Online Play, Online Connection:
A Longitudinal Social Network Analysis of BZFlag

John C. Paolillo and Daniel O. Kutz
School of Library and Information Science
Indiana University
Bloomington, Indiana
{paolillo, dokutz}@indiana.edu

Abstract

In this paper we investigate individual and collective activity patterns in BZFlag, an open-source, online, multi-player, 3D tank-battle game, with tens of thousands of active users worldwide. We present analysis of nearly two full years' worth of this data, sampled at five minute intervals. The results of this analysis reveal a highly-skewed distribution of activity among players, where the vast majority of players have relatively limited activity on the servers. A distinct minority of players log a large number of playing hours, and exhibits a highly-structured pattern of interaction. In the aggregate picture, there is highly robust periodic activity on BZFlag at several levels. On the local scale, daily patterns predominate, with significant half-day, 2-hour and shorter periodic patterns. On the global scale, equally robust long-term periodic fluctuation is observable. Yearly patterns, clearly pertaining to the ebb and flow of the North American school year, are strongly evident. The results of this study underscore the importance of the time dimension in understanding patterns of social activity in online games, and in understanding social relationships more generally.

1. Introduction

Online multiplayer games, though a relatively new entertainment form, are fast becoming a primary mode of interaction for many people in which meaningful social relationships are formed and maintained. The inherent social nature and the resultant bonds formed are attributes that make online gameplay inherently appealing (Steinkuehler and Williams, 2006; Williams et. al., 2006). Yet, recent work on the social nature of gameplay focuses on interaction within the game, typically for a small subset of the game's players. For example Jakobsson and Taylor (2003) and Ducheneaut, et al. (2006) investigate how one's social network (e.g. guild) is used to address gameplay challenges in massively multiplayer games, while Steinkuehler and Williams (2006) address how online gameplay as a "third place" can help establish social bridging. Less research has been focused on the large scale social and temporal patterns of participation in online games. How does participation in an online game unfold over time, and how does that participation influence an individual's social standing among other participants? To what extent and for whom do online social activities translate into meaningful social relations? What group-level patterns exist in the collective activity, and how do they relate to individual patterns of participation?

1.1. Social networks

Online interactions can have an important role in a player's social circle. This is especially the case when people's opportunities for face-to-face contact are limited, or when teens are confined to home after school, such that electronic forms of social connection assume greater importance than they otherwise might. Furthermore, when players from around the world gather together in online spaces to conduct collaborative quests, coordinate collective actions, and organize themselves for inducting new members and maintaining group cohesion for future activities, then social norms are modeled and enforced, and interactional strategies are transmitted in proportion to the time people are individually and collectively engaged.

Social Network Analysis (SNA) is the area of sociological inquiry that most directly addresses these issues; the basic unit of analysis in SNA is that of the social relation or "tie" between individuals (Degenne and Forsé 1999;

Wassserman and Faust 1994). Ties may be of any type and still be suitable for social network analysis: kinship, marriage, friendship, economic transactions, etc. are all different kinds of ties investigated and contrasted in SNA. The choice of tie determines the nature of the analysis, and the resulting range of interpretations. In all cases, the analysis results in a structural representation in which a number of community and individual characteristics can be read. Chief among these is centrality: actors that have more ties to other actors are more central. Centrality is often associated with greater power, and the expression and enforcement values within the community, whereas peripheral members of the community, with fewer ties to those inside, potentially have more exposure to outside-community influences by functioning as bridges to other networks (Degenne and Forsé 1991).

A common correlate of centrality is tie strength, first discussed in Granovetter (1973). The strength of a tie is directly proportional to the duration, frequency, and intensity of contact shared between two actors. Two people with longer, more frequent and more intense social contact, such as close friends and family members, share a stronger tie than two who do not, such as casual acquaintances and people who only meet incidentally through the work setting. Moreover, duration, frequency and intensity of contact are often correlated, so that one need not always measure all three to determine tie strength. Strong ties also tend to be transitive, because people have limited time with which to engage in social life. If two people have a strong tie with one another, and one of those has a strong tie to a third, then it is highly likely that the other has a strong tie to the third, simply because the time spent together with the first person tends to bring them in contact with one another. Tie strength itself further correlates with the types of resources and opportunities available to people. In particular, people with larger numbers of weak ties may have access to a wider range of resources such as job opportunities, whereas those with more strong ties may benefit more from support within their local community and family networks.

Since its introduction in Granovetter (1973), tie strength has been used to explain a variety of different social phenomena, from resource mobilization (Granovetter 1974) to linguistic variation (Milroy and Milroy 1992). Most research infers tie strength indirectly from social connection measured by other means, such as interviews or questionnaires. Few studies actually examine the role of frequency and duration of contact in structuring social relations more directly. In large part, the reason is methodological: it is hard to observe enough of even one actor's social activity to be able to make the appropriate inferences to tie strength. Recent technological innovations, such as the spread of mobile phones, make such observations possible today (Eagle and Pentland 2006); online communication environments, such as online games, also facilitate similar observations. Online games specifically contribute the potential to understand social aspects of the "ludisphere", and how it affects our social life more generally.

1.2. Social networks and gameplay

When players discuss online gameplay they often identify temporal factors as an influence on their experience. Experienced players complain that early afternoon is a bad time to play because younger, less mature players get out of school and start playing at that time. A player of a real-time strategy game may not be able to find enough skilled participants to play during summer months because "it is summer and everyone is outside enjoying the weather." Or a raid leader in a massively multiplayer online game might try to organize a group activity schedule that is opportune for the majority of guild members so that an uninterrupted block of time can be committed to gameplay. These considerations highlight both the social and temporal aspects of online games.

Research on the formation and maintenance of social networks in gameplay has focused on the formation of social connections within a game as a subset of the gameplay community such as a guild or player groups. For example Nardi (2006) examines social organization and collaboration within a game as a process of learning, collaboration, and community building. Steinkuhler and Williams (2006) use discourse analysis to examine communication among players in order to understand if massively multiplayer online games can serve as a "third place" (Oldenburg, 1989) and provide the opportunity for the creation of weak and strong social ties. Jakobsson and Taylor (2003) identify that social networks play an important role in how players engage game content. Among their findings, players will switch their "characters" in order to suit the needs of the party. These role changes are dependent on the establishment of awareness of existing social ties between players and their multiple character personas.

