On the Geographic Distribution of On-line Game Servers and Players

Wu-chang Feng   Wu-chi Feng
OGI@OHSU
{wuchang,wuchi}@cse.ogi.edu *

ABSTRACT

With a shift in the on-line gaming landscape from individually hosted game servers, to gaming services centrally hosted by game publishers, game console manufacturers, and third-party infrastructure providers, it is becoming increasingly important to understand the geographic distribution of current game servers and players. Much like content-distribution networks are key in improving user web experience, game server placement is key in improving game player experience. This paper explores the current geographic distribution of a global set of servers for several popular on-line games as well as the geographic distribution of a set of players for a particular on-line game server. Our results quantify the breakup of current game servers across continents and show, quite suprisingly, that players do not necessarily migrate to servers that are geographically close.

1. INTRODUCTION

With the increasing popularity of on-line PC games over the last several years and the emergence of on-line console gaming such as Sony's Playstation 2 and Microsoft's Xbox Live, it is becoming important for games and game providers to deliver reliable, low-latency service to game players. Much as content-distribution networks are key to delivering fast and reliable web service to clients, game server placement is instrumental in delivering high-quality gaming experiences to players. Ideally, players would like to reside either on the same LAN as the server hosting their game or as close to it as possible. This is especially true for the class of games known as first-person shooters (FPS) [4]. As lag is a major component in determining the quality of the experience, placing game servers in locations that have low latencies to the player population is an essential task.

Over the last several years, the landscape of on-line games has increasingly shifted from game servers hosted by individuals to game servers that are run and tightly controlled by game publishers [1], console makers [2], and even third-party service providers [3]. Some of the reasons for this shift include the desire to control the game's quality and user experience as well as to prevent cheating. One of the challenges facing this approach is the ability to place servers in close proximity to users. As many games are extremely sensitive to latency, placing servers geographically and topologically close to players is extremely important.

We hypothesize that the geographic distribution of game servers run by individuals closely matches the geographic distribution of game players themselves and that it is desirable for hosted game services to match the current geographic distribution of individually run game servers. Because of this, it is extremely valuable to have a characterization of the geographic distributions of current game servers and game players. In this paper, we present our initial work in the geographic characterization of game servers for several popular FPS games as well as a geographic characterization of the game players for one particular, popular game server that we host ourselves. The questions we seek to answer in our study are the following:

- *Where are current game servers located?* Locations of individual game servers gives us an indication as to where players are. As a result, such information can be used to properly size and place clusters of servers for hosted gaming services [3].
- *How important is geographic proximity to game servers?* Recent work indicates that for a class of on-line games, lag is a clear determinant in player experience and performance [5, 6, 7]. It has been shown that players tend to migrate to near-by servers. Since it is often infeasible for hosted gaming services to have data centers in every location throughout the world, quantifying the tolerable delay is important in determining how distributed these data centers must be. If the geographic distribution of players is tightly correlated with the geographic location of the server, then hosted game

*This work is supported by the National Science Foundation under Grant EIA-0130344 and the generous donations of Intel Corporation. Any opinions, findings, or recommendations expressed are those of the author(s) and do not necessarily reflect the views of NSF or Intel.

NetGames '03 May 22-23, 2003, Redwood City, California USA

services must employ clusters in many geographic locations. If the correlation is weak, then placing clusters in a few strategic locations is all that is required.

Section 2 describes the methodology used to collect and generate geographic information on game servers and players in this study. Using this methodology, Section 3 describes our initial work on analyzing the geographic properties of a set of game servers and of a player population of a particular game server. Section 4 concludes with a discussion of the implications that these results have on the placement of game servers.

2. METHODOLOGY

2.1 Obtaining server and player addresses

Most on-line games implement a centralized registry server that individual game servers register with at startup. As part of the game client, this global list of game servers can then be downloaded to support an in-game server browser that allows players to select an appropriate game server. Central to this server selection process is the round-trip latency from the client to each server in the list, often referred to as the *ping time*. To generate a ranked list of servers based on their ping time, each client can take the list of server IP addresses, individually ping each one, and sort them for the player. Given this, it is relatively easy to obtain a list of server IP addresses per game at a particular time. In this study, we installed game clients for Counter-Strike [8], Battlefield 1942 [9], and Unreal Tournament 2003 [10]. For each game, we then initiated a query to generate a ranked list of game servers based on their "ping" times and dumped the resulting traffic. After parsing this traffic stream, we recovered the complete list of servers for each of the above games. Table 1 describes the list of servers collected using this method for each game.

