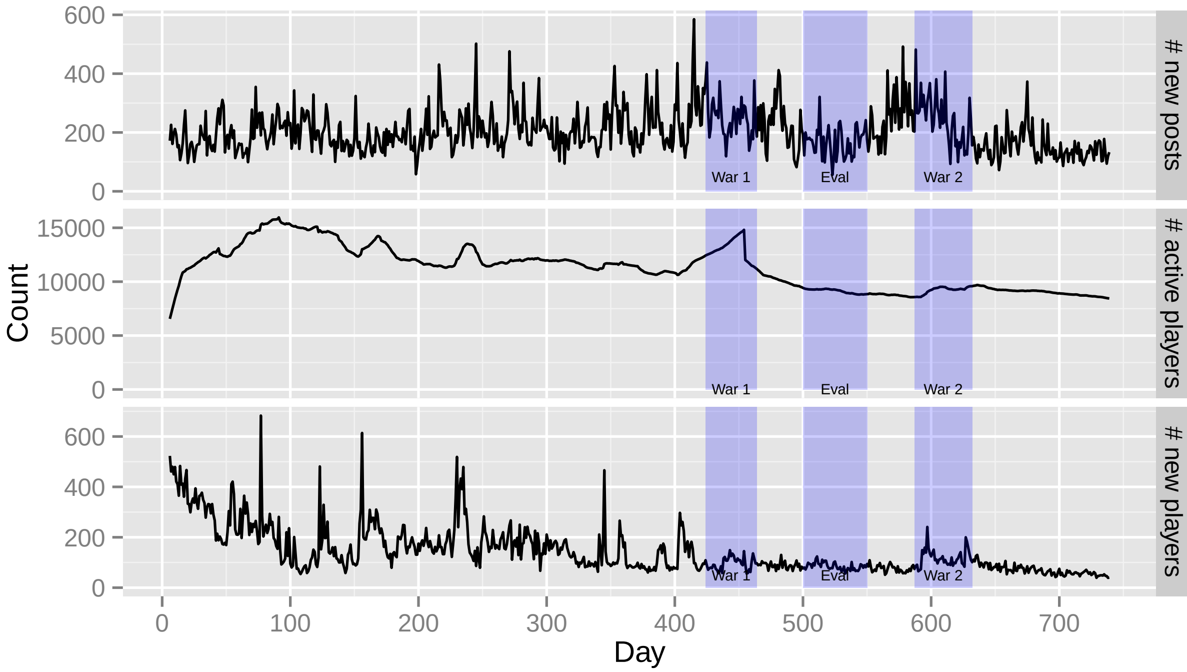


Figure 7.8. New posts per day from day 5 to day 739. Highlighted time periods indicate the two wars and the evaluation period.

Chapter 8

Public and Private Interaction

Introduction

Social media (such as Facebook, Twitter) has allowed researchers to induce large social networks from easily accessible online data. However, relationships inferred from social media data may not reflect real interaction. First, individuals who interact using social media understand (to a certain extent) that it is a public forum for communication, and hence may limit what they say. Secondly, the relations expressed may not represent the full set of relations an individual has in the "real world". One may have friends who do not use Facebook, and thus that relationship would be missing. Thirdly, extraneous relations may be present in social media that do not occur in the real world. For instance, I may "follow" the twitter account of a celebrity, but that is not a real indicator of a relationship. In many cases, what we want is the "private" social network that identifies the true relationship in real world. We can view the social network from social media as reflecting some of the relationship from the real, hidden, private network.

The main question of this work is: How does the public social network reflect the private social network? In this paper we are interested in uncovering the "private" social network that identifies strong relationship in the real world from the "public" social network. We begin to address this question by studying public and private interactions between players in a Massively Multiplayer Online Game (MMOG). Our data set contains posts by players in a Usenet-like public forum within the game, personal messages exchanged between players, and their relationship data collected from an online game *Game X*.

Our goal is to understand whether players' public interaction knowledge can help us discover the social ties among players. We hypothesize that if players communicate with each other publicly (e.g. posting on the same forum topic, referencing another player in a post, etc.), they are more likely to have a social tie such as coming from the same nation or interest group (i.e. guild).

Why Massively Multiplalyer Online Games (MMOG)?