Kolo and Baur (2001) look at the larger demographics of gameplay in order to get an understanding of who is participating. The authors conduct a survey among Ultima Online players in order to obtain demographic information, the frequency and duration of gameplay, what time players are online, and the type of gameplay they engage in.

Their results show that the social experience is one of the most appealing attributes of gameplay. Investigating the demographics of players and motivations of play, Yee has been running the Daedalus Project (http://www.nickyee.com/daedalus) for several years, which documents via surveys player motivations and demographics of those who participate in massively multiplayer online role-playing games (MMORPG). As part of a long-term project, Seay et. al. (2003), use an online survey to understand how player groups develop and move between various games. To better understand the behavior of "hardcore" (those that have an above average commitment to the game) and "casual" (those that see gameplay as a leisure activity) players, Fritsch et. al. (2006) survey the play styles of real time strategy, first person shooter, roll playing games, and sports games players. Results from their study show significant differences in hours per week dedicated to gameplay, with role playing games receiving the largest amount of gameplay (32.8 hours), sports games the least (11.2 hours), and first person shooters and real time strategy game players spending 14.9 and 18.3 hours respectively.

Researchers have used automated means as well as self-report data to obtain information about the social and temporal aspects of online games. Chen et. al. (2006), looking to improve on game design, conducted a packet trace and monitored network activity of ShenZhou Online to identify large-scale player interactions. Their findings show that teammates stay closer together, that there is a heavy-tailed distribution of players across the space, and that players engaging socially also spend more time in the game. By using automated scripts, Ducheneaut et. al. (2006) capture basic in game demographic information in an attempt to document patterns of gameplay over an 8 month period in World of Warcraft, a massively multiplayer online game. Their scripts track player metrics in an attempt to identify the progression of a player's gaming experience. Results show that as players invest more time into their character (identified by their "level"), they are also more likely to spend more hours playing the game. In a similar study, Williams et. al. (2006) attempt to identify the size and nature of player guilds in a massively multiplayer online game. Through participant observation, and by tracking the grouping patterns of players, the authors are able to establish a typology of guilds, identify the most central figures, the establishment of sub-groups within the guild, and the rate at which players leave and join a guild or a game (Williams et. al. 2006).

Gameplay research that focuses on the temporal dimension tends to examine the role of time within a game itself. Klastrup (2006) examines how the implementation of "death" in a game (and its associated time penalty) can shape the social structure of game. Lindley (2005), taking a semiotic approach, examines the different roles that time can take in a game, simulation, and narrative. Other scholars examine the connection between game time and realtime. For example, Juul (2004) provides a time model that attempts to capture the interplay between game time, real-time and the player's experience of time. Time-and space within a game are also often closely linked together, and can serve as a throttling mechanism by designers to create impressions of size (for example by limiting movement speeds within an game environment), or temporal and spatial attributes of the environment are used by players to distinguish themselves from others (see, e.g., Schroeder 1997).

However there is also a temporal component outside of the game environment that influences gameplay. Real world events such as school schedules, work hours, vacations, holidays, all influence who engages in online gameplay and at what time. Within the context of a game itself, content may become available at a certain time, being consumed for as long as it remains novel, interesting or challenging to players. Content can also become unavailable as servers or various parts of the network undergo outages, or as software versions are upgraded. Players may go online only to find their favorite hangouts unavailable, and then either look for a place to re-establish their prior social contacts, or look for a new social milieu entirely. Cheaters may also succeed in disrupting gameplay, and cause players to find other venues for gameplay. Methods are needed that allow us to make precise observations of this nature, and to determine the nature of the effect of internal and external events on gameplay.

2. BZFlag

To address the role of large-scale temporal patterns and social ties in gameplay, we examine BZFlag, a graphical, online multiplayer open source capture-the-flag type game in which players navigate tanks across a playing field. BZFlag (Battle zone capture the Flag) is an online multiplayer "capture the flag" type game. It is a graphical game, in which the player occupies a first person perspective, similar to a first person shooter, except that the player drives a tank, instead of inhabiting the role of a humanoid avatar. The mechanics of BZFlag otherwise follow those of a first person shooter. Players connect to one of several servers and engage in team-based play on a map. A server can host a single map (often the case for the more popular servers), or it can rotate through a collection of maps. On average, a running game is made up of two teams with 16 players per team. However the game engine can support up to 200 players, and other forms of gameplay, such as free for all matches. A single shot kills a player, but death is temporary and a player can "respawn" after a short delay (a few seconds). Players normally respawn at a random location on the map with a new tank at full health. As there is minimal time penalty for death and no permanent elimination during gameplay, BZFlag matches result in a fast-paced gaming experience, with players quickly rejoining combat after respawning. Successful gameplay primarily depends on the player having a good spatial awareness of the map layout, taking advantage of "superflags" (powerups) that provide special abilities, and coordinating with teammates. BZFlag also has gameplay leagues, which is the preferred form of engagement for experienced players. In some leagues, superflags are not enabled, and team coordination and spatial awareness become the defining factors as players test their playing skills against each other.

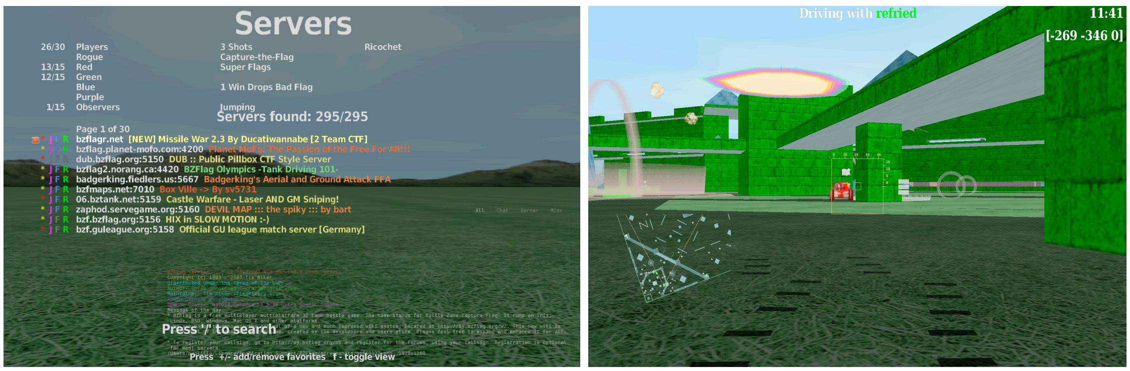

Figure 1. Screenshots of the server selection screen (left), and the game interface (right).