Game servers as of Tue Nov 26 2002	
Counter-Strike	50374
Battlefield 1942	2369
Unreal Tournament 2003	2172

Table 1: Number of servers

While generating a global list of servers is relatively easy, obtaining a global list of players is difficult. Because access to centralized authentication and game servers that contain this information requires coordination with game publishers, we instead focus our study on the geographic distribution of players for a particular, popular game server. To this end, we obtained the trace from an extremely popular Counter-Strike server, cs.mshmro.com [4]. This server is generally busy 24 hours a day, 7 days a week. By parsing this trace, the IP addresses of connecting players over time was obtained. Table 2 describes the trace from which this data was collected.

Server trace information	
Server Name	cs.mshmro.com
Start Time	Thu Apr 11 08:55:04 2002
Stop Time	Thu Apr 18 14:56:21 2002
Total Time	7 d, 6 h, 1 m, 17.03 s
Player Sessions	16030

Table 2: mshmro trace information

2.2 Deriving geographic locations

Given a list of server IP addresses for each game, we employed a commercial geographic mapping tool to map the IP addresses above into geographic locations [11]. While the tool itself is powerful, IP address to geographic location mapping is an inexact process. Thus, it was impossible to resolve all of the IP addresses obtained into geographic locations. Over the course of our evaluation, the tool was able to generate geographic locations for over 60% of the IP addresses collected. It is important to note that this study represents a pathologic skew in the database directly impacts the accuracy of the results presented. The tool itself is being updated continuously, allowing for a more accurate characterization over time.

3. EVALUATION

3.1 Game servers

Currently, as of March 2003, one of the most dominant on-line games is Counter-Strike. Figure 1 shows the per-longitude histogram of Counter-Strike servers over a map of the world. As the figure shows, the servers appear to be primarily distributed across North America, Europe, and Asia. In North America, a large number of servers reside along the west coast while the rest are spread out across the midwest and the east coast. In Europe, the servers are spread across the continent fairly evenly. Finally, in Asia, the servers are clustered in the east.

In order to compare the geographic distributions across games, Figure 2 plots the longitudinal cumulative density function (CDF) of all Counter-Strike, Battlefield 1942, and Unreal Tournament 2003 servers as of November 2002. As with the histogram plot, the figure is overlaid on top of the map of the world in order to more clearly show its relation to geographic locations. As the figure shows, the distributions of game servers for each game closely match each other with the only variant being Unreal Tournament 2003, which contains more servers that are based in North America versus the other two games. The CDF also shows that for Counter-Strike and Battlefield 1942, most of the servers reside in North America and Europe and are evenly split between the two with approximately 40-45% of them residing in each continent.

In the previous figures, it is assumed that most of the servers reside in continents in the northern hemisphere. In order to demonstrate that this is the case, Figure 3 shows