Massively multiplayer online games attract millions of players to a shared, virtual world. While the term MMOG may encompass a variety of genres, we are interested in the large portion of games that are often labeled as "role playing games." In these, players create an avatar that represents them in the virtual world[1].

Many MMOGs are appealing for their complex economies and social structures (e.g. Framville, Second Life, World of Warcraft, Eve Online). Often games contain player-created and -controlled "guilds" or "corporations" that other players can join. These groups regularly have conflicts and interactions in the world. In some instances, long term (approximately a year of real world time) espionage has been conducted [25].

MMOGs have several advantages as a method of gathering data. First, we can gather data on a large number of people with a diverse background. MMOGs are played by millions of players, and contrary to belief, MMOGs have a wide array of player types [64]. Secondly, data gathered from MMOG are high in experimental realism [50] and reflect rich social dynamics. Many MMOG players willingly spend hours playing the game (average 22 hours per week [64]) with high commitmment to their game characters. In addition, events in games often occur faster than in the real world so one can see the rise and fall of organizations within the game. Thirdly, MMOGs provide a unique way to observe communication and behavior of players. In-game forums and messaging data can be gathered along with player action data, and this gives us one way of addressing the "radical chic" problem, by explicitly studying the correlation between communication and behavior. Finally, all data are generated in a virtual world, so most privacy concerns are minimal.

The main criticism against MMOG data is that in-game player behaviors and social patterns are not the same as those of real world[2]. This is an ongoing endeavor, and there is a literature of studies showing that in-game behaviors and social patterns do reflect real world patterns [65, 16].

Section 1 provides an overview of our MMOG. Section 8 outlines our data collection and models trained to predict social ties among game players. Section 8 discusses the results and suggests future work.

Experiments

In this section we investigate whether players' public interaction knowledge can help us discover the hidden private network among the players. To do this, we choose to train a number classifiers to predict whether two players are from the same guild. Guilds are mainly

[1] In the following we use the term "players" to refer to the avatars within the game.

[2] The "mapping principle" [61] is a term used to describe which behaviors in virtual space "map" to the real world.

created to allow members to cooperate and gain physical and economic control in the game. Unlike nation membership, guild membership is closed; one has to be approved by other members in the guild to join. For these reasons, we think guild membership could be a good criteria for uncovering the private social network among the game players.

Methods

We collected data from 50 days in the game (days 500-550) on pairs of players who have interacted either publicly or privately. This period has a relatively stable rate of posts per day, new players per day, and active players per day. For this work, we further defined three types of public interaction between two players as the following:

Co-posters: Co-posters of player p are all players who have posted in a topic that player p has also posted in.

Co-quoters: Co-quoters of player p are all players who have posted in a topic and quoted any of player p's posts.

Co-referencers: Co-referencers of player p are all players who have posted in a topic and referenced player p in the post.

For each pair of players who have publicly or privately interacted, we collected the following information, which constructs our dataset:

Membership Information (yes/no): whether two players are from the *same guild, same nation*, *same race*, or the *same agency.*

Relationship Information (yes/no): whether two players are *friends, foes*, or of the *same sex.*

Public Interaction (numerical): how many times two players have *co-posted, co-quoted*, or *co-referenced.*

Private Interaction (numerical): how many times two players have exchanged *personal messages*, or got involved in *tradings* and *combats.*

Player Proximity (numerical): *a* measure of *geographical distance* between two players.

To study whether public interaction information improves prediction of the private network, we constructed different feature sets to train the classifiers. Feature Set 1 serves as the baseline, consisting of features that are easily available to all other players in the game. Feature Set 2 adds public interaction information in addition to Feature Set 1. Feature Set 3 includes a richer set of information about the players. Our hypothesis is that models trained on Feature Set 2 will predict same guild membership significantly better than models trained on Feature Set 1 and models trained on Feature Set 3 will further improve the prediction.

Table 8.1. Paired-samples t-tests for resampling results

Feature Set	Models	Accuracy	Precision	Recall	F-Score
Feature Set 1	Linear Classifier	.694	.673	.753	.711
	Boosted Tree	.694	.673	.753	.711
	SVM	.694	.673	.753	.711
Feature Set 2	Linear Classifier	.706	.693	.740	.716
	Boosted Tree	.709	.698	.736	.717
	SVM	.724	.685	.828	.750
Feature Set 3	Linear Classifier	.739	.697	.842	.763
	Boosted Tree	.767	.721	.873	.790
	SVM	.745	.678	.934	.786

Feature Set 1 (Baseline): same-nation, same-race.