While many commercial first person shooters provide players with various customization features, BZFlag is unusual in that it is a fully open source multi-platform game. Gameplay is free, and participants can contribute not only new maps but also view and submit changes to the source code. As a result, BZFlag has been under constant revision and evolution since its initial release in 1993. A player has thousands of maps and hundreds of servers to choose from. Logging into BZFlag one can choose among more than 200 servers to participate on. However, on average there are between ten and twenty servers that are the most popular and have a majority of the players present at one time. Though the game allows anonymous play, servers frequently require players to be authenticated with a central server, which also provides users with a login identity for the official BZFlag forums, on which there currently are more than 11,000 registered players. The open source nature of BZFlag has its negative aspects as well: cheaters can readily modify the client source code to give themselves an advantage, leading to a number of ongoing conflicts between cheaters, server hosters and admins.

BZFlag has a very active and dynamic community with players from all over the world. Many of these players eventually become involved in community maintenance in various other ways, as administrators on popular servers, creators of new maps to be consumed by the community, server hosters, maintainers of archived games, gatherers of gameplay statistics and developers of the open-source software that supports the community. As the resources required for involvement are relatively low, individual players' commitments to the activities maintaining BZFlag also tend to be low, particularly among those new to hosting servers. In the past year, three major hosters of BZFlag servers, pythonian, nn, and bzfx, have all undergone major disruptions in service. Pythonian was transferred from one player to another, and eventually died because of disputes between the hoster and the BZFlag developers over issues of map authorship credit and control of cheaters; nn began his service by offering to host anyone's map and make the author an admin, but later ran into similar problems after flooding the central BZFlag server list with hundreds of empty servers. Later, after moving his service from a home broadband connection to a commercial-grade connection, he transferred his service to another player and eventually it died out. Out of these three, only bzfx retains a modest presence among BZFlag servers today, by currently staging a second comeback after twice losing its hosting one and two years ago, due to disputes over services borrowed from other BZFlag hosters.

These server-related events point to an important aspect of BZFlag: its structure is essentially a federation of players, map creators, hosters, server admins, and developers, whose individual activities are governed by alliances and cooperative agreements with others, and whose total collective activity defines the social environment. As the relationships and agreements change, the environment is re-structured, and ripple-effects can be felt throughout the entire community of players. All of these relationships are more or less visible via the different forms of electronic communication that constitute them: whether on the BZFlag forums, League and team websites, fan sites, the automatically archived developer IRC channel, or the game servers themselves. The availability of comprehensive logs of some of these communications invites us to study them to learn about the temporal dynamics of social processes in game-space.

For this study we were provided access to 665 days of player connection activity recorded at five minute intervals from October 27, 2006 to August 22, 2008. The data are those behind the fan site bzstats.strayer.de, run by Andreas Püschel, who provided us with several custom scripts that allow us to query his database. In the data, server records contain the server name, its IP address, and game parameters such as game mode, maximum number of players, and physics settings. Player session records contain the player's callsign (nickname), the connection time and date (beginning and ending), the server's ID, the player's team membership, and game score. From this data, we are able to compare connection profiles for individual players and servers, and to construct longitudinal social network maps of connection among different users.

3. Methodology

SNA most typically addresses one-mode data, such as friendship, kinship, etc. in which the nodes at either end of the relationship are the same type (e.g. people). Such data has two representations: as a sociomatrix, or as a sociogram (Degenne and Forsé 1999). A sociomatrix is a square matrix in which actors are the rows and columns of the matrix, and a tie or its absence, from one actor to another, is indicated by a 1 or 0 in the appropriate cell (weighted forms are also available, where values other than 0 or 1 appear in a cell). A sociomatrix generally represents a directed relation, i.e. one in which the two participants in a relation have different roles (e.g. parent-child in one kind of kinship relation). Relations can also be undirected (e.g. "mutual friend"), in which case the sociomatrix representation is symmetric about the diagonal. While sociomatrix representations are useful for mathematical and statistical manipulation, sociograms, also known as "social network diagrams", are more commonly presented for interpretation. In a typical sociogram, individuals are represented as nodes, and the presence of a tie is represented by a line between two different nodes. Different layout algorithms are used to try to situate nodes that share a lot of mutual connections near one another; Fruchterman-Reingold and Kamada-Kawai are probably the two most popular layouts used in SNA today. Other layouts are based on spectral decomposition techniques such as Principal Components Analysis (PCA) or Multi-Dimensional Scaling (MDS) of the sociomatrix or related matrices (Seary and Richards 2003).

While one-mode data is the norm in SNA, it is often necessary to work with two-mode or "actor-event" data, in which different types of node occupy each end of the relation. Examples of this latter type of data are memberships on the boards of directors of different companies, voting records from Supreme Court cases and congressional legislation, and participation of individuals in community events. Under the assumption that two actors sharing participation in the same event is equivalent to the same actors sharing a tie of some sort, two-mode data is readily converted into one-mode data via matrix multiplication: if the actor-event matrix is B, where rows are actors and columns are events, the actor-actor matrix $A = BB^T$, and the event-event matrix $E = B^TB$ (Degenne and Forsé 1999; Wasserman and Faust 1994). Investigation of the temporal dimensions of social contact further complicates this situation by introducing a third mode: time. Unlike two-mode data, there is no general way of handling three-mode data. Moreover, time as the third mode has special properties that require careful treatment.