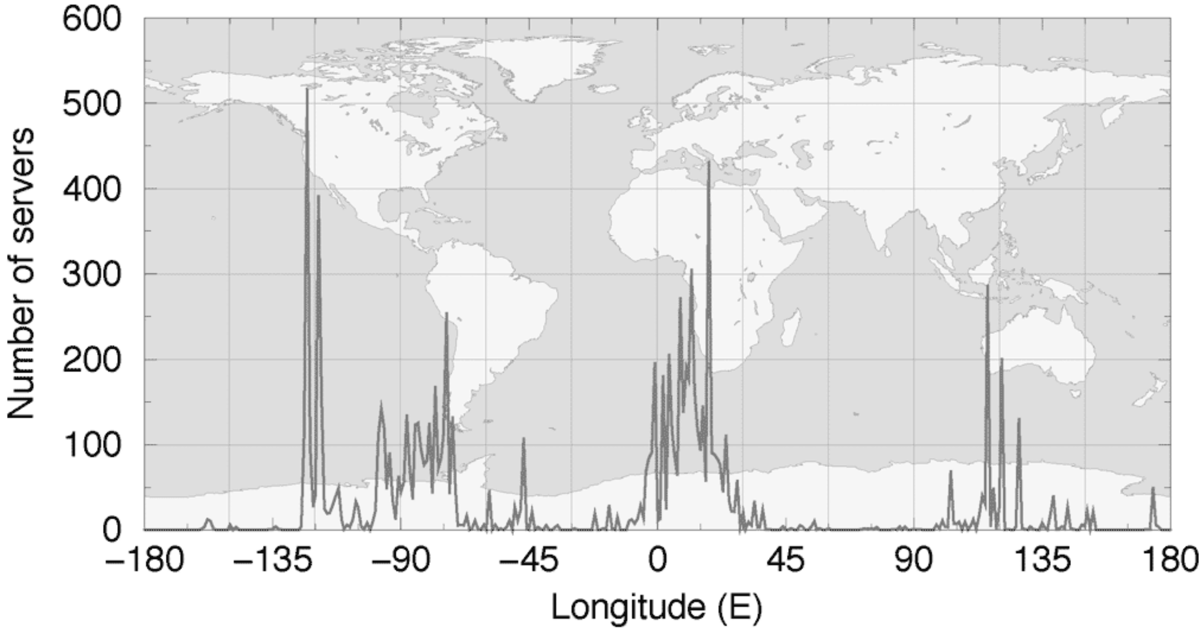

Figure 1: Longitude histogram of Counter-Strike servers

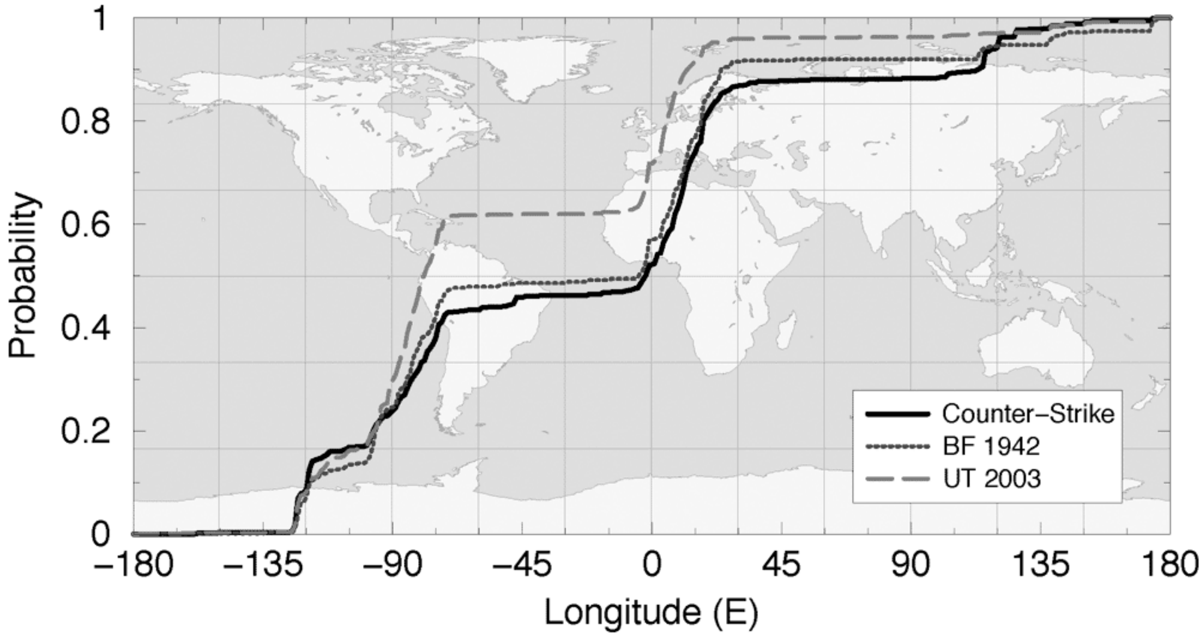

Figure 2: Longitude CDF of game servers