Feature Set 2: same-nation, same-race, co-posts, co-references, co-quotes.

Feature Set 3: same-nation, same-race, co-posts, co-references, co-quotes, same-sex, same-agency, are-foes, are-friends, num-private-msg, num-trades, num-combats, distance.

We trained three types of classifiers for each feature set using models available in R: generalized linear models (LM), boosted decision trees (BT), and support vector machines (SVM). The caret package [37] for R was used for pre-processing the data and tuning/training the models.

Because the original dataset is highly unbalanced, mostly consisting of data points where two players are not in the same guild (i.e. negative samples), a classifier trained on this data will achieve high performance by predicting that the sample belongs to the negative class. To avoid this, we balanced the data by randomly sampling the same number of negative samples and positive samples, which resulted in 65076 samples in total. 75% of the samples were used for training and 25% for testing. Training was done using a 5-fold cross validation process repeated 5 times.

Results

We use accuracy, precision, recall, and F-score to discuss the performance of the trained models. Table 8.1 shows the results. Comparing the results of Feature Set 1 (baseline) to Feature Set 2, we see that including public interaction information does improve predicting whether two people are from the same guild. Adding a richer set of information about the players (Feature Set 3) further improves the prediction. For Feature Set 1, further investigation on which feature had a large contribution to the training process reveals that all three models only used feature *same-nation* for training. This led to all three models producing the same accuracy, precision, recall, and F-score values. SVM had the highest

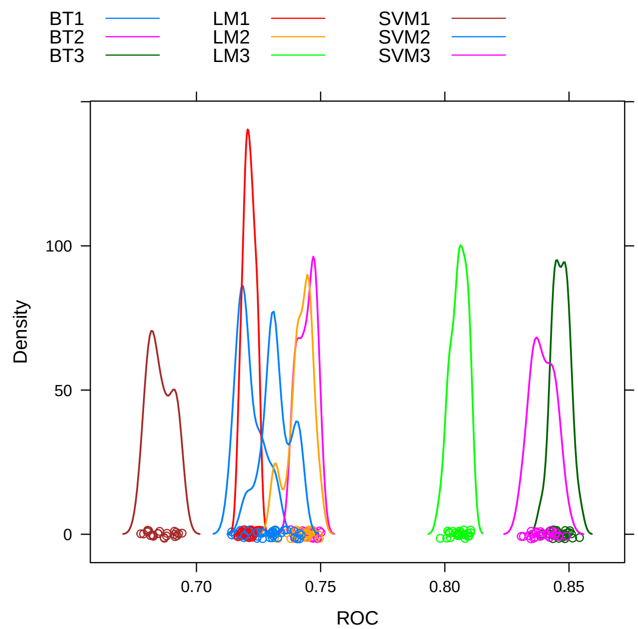


Figure 8.1. Density distribution plot of model resampling with area under ROC curves as a metric.

F-Score for Feature Set 2, however decision tree outperformed SVM for Feature Set 3. Linear models generally had the poorest performance in all three feature sets.

To further compare the models, we look at their resampling distributions with the area under ROC curves as a metric. Since we used a 5-fold cross validation method repeated 5 times, we have 25 resampling measurements for each model. Figure 8.1 plots the resampling results. A paired two-sample t-tests were also conducted to see whether these differences are statistically significant, which is shown in Table 8.2. The upper right hand side of the table shows the mean differences between a pair of models and the lower left hand side shows the p-values. The bottom three rows show that for each model type, the models trained on different feature sets are statistically different from each other (e.g. SVM1 vs. SVM2 vs. SVM3). The results shown in Table 8.1 and Table 8.2 let us conclude that using public interaction information significantly improves the prediction of whether two people are from the same guild, thus supporting our hypothesis. Results also show that using a richer set of information about the players further improves the prediction.

Conclusions and Future Work

In this chapter we address the question of whether we can use public social interaction data to uncover the private social ties using data taken from a Massively Multiplayer Online Game. We begin to answer this question by training models that classify whether two players are from the same guild using their public and private interaction data, as well as

Table 8.2. T-tests for the resampling results (LM: Linear Model, BT: Boosted Tree, SVM: Support Vector Machine).