Our approach to this problem was to treat the time mode as primary, and to express the other two modes in terms of their relation to time. This has the effect of bringing temporal patterns into focus more clearly than they otherwise might be. A potential downside of this approach is that we cannot independently estimate the effects of the three modes, and so we cannot evaluate the relative contributions of e.g. server and player patterns to those of the temporal patterns we observe. However, given the size of the data set (nearly a million time slices, tens of thousands of distinct players and thousands of distinct servers), there are serious problems of tractability in attempting to handle even one mode data (e.g. player-player data); to the extent that our approach makes temporal patterns visible, while permitting us to draw social network maps of player relationship, this loss is not serious. A second aspect of our approach is that, because we anticipate that there will be regularities in the time-dimension, we can aggregate over different periods, e.g. days or 5-minute sessions, depending on the desired periods, to reveal these patterns. Judicious choice of the periods to investigate is necessary to provide the optimal analysis.

These considerations led us to arrange our data in the following way. To investigate day-scale and shorter activity patterns, we created a large incidence matrix in which the rows were player-server combinations, i.e. instances of a specific player being logged into a specific server, and the columns were individual 5-minute time slices. When a particular player is logged in to a server in a particular time slice on a particular day, a "1" is recorded in the corresponding cell. If a player is not logged into that server in a given time-slice, a "0" is recorded. If a player is never logged into a given server on a given day, then the corresponding player-server row is simply omitted (i.e. we drop all rows containing only 0). The matrix is constructed of the row-wise concatenation of all the active player-server combinations for each day, for all 665 days in our sample. For longer-scale patterns, we reversed the role of days and time slices, so that the columns were the 665 days of the sample, and player-server combinations were tallied, one time slice at a time, for all 288 slices in a day.

Both arrangements permit us to conduct a PCA, via computing a Singular Value Decomposition (SVD) of the respective correlation matrix (Basilevsky 1994). Normally, the matrices described above would be analyzed by PCA by computing SVD directly on a z-score normalized matrix Z. SVD provides the relation $Z = P\Lambda Q^T$, where P and Q are orthonormal matrices projecting the rows and columns, respectively, of the original matrix Z into a common Principal Components (PC) space; the diagonal matrix of singular values L represents the relative contribution of the different dimensions of this space to the observed values of the data. Choosing the largest of these values, and consequently the dimensions of shared variation contributing the most to the data, results in a reduced-dimensionality space (with respect to the number of columns in the original matrix) that accounts for the greatest proportion of variation in the original data. To interpret the analysis, we generally take Q to represent loadings of the columns on the PCs, and look at scaled PC scores to understand the distribution of the rows, which we can compute as $ZQ = P\Lambda$; whichever side is easiest to compute in a given application is used. In practice both of our data matrices are too large for current software to treat this way efficiently (they would need to occupy several gigabytes of active memory). Hence, our process involved computing the correlation matrix $R = Z^TZ$ first, which could be done space-efficiently by iteratively aggregating pieces of the cross-product matrix and adjusting the values afterward. The SVD of the correlation matrix is given by $R = Q\Lambda^2Q^T$, and Q computed this way can be used to compute PC scores as ZQ.

PCA computed on the three-mode data this way allows us to identify a number of important features of BZFlag activity for our interpretation. First, since the time slices themselves represent different lags in time, PCA itself represents a Discrete Fourier Analysis (Basilevsky 1994), and the loadings of the time slices on the different PCs tell us about the periodicity and phase associated with the different PCs, while the Λ^2 magnitudes tell us about the relative sizes of those periodic patterns. Second, players and servers can both be projected into the PC space, by aggregating over the relevant player-server PC scores. From these projections, it is possible to generate SNA layouts in which a player or server is located according to an "average" association with their playing times. Player-player sociograms could thus be generated for individual days, where the ties present represent shared playing time on a server for that specific day. This approach was used to generate a social network "movie" with one frame for each of the 665 days.

As there is no off-the-shelf software that performs this analysis, we accomplished this with a combination of stock database and statistical analysis software, with specialized scripts written to fill functional gaps in the process. Database dumps from bzstats.strayer.de were formatted using scripts written in the Icon Programming Language (http://www.cs.arizona.edu/icon/), and imported into a PostgreSQL database. Statistical analysis and social network layouts were computed in R (R Core Development Team 2008), using the RdbiPgSQL package (http://www.bioconductor.org/packages/release/bioc/html/RdbiPgSQL.html) to connect with the database. Special views were composed in SQL and analysis routines were written in R to accomplish the various analysis tasks. In addition, standard statistical analysis features of R (e.g. time series analysis and linear modeling) were used at different stages of the analysis.

4. Results

The distribution of playing time logged by players over the 665 days of the sample is highly skewed, with a minority of players logging large numbers of hours. Figure 1, in which players are displayed by total playing time and rank-order on log-log coordinates, shows that this distribution has the form of a power law, with a slope of approximately -0.8; two “bulges” in the curve, however, suggest that there might be two distinct distributions here. Those above 100,000 minutes (about 69.5 days solid playing time) are characterized by highly dedicated, possibly even “addicted” players who have a significant amount of their waking life devoted to the game. A minority of these players actually represent “loggers” which are proxies set up by server owners as observers which record activity (especially chat) for later examination; the extreme values of these (in excess of 957,600 minutes = 665 days) come from loggers that run on multiple servers. Those below 100,000 minutes, peaking around 10,000 minutes (= 7 days) of playing time over the 665 day period, represent more typical playing patterns. While the vast majority of players occupy the long tail from 10,000 minutes to 5 minutes, only a small minority of players log a single 5 or ten minute session on BZFlag. Hence, the distribution is unlike that of a Zipf distribution, which has half of its probability mass or more in the lowest rank.

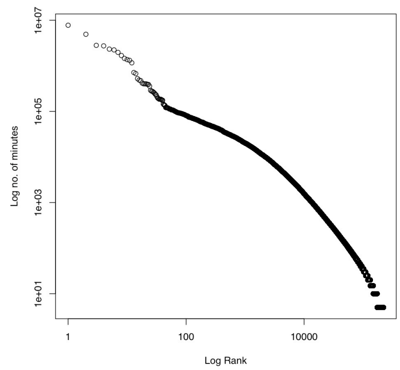

Figure 1. Rank-order distribution of total player time across 665 days in minutes, log-log scale.

This distribution is reasonably characteristic of degree distributions in a large social network. A few observations are in order, however. First, a unknown number of players will use a callsign once, whether as a joke, an attempt to remain hidden, or as a typing error, and never use it again. This means that an indeterminate number of these player identities potentially come from regular players who return but only under other aliases. Other users may or may not be aware of the shared identities across these (and even regularly returning) identities. Hence, some part of the social structure of BZFlag players may not be properly represented by the available information. At the same time, a majority of players with high playing times are indeed frequent players of the game. Since these players’ activity characterizes the predominant patterns of the game, the analysis of players’ interconnections should be meaningful.

The PCs computed from the 288 time slices indicate very robust periodic activity at the daily time-scale and less, while those computed from the 665 days indicate less robust periodic activity around an annual scale, with more noise present. Scree plots of the Λ^2 values from each of these PCAs are given in Figures 2 and 3, alongside plots of the loadings on the first several PCs for each time slot or day, respectively, indicating the nature of the periodic activity in each analysis.

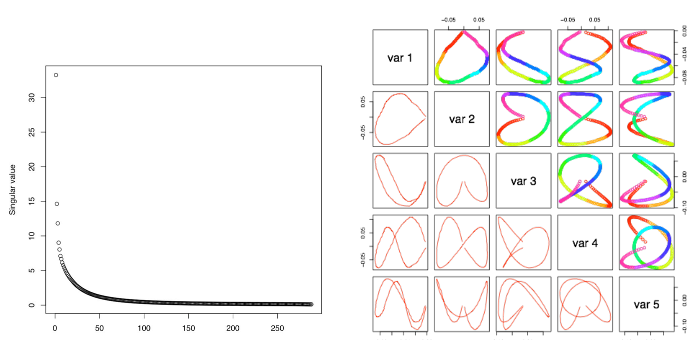

Figure 2. Scree plot and loadings on the first five PCs, 288 five-minute time slices.

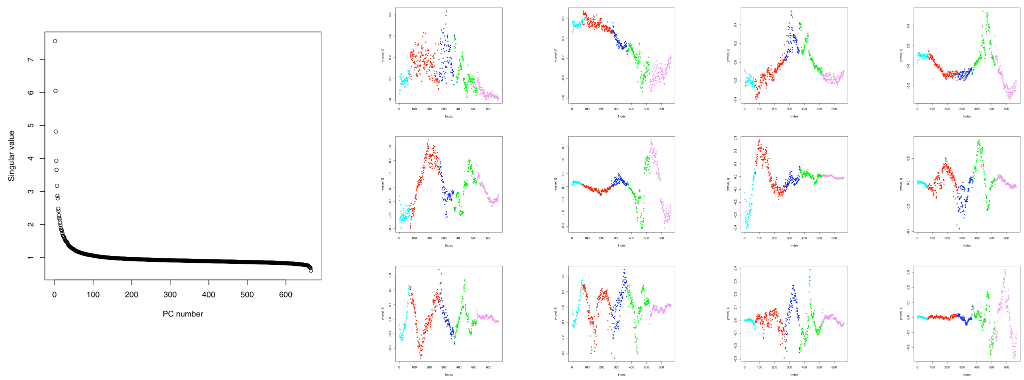

Figure 3. Scree plot and loadings on the first twelve PCs, 665 days.

In Figure 2, the scree plot indicates a fairly typical fall in PC variance over the full range of PCs. The first PC accounts for roughly 11.4% of the variance, the second, another 5%, etc., with diminishing variance corresponding to each additional PC. The "elbow" of the scree plot falls somewhere around 25 PCs in, where 57.25% of the variance is accounted for, although the transition to the more level portion of the plot is fairly smooth, and it is hard to assign a clear stopping point for extracting PCs (cf. Basilevsky 1994). The loadings of the time slots on the first five PCs show clear periodic activity: the first two time slots describe a closed loop, indicating that PC 1 and 2 describe full-day cycles, 90 degrees out of phase with each other. PC 3 through 5 (and higher PCs) add ½ wavelength into each full-day cycle; the resulting scatterplots describe "spirograph" patterns representing the phase and wavelength differences near-perfect sinusoids for each pair of PCs. Hence, we have a harmonic periodic pattern with a daily cycle, and steadily diminishing amplitudes of each harmonic, as frequency increases. Higher harmonics (e.g. above PC 25) have almost identical amplitudes, representing strictly white noise among the higher frequency spectrum.

While it is unlikely that each sinusoidal component can be given an independent interpretation, the transition to noise around PC 25 indicates that the shortest typical period we can resolve in overall BZFlag activity is around 115 minutes in length (almost 2 hours), give or take. Reflecting this, we clustered the PC loadings for each day, weighted by the Λ^2 values using Ward's clustering method, and selected a cut with 10 clusters; these clusters are plotted as distinct chromatic color bands on the plots above the diagonal in Figure 2 (those below the diagonal are the same plots, reflected about their diagonal, but drawn with lines to clarify the phase relations). The color mappings clearly indicate that contiguous time-periods are clustered together, suggesting that time slices closer together have more similar activity on BZFlag than times further apart. There is also no sharp discontinuity in these plots, suggesting that on a daily level, there are smooth transitions from one phase of BZFlag activity to another. In other words, there is a robust daily cycle of activity that is highly regular. Various techniques of spectral analysis are able to identify a weekly change in this daily cycle as well, although this pattern is weak in comparison.

Figure 3 shows that a different situation holds for the longer-term patterns. The elbow in the scree plot is much higher, around 50 or so, and the highest values of Λ^2 are much lower, beginning at just under 8, or 1.13% of the total variance. The proportion of overall variance accounted for at the elbow (PC50) is only 15.34%, indicating lower overall values of correlation in this analysis, and hence a greater proportion of noise than in the analysis of daily patterns. Phase plots of the PCs are less revealing in this analysis as well, because of the increased noise, although a similar overall pattern of periodicity is observable in the PCs, with PC 1 and 2 both expressing an approximately 2-year cycle, and each successive PC adding in ½ wavelength into the two-year period; these regularities are confirmed by spectral analysis of the PCs. The noise component of the PCA begins at a period around 29.5 days, or approximately one month.