	LM1	BT1	SVM1	LM2	BT2	SVM2	LM3	BT3	SVM3
LM1		0.000	0.036	-0.021	-0.023	-0.011	-0.084	-0.125	-0.085
BT1	1.000		0.036	-0.021	-0.023	-0.011	-0.084	-0.125	-0.085
SVM1	<2.2E-16	<2.2E-16		-0.056	-0.059	-0.047	-0.120	-0.161	-0.121
LM2	0.000	0.000	<2.2E-16		-0.003	0.010	-0.064	-0.105	-0.064
BT2	<2.2E-16	0.000	<2.2E-16	1.000		0.012	-0.061	-0.102	-0.062
SVM2	0.000	0.000	<2.2E-16	0.000	0.000		-0.073	-0.114	-0.074
LM3	<2.2E-16	<2.2E-16	<2.2E-16	<2.2E-16	<2.2E-16	<2.2E-16		-0.041	0.000
BT3	<2.2E-16	<2.2E-16	<2.2E-16	<2.2E-16	<2.2E-16	<2.2E-16	<2.2E-16		0.041
SVM3	<2.2E-16	<2.2E-16	<2.2E-16	<2.2E-16	<2.2E-16	<2.2E-16	<2.2E-16	1.000	

their relationship data. Results show that by using public interaction information, we can significantly improve the prediction of guild co-membership compared to only using easily available, extrinsic information such as nation or race association. Using the relationship data improves the prediction even more.

This work has several possible extensions. Currently we are only looking at guild co-membership. A natural extension would be to study whether we can predict other types of private interactions, such as how often they exchange personal messages or participate in trades/combats, which will further reveal the private networks among players. Another extension would be to analyze the context of players' posts instead of merely looking at the frequency to better understand the relationship among players. Finally, we may analyze player characteristics to study whether there are certain interaction patterns (e.g. leaders of nation/guild may get more incoming messages) depending on these characteristics.

Chapter 9

Conclusions

Games have been with us for millenia [1], and current data indicates games will only continue to be played. Indeed, the rise of powerful processors on mobile phones has allowed people to play games everywhere [2].

Given their immense popularity, games provide an intriguing – and vast – source of data on human behavior. In this LDRD we established the utility of game data for a variety of purposes at Sandia. It is our hope that future projects will utilize the game data we have gathered, and also gather new data.

There are many open questions and much further work to be done. It is clear that not all behavior translates from the real-world to the game world and vice-verso. The anonymity of players within the game world allows them to take on roles, for instance, that may not reflect their "real world" nature. Actions taken within games (such as attacking another player) may never occur within a real-world context.

We must be careful in how we translate results from the gaming context to the real world. Rather than thinking of mapping between behaviors, it may be more useful to think about the underlying motivation and issues, with the specific behavior a product of the domain. That is, what motivates someone to take an action with potentially long term harmful impacts to oneself may be the same as in the game and the real world. In the game, this action may be to attack a player, in the real-world this may be to send a hurtful email. The actions may differ, but the underlying factors that induce such behavior may not.

If this hypothesis is correct, games can provide a vast and detailed source of data on human behavior that could transform Sandia's ability to develop and validate Agent Based Models. The overall goal of this LDRD was to evaluate the potential impact of games within Sandia. We hope that the results described above lead the reader to the conclusion we see: gamex can transform Sandia's ability to construct and evaluate ABMs, thus allowing Sandia to anticipate national security threats.

[1] One of the oldest known board game is called *Senet* and was played in Egypt around 3500 B.C.

[2] Games are the largest segment of mobile applications and account for 17% of all apps in the Apple App Store, and 15% of all apps in the Google Play store, `http://www.portioresearch.com/en/blog/what-apps-are-people-using.aspx`

References

[1] Cosmopolis. http://www.casos.cs.cmu.edu/projects/project.php?ID=7&Name=MMOG%20-%20Cosmopolis, 2012.

[2] MMOG data. http://users.telenet.be/mmodata/Charts/TotalSubs.png, 2012.