Unlike the daily patterns, however, the periodicity of these patterns is subject to greater noise and damping effects, generally at one or both ends of the sample. Some of these patterns can be observed from the plots of the first twelve loadings in Figure 3. The *x* axis of these plots are the 665 days in order, and the colors correspond to "eras" based on which servers were most popular during each period (cluster analysis of the loadings does not cluster contiguous time periods together in this PCA). These eras are: *Boxy Wars* (cyan), from the beginning of the sample up to Dec. 31, 2006; *Missile Wars 1* (red), from Jan. 1 to mid-June, 2007; *Bloodbath 2* (blue), mid-June to late September 2007; *Missile Wars 2* (green), from October 2007 to May 2008; and a *mixed* period (pink) in which Missile Wars 2 still mostly dominates but other servers have enjoyed resurgence, from June 2008 to the end of the sample. Remarkably, the boundaries of these events appear to coincide with periodic and damped phases of the PCs.

We explain these patterns and the difference between them and the daily patterns as follows. When a server is popular, such as Boxy Wars (originally the main bzfx server), it supports a cohort of players who can expect to find others to engage with in developing and expressing various game skills and strategies, at regular temporal intervals. This cohort is expressed in our analysis as relatively strong correlations across different time slots or days, in the respective analyses. When such a server loses its hosting, as happened on Boxing Day in 2006 with Boxy Wars, the support for the cohort no longer exists: other maps/server require different strategies, and one cannot expect to find the same people on them. Hence, at the point of such an event, players no longer congregate regularly at the same times on different days, and the cohort either dissipates, or re-forms on a different server, generally with a different membership. This results in a damping of the periodic patterns, when a cohort is lost, or amplification of the patterns when a new cohort arises on a newly popular server.

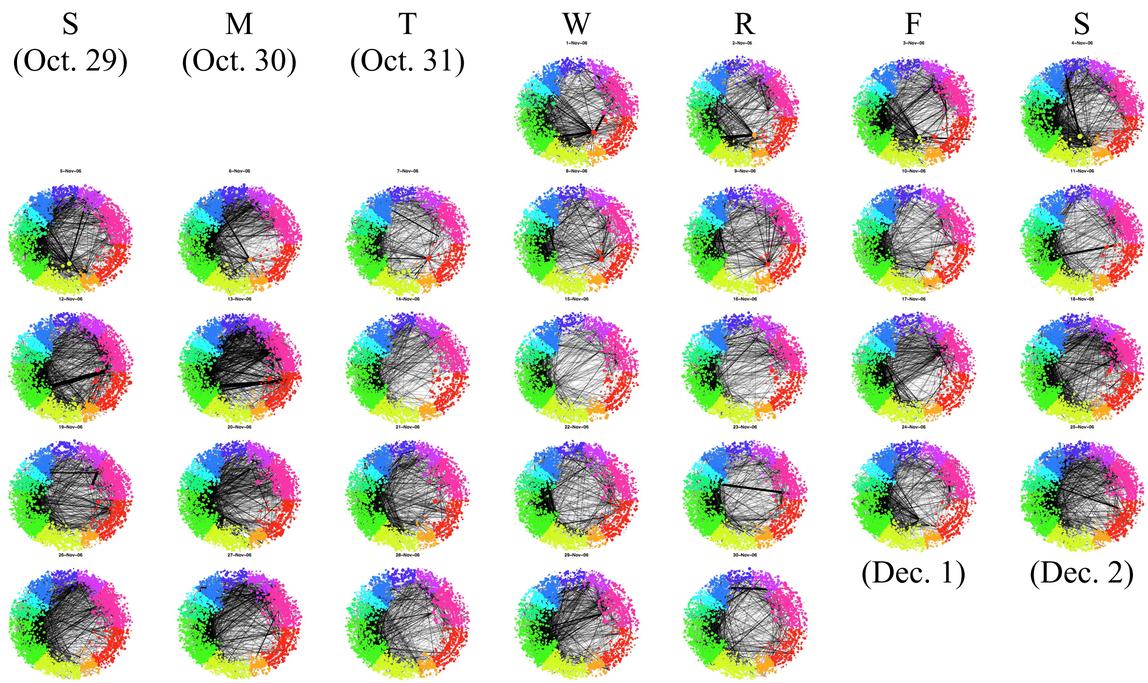

Figure 4. Calendar of social network maps for BZFlag players in the month of November, 2006.

To get a clearer idea how these patterns are manifest in the social connection among players, sociograms were constructed for each of the 665 days in the sample by locating each active player on any given day using a modified principal components scatter-plot for node locations and aggregating player-player contacts according to their number of shared sessions on any given server; contacts exceeding 80 minutes of shared time were plotted. The first two Principal Components were used to plot player locations; normally, this places central nodes on the edges of the graph. To give a more natural layout organized a bit more like force-directed layouts, we "rolled" the layout into a doughnut shape by preserving the angle from PC1 and 2, and subtracting the magnitude of a point on all PC's from the unit circle radius; this has the effect of mapping higher (more central) values toward the center of the diagram, with peripheral points being spread toward the unit circle. Thirty such plots are shown in Figure 4 for the month of November, 2006 (during the Boxy Wars era), arranged in calendrical form, so that daily and weekly patterns are evident. Player nodes are colored according to the proximity of each player on a given day to one of 10 time-slot clusters, in other words, the approximate center of their playing time for that day. The red cluster begins at 00:00 UTC, with clock time increasing clock-wise, so that 12:00 UTC is roughly 180 degrees from midnight, at the beginning of the aqua cluster.

A number of patterns are evident in this monthly display. The first is that the density of connection is greatest from the yellow through the cyan time periods, in other words, during the late morning to early afternoon times in Europe, or early to mid-morning US continental times, and late evening in East Asian and Australian time zones. A second cluster of contacts, though not as dense and not as consistent in appearance, begins from the blue range and extends to the magenta, representing late afternoon to evening European times and late morning to afternoon US times. Densities in both clusters are more evident on the weekend days, from Friday to Monday (which includes part of Sunday in US time zones). Tuesday through Thursday show the thinnest overall activity. Finally, a small number of players appear, generally in the early morning hours UTC, with connections to players in most of the other time ranges. These are either highly active players, or possibly loggers on highly popular servers.

To examine how individuals might move through the social environment of BZFlag, we examined the “careers” of three prominent individuals with more than 500 days logged in the sample: T, with 513 days; A with 501 days; and R, with 625 days. All three are central participants in that they have authored highly popular maps, and have hosted their own and others’ maps on their own servers. We plot the careers of these players in terms of the first six PCs of the time-slice analysis in Figure 5, with PCs 1 and 2 in the first plot, PCs 3 and 4 in the second, and PCs 5 and 6 in the third. Observed values for each player on each day are plotted as points, along with smoothed lowess curves indicating the central trend for each player. Numbers along the lowess curves indicate the day value for each smoothed point, at ten-day intervals. In these plots, positions further from the origin suggest closer association (higher correlation of playing times/servers) with a particular cohort, represented by whichever PC (and associated periodic pattern) is involved. Positons closer to the origin suggest weaker association with any particular cohort. Moreover, points from opposite ends of any of these plots represent complementary cohorts, i.e. those whose server/time patterns are not shared. Hence, these plots allow us to interpret the careers of different players in terms of shared social contacts over time.

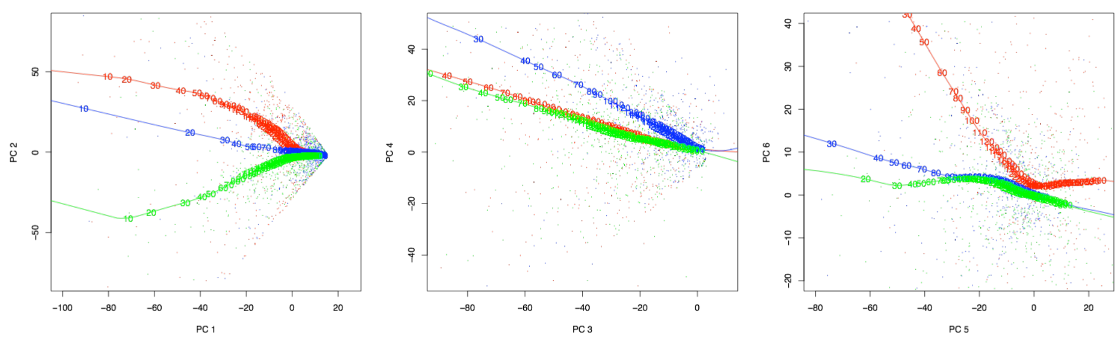

Figure 5. Careers of three prominent players: T (red), A (blue) and R (green).

All three plots show a movement from left to right in all three careers. The first plot, in which the angle from the origin to each point corresponds roughly to the average time of day for that player’s participation on a given day, shows that the typical playing times of each player migrated over the 665 day period, from early afternoon UTC for T, close to noon for A and late morning for R, toward times that are closer together and shifted toward a bit after midnight UTC for all three players. This corresponds to evening hours in the Eastern US time zone, where all three players reside. At the same time, it is clear from the broad spread of all three colors of points that the three players exhibit great variability in their playing times, so these trends are only a rough index at best.

The second plot shows all three players converging on the origin of PCs 3 and 4 from roughly the same direction. This suggests that there was a cohort in the early days of these players’ experience which dissipated relatively quickly. One potential candidate for this cohort would be the Boxy Wars cohort; T and R were both frequent players on the Boxy Wars server, and they formed an alliance and friendship based on that experience, which carried over into other eras, when T helped R with hosting one of his maps, that later became one of the more popular maps on BZFlag, note that T and R follow almost identical trajectories in this plot.

The third plot shows A and R approaching the origin on similar trajectories, and exiting on the right side as well. T approaches from a different direction, and exits on a different trajectory as well. This suggests that there are elements of T's early experience that are not shared much with that of A and R; similarly, T's later experience (in which time T plays less and eventually largely leaves BZFlag) places him in contact with a cohort distinct from both his own early experience, as well as that of A and R's earlier and later experience. For a period in the middle of their careers, all three players experience relatively similar cohorts, somewhat more like those in their early experience, but closer to the origin, suggesting that the early cohort may be diluted to some extent by increased contact with a broader range of players, as all three logged greater online hours.

5. Discussion

The analysis of social and temporal patterns in BZFlag activity highlights a number of important observations. First, although the distribution of player activity is highly skewed, and a majority of players play for only a short period of time, the social environment of BZFlag is highly structured. Time plays an important role in the structuring of the environment, with the majority of activity concentrated in late morning and early afternoon UTC, a time when players in many time zones have waking hours that they can devote to online gaming. A second and smaller group of players is active primarily during the US afternoon and evening times, when fewer international players are likely to be found. Hence, there is a natural division in the community by continental and time zone boundaries. These boundaries are not absolute, and there are frequent periods of strong connection between the players of the American and Eurasian peak hours.

A second major observation is that there are patterns of periodic activity both on the daily scale and on longer scales, although the longer-scale patterns are subject to greater noise, as well as damping effects. These effects are explicable as the emergence and dissipation of cohorts, caused by server hosting events, such as the introduction of a new map that rapidly rises to popularity, or the sudden loss of hosting services through a failure of an alliance between two or more players, admins, hosters or other participants. These events have broad consequences for the cohesion of social cohorts supported by different servers, and for the community structure as well, but the harmonic properties of the PCA do not appear to be affected. In some sense, this suggests that there are different levels of scale in which such server hosting events play out. On the smaller, interpersonal scale, changes are largely unpredictable, and cohorts are relatively fragile, being subject to sometimes tenuous hosting alliances for the existence of spaces in which people's preferred social activities can take place. On a larger, global scale, the patterns are almost deterministic, being governed by harmonic periodic activity in the patterns of players connecting to servers.

A further set of patterns we have observed pertains to the careers of individuals in the context of the global patterns of activity. Individuals, through their activity on servers, through creating maps and hosting them, and in working out alliances with other players for the maintenance of gameplay environments, experience trajectories through the social environment that place them in contact with different cohorts at different times of their careers. In addition, server hosting events form part of the social experience of an individual in the game. These patterns are expressed as trajectories through the principal components space in our analysis, which we can compare in order to better understand the overall development of the social environment, as well as an individual's experience within it.