[3] Alessandro Acquisti, Maarten Sierhuis, William J Clancey, and Jeffrey M Bradshaw. Agent based modeling of collaboration and work practices onboard the international space station. In *Proceedings of the Eleventh Conference on Computer-Generated Forces and Behavior Representation*, volume 8, pages 315–337, 2002.

[4] Icek Ajzen. The theory of planned behavior. *Organizational behavior and human decision processes*, 50(2):179–211, 1991.

[5] T. Alves and L. Roque. Using Value Nets to Map Emerging Business Models in Massively Multiplayer Online Games. In *Proceedings of the Pacific Asia Conference on Information Systems*, 2005.

[6] S. E. Asch. Effects of group pressure upon the modification and distortion of judgments. In H. Guetzkow, editor, *Groups, Leadership, and Men*, pages 177–190. Carnegie Press, Pittsburgh, PA, 1951.

[7] G. Backus, M. Bernard, S. Verzi, A. Bier, and M. Glickman. Foundations to the unified psycho-cognitive engine. Technical Report SAND2010-6974, Sandia National Laboratories, 2010.

[8] Eytan Bakshy, Dean Eckles, Rong Yan, and Itamar Rosenn. Social influence in social advertising: Evidence from field experiments. *SSRN eLibrary*, June 2012.

[9] Osman Balci. Validation, verification, and testing techniques throughout the life cycle of a simulation study. *Annals of Operations Research*, 53:121–173, 1994.

[10] Olivier Barreteau, Franois Bousquet, and Jean-Marie Attonaty. Role-playing games for opening the black box of multi-agent systems: method and lessons of its application to senegal river valley irrigated systems. *Journal of Artificial Societies and Social Simulation*, 4(2), 2001.

[11] Gnana K Bharathy and Barry Silverman. Validating agent based social systems models. In *Proceedings of the Winter Simulation Conference*, pages 441–453. Winter Simulation Conference, 2010.

[12] J. C. Bohorquez, S. Gourley, A. R. Dixon, M. Spagat, and N. F. Johnson. Common Ecology Quantifies Human Insurgency. *Nature*, 462(7275):911–914, 2009.

[13] Seraphina Brennan. Rumor: Band of brothers breaks apart in eve, goonswarm responsible, February 2009.

[14] Kathleen M. Carley. Validating computational models. Technical report, Carnegie Mellon University, 1996.

[15] Kathleen M. Carley. Validating computational models. Technical report, Carnegie Mellon University, 1996.

[16] Edward Castronova, Dmitri Williams, Cuihua Shen, Rabindra Ratan, Li Xiong, Yun Huang, and Brian Keegan. As real as real? macroeconomic behavior in a large-scale virtual world. *New Media and Society*, 11(5):685–707, 2009.

[17] Munmun De Choudhury, Michael Gamon, Scott Counts, and Eric Horvitz. Predicting depression via social media. In *Proceedings of the 7th International AAAI Conference on Weblogs and Social Media (ICSWM 2013)*, 2013.

[18] A. Clauset, C. R. Shalizi, and M. E. J. Newman. Power-law Distributions in Empirical Data. *Arxiv preprint arxiv:0706.1062*, 2007.

[19] Jeffrey S. Dean, George J. Gumerman, Joshua M. Epstein, Robert Axtell, Alan C. Swedlund, Miles T. Parker, and Steven McCarroll. Understanding anasazi culture change through agent-based modeling. In *Dynamics in Human and Primate Societies: Agent-Based Modeling of Social and Spatial Processes*, pages 179–205. Oxford University Press, New York & London, 2000.

[20] N. Ducheneaut, N. Yee, E. Nickell, and R. J. Moore. Alone together?: Exploring the Social Dynamics of Massively Multiplayer Online Games. In 407–416, editor, *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2006.

[21] Giorgio Fagiolo, Alessio Moneta, and Paul Windrum. A critical guide to empirical validation of agent-based models in economics: Methodologies, procedures, and open problems. *Computational Economics*, 30(3):195–226, 2007.

[22] Jay W. Forrester and Peter M. Senge. Tests for building confidence in system dynamics models. *TIMS Studies in Management Sciences*, 14:209–228, 1980.

[23] Malcolm R. Forster. Key concepts in model selection: Performance and generalizability. *Journal of Mathematical Psychology*, 44:205–231, 2000.

[24] Donelson R. Forsyth. *Group Dynamics*. Wadsworth Cengage Learning, 6 edition, 2014.

[25] Tom Francis. Murder incorporated. *PC Gamer Magazine*, page 90, 2006.

[26] Eric Gilbert. Phrases that signal workplace hierarchy. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*, CSCW '12, pages 1037–1046, New York, NY, USA, 2012. ACM.

[27] Eric Gilbert and Karrie Karahalios. Predicting tie strength with social media. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, pages 211–220, New York, NY, USA, 2009. ACM.

[28] Martin Greenberger, Matthew A. Crenson, and Brian L. Crissey. *Models in the Policy Process: Public Decision Making in the Computer Era*. Russel Sage Foundation, New York, 1976.

[29] Paul Guyot and Shinichi Honiden. Agent-Based participatory simulations: Merging Multi-Agent systems and Role-Playing games. *Journal of Artificial Societies and Social Simulation*, 9(4), 2006.

[30] Mohammad Hashemian, Dylan Knowles, Jonathan Calver, Weicheng Qian, Michael. C. Bullock, Scott Bell, Regan L. Mandryk, Nathaniel Osgood, and Kevin G. Stanley. iEpi: an end to end solution for collecting, conditioning and utilizing epidemiologically relevant data. In *Proceedings of the 2nd ACM international workshop on Pervasive Wireless Healthcare*, MobileHealth '12, pages 3–8, New York, NY, USA, 2012. ACM.

[31] Macartan Humphreys and Jeremy M. Weinstein. Who Fights? The Determinants of Participation in Civil War. *American Journal of Political Science*, 52(2):436–455, 2008.

[32] W. Jager, M.A. Janssen, H.J.M. De Vries, J. De Greef, and C.A.J. Vlek. Behaviour in commons dilemmas: Homo economicus and homo psychologicus in an ecological-economic model. *Ecological Economics*, 35:357–379, 2000.

[33] Marco A. Janssen and Elinor Ostrom. Empirically based, agent-based models. *Ecology and Society*, 11(2), 2006.

[34] Luke Karmali. World of warcraft down to 7.7 million subscribers, July 2013.

[35] Leonard Kaufman and Peter J. Rousseeuw. *Finding Groups in Data: An Introduction to Cluster Analysis*. Wiley-Interscience, 1 edition, March 2005.

[36] Jack P.C. Kleijnen. Verification and validation of simulation models. *European Journal of Operational Research*, 82:145–162, 1995.

[37] Max Kuhn. Building predictive models in r using the caret package. *Journal of Statistical Software*, 28(5):1–26, 2008.

[38] Charles Kurzman. Why is it so hard to find a suicide bomber these days?, 2011.

[39] Xiaochen Li, Wenji Mao, Daniel Zeng, and Fei-Yue Wang. Agent-based social simulation and modeling in social computing. In Christopher Yang, Hsinchun Chen, Michael Chau, Kuiyu Chang, Sheau-Dong Lang, Patrick Chen, Raymond Hsieh, Daniel Zeng, Fei-Yue Wang, Kathleen Carley, Wenji Mao, and Justin Zhan, editors, *Intelligence and Security Informatics*, volume 5075 of *Lecture Notes in Computer Science*, pages 401–412. Springer Berlin / Heidelberg, 2008.

[40] Eric T Lofgren and Nina H Fefferman. The untapped potential of virtual game worlds to shed light on real world epidemics. *The Lancet infectious diseases*, 7(9):625–629, 2007.

[41] Neil Long. Eve online passes 500,000 subscribers, February 2013.

[42] Anmol Madan and Alex (Sandy) Pentland. Modeling social diffusion phenomena using reality mining. In *Proceedings of the 2009 AAAI Spring Symposium on Human Behavior Modeling*, 2009.

[43] Martin Maechler, Peter Rousseeuw, Anja Struyf, Mia Hubert, and Kurt Hornik. *cluster: Cluster Analysis Basics and Extensions*, 2012.

[44] Daniel McFadden. Qualitative response models. In Werner Hildebrand, editor, *Advances in Econometrics*. Cambridge University Press, New York, 1982.

[45] Laura A. McNamara, Timothy G. Trucano, George A. Backus, Scott A. Mitchell, and Alexander Slepoy. R&D for computational cognitive and social models: Foundations for model evaluation through verification and validation. Technical Report SAND2008-6453, Sandia National Laboratoriesi, 2008.

[46] Scott Moss and Bruce Edmonds. Sociology and simulation: Statistical and qualitative cross-validation. *American Journal of Sociology*, 110:1095–1131, 2005.

[47] Scott Moss and Bruce Edmonds. Sociology and simulation: Statistical and qualitative Cross-Validation. *American Journal of Sociology*, 110:1095–1131, 2005.

[48] William L. Oberkampf, Timothy G. Trucano, and Charles Hirsch. Verification, validation, and predictive capability in computational engineering and physics. *Applied Mechanics Reviews*, 57:345–384, 2004.

[49] S. Papagiannidis, M. Bourlakis, and F. Li. Making Real Money in Virtual Worlds: MMORPGs and Emerging Business Opportunities, Challegens and Ethical Implications in Metaverses. *Technological Forecasting and Social Change*, 75(5):610–622, 2008.

[50] Harry T. Reis and Charles M. Judd. *Handbook of research methods in social and personality psychology*. Cambridge University Press, 2000.

[51] David O. Sears. College sophmores in the laboratory: Influence of a narrow data base on social psychology's view of human nature. *Journal of Personality and Social Psychology*, 51(3):515–530, 1986.

[52] C. A. Steinkuehler. Learning in Massively Mulitplayer Online Games. In *Proceedings of the 6th International Conference on Learning Sciences*, pages 521–528, 2004.

[53] C. A. Steinkuehler. Cognition and Literacy in Massively Multiplayer Online Games. In *Handbook of Research on New Literacies*, pages 1–38. Mahwah NJ: Erlbaum, 2008.

[54] Catherine A. Sugar, Gareth, and M. James. Finding the number of clusters in a data set: An information theoretic approach. *Journal of the American Statistical Association*, 98:750–763, 2003.

[55] Ron Sun, editor. *Cognition and Multi-Agent Interaction: From Cognitive Modeling to Social Simulation*. Cambridge University Press, December 2005.

[56] M. Szell and S. Thurner. Measuring Social Dynamics in a Massively Multiplayer Online Games. *Social Networks*, 32(4):313–329, 2010.

[57] Pang-Ning Tan, Michael Steinbach, and Vipin Kumar. *Introduction to Data Mining*, chapter Cluster Analysis: Basic Concepts and Algorithms. Addison Wesley, 2005.

[58] Yla R. Tausczik and James W. Pennebaker. The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, 29(1):24–54, March 2010.

[59] M. Taylor. *Rationality and Revolution*. Cambridge University Press, 1988.

[60] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2012.

[61] Dmitri Williams. The mapping principle, and a research framework for virtual worlds. *Communication Theory*, 20(4):451–470, November 2010.

[62] Timothy D. Wilson, Elliot Aronson, and Kevin Carlsmith. The art of laboratory experimentation. In Susan T. Fiske and Daniel T. Gilbert andGardner Lindzey, editors, *The Handbook of Social Psychology*, volume 1. John Wiley & Sons, 5 edition, 2010.

[63] Paul Windrum, Giorgio Fagiolo, and Alessio Moneta. Empirical validation of agent-based models: Alternatives and prospects. *Journal of Artificial Societies and Social Simulation*, 10(2), 2007.

[64] Nick Yee. The demographics, motivations, and derived experiences of users of massively multi-user online graphical environments. *Presence: Teleoperators and Virtual Environments*, 15(3):309–329, June 2006.

[65] Nick Yee, Nicolas Ducheneaut, Les Nelson, and Peter LIkarish. Introverted elves & conscientious gnomes: The expression of personality in world of warcraft. In *Proceedings of CHI 2011*, 2011.

DISTRIBUTION:

1	MS	1004	J. Whetzel, 6134
1	MS	1161	T. Stalker, 5447
1	MS	1188	A. Nanco, 6114
1	MS	1327	P. Bennett, 1463
1	MS	1327	A. Bier, 1462
1	MS	1327	K. Lakkaraju, 1462
1	MS	1327	J. Wagner, 1462
1	MS	9407	J. Lee, 8112
			,
1	MS	0899	Technical Library, 9536 (electronic copy)
1	MS	0359	D. Chavez, LDRD Office, 1911

Sandia National Laboratories