6. Conclusion

The results of this study underscore the importance of the time dimension in understanding patterns of social activity in online games, and in understanding social relationships more generally. Tie strength is not constant over time, and ties responsible for community cohesion require continuous maintenance. The overall trajectory of a social activity, and the relation of individuals' patterns to that of the group, are better understood when the time dimension is taken into account in a way that allows one to resolve temporal patterns on different levels of scale simultaneously. Building this account is not a simple matter and a number of challenges had to be faced to accomplish this. The nature of these challenges are potentially instructive for other research in online games or other communication and social activities that can be similarly logged and investigated.

One of the main challenges in the project was dealing with a large and dynamic dataset in which the active players and servers are constantly changing. On a technical level, this data characteristic complicated our network analysis and longitudinal analysis, requiring novel data arrangements to reveal the patterns of interest. Of particular interest here was the need to address the temporal aspects of player behavior, in particular regular periodic patterns of connection. This required an approach that would allow the application of time-series analysis techniques. The approach adopted did not allow us to make independent estimates of the relative contributions of individuals and servers to the patterns we seek. Hence, our analysis functions as an exploratory mode of research that merely helps identify patterns of interest for further investigation. Future research will need to adopt different methods to confirm these patterns and test them for statistical significance.

This research has other potential benefits as well. For example, it provides some guidance for game designers in the construction of more sophisticated metrics for evaluating the success of changes in a game environment. All online games can be expected to undergo development, addition and removal of features, and infrastructural support such as servers. These changes can have a large impact on the social environment and how individuals experience it. At present, game companies tend to use relatively simple metrics, such as the number of concurrent players or subscribed players. But these numbers are frequently contentious, because it is often unclear both how they are calculated, and what they mean. Player growth or decline as a metric of player interest in a game is but one of the broader temporal indicators; our research here indicates that one should also be able to identify other fluctuations such as daily, weekly, and seasonal patterns. Primetime hours, holidays, school and work breaks, and time of year can all influence who is online at the same time. Punctuated events, such as the release of a new player map, a version update of the game itself, or server problems could all drive increases or decreases in player participation, as well as the emergence and dissipation of player cohorts.

These observations give further reasons to investigate more closely the temporal and social patterns of other online games, as well as other activities, both online and off. By understanding the temporal patterns of social and ludic activities, and how they influence the expression and experience of social connection, we can better appreciate the meanings of our game activities and the social worlds they open up to us.

7. References

Basilevsky, A. (1994). *Statistical Factor Analysis and Related Methods*. New York: Wiley.

Chen, K.-T., Huang, P., Huang, C.-Y., & Lei, C.-L. (2005). Game traffic analysis: an MMORPG perspective. *Proceedings of the international workshop on Network and operating systems support for digital audio and video*. New York, NY: ACM.

Degenne, A.; and Forse, M. (1999). *Introducing Social Networks*. Thousand Oaks: Sage.

Ducheneaut, N., Yee, N., Nickell, E., & Moore, R. J. (2006). Building an MMO with mass appeal: A look at gameplay in World of Warcraft. *Games and Culture*, 1(4), 281-317.

Eagle, N.; and Pentland, A. (2006). Reality Mining: Sensing complex social systems. *Personal and Ubiquitous Computing*, 10 (4), 255-268.

Fritsch, T., Voigt, B., & Schiller, J. (2006). Distribution of online hardcore player behavior: (how hardcore are you?). *Proceedings of 5th ACM SIGCOMM workshop on Network and system support for games*. New York: ACM Press.

Granovetter, M.S. (1973). The strength of weak ties. *American Journal of Sociology*, 78: 1360-80.

Jakobsson, M., & Taylor, T. L. (2003). The Sopranos meets Everquest: Social networking in massively multiuser networking games. *fineArt forum*, 17(8).

Juul, J. (2004). Introduction to game time. In N. Wardrip-Fruin and P. Harrigan, eds., *First Person: New Media as Story, Performance, and Game*, 131-142. Cambridge, MA: MIT Press, 2004. Retrieved September 15, 2008 from http://www.jesperjuul.net/text/timetoplay/

Klastrup, L. (2006). Death matters: Understanding gameworld experiences. *Proceedings of the 2006 ACM SIGCHI international conference on Advances in computer entertainment technology*. New York: ACM Press.

Kolo, C., & Baur, T. (2004). Living a virtual life: Social dynamics of online gaming. Game Studies: *The International Journal of Computer Game Research*, 4(1). Retrieved September 15, 2008 from http://www.gamestudies.org/0401/kolo/

Lindley, C.A. (2005). The semiotics of time structure in ludic space as a foundation for analysis and design. *Game Studies: International Journal of Computer Game Research*, 5(1). Retrieved September 12, 2008 from http://www.gamestudies.org /0501/lindley/

Milroy, L.; and Milroy J. (1992). Social network and social class: Towards an integrated sociolinguistic model. *Language in Society*, 21:1, 1-26.

Nardi, B., & Harris, J. (2006). Strangers and friends: Collaborative play in World of Warcraft. *Proceedings of CSCW 2006* (pp. 149-158). New York: ACM Press.

R Development Core Team. (2008). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Schroeder, R. (1997). Networked worlds: Social aspects of multi-user virtual reality technology. *Sociological Research Online*, 2(4).

Seary, A.J.; and Richards, W.D. (2003). Spectral methods for analyzing and visualizing networks: An introduction. In R. Breiger, K. Carley, and P. Pattison, eds. *Dynamic Social Network Modeling and Analysis: Workshop Summary and Papers*, 209-228. Washington, D.C.: National Academies Press.

Seay, A. F., William, J. J., Kevin Sang, L., & Robert, K. (2003). Project Massive 1.0: Organizational commitment, sociability and extraversion in Massively Multiplayer Online Games. *Proceedings of the Level Up Conference*.

Steinkuehler, C., & Williams, D. (2006). Where everybody knows your (screen) name: Online games as "third places." *Journal of Computer-Mediated Communication*, 11(4). Retrieved June 15, 2008 from http://jcmc.indiana.edu/vol11/issue4/steinkuehler.html.

Williams, D., Ducheneaut, N., Xiong, L., Zhang, Y., Yee, N., & Nickell, E. (2006). From tree house to barracks: The social life of guilds in World of Warcraft. *Games and Culture*, 1(4), 338-361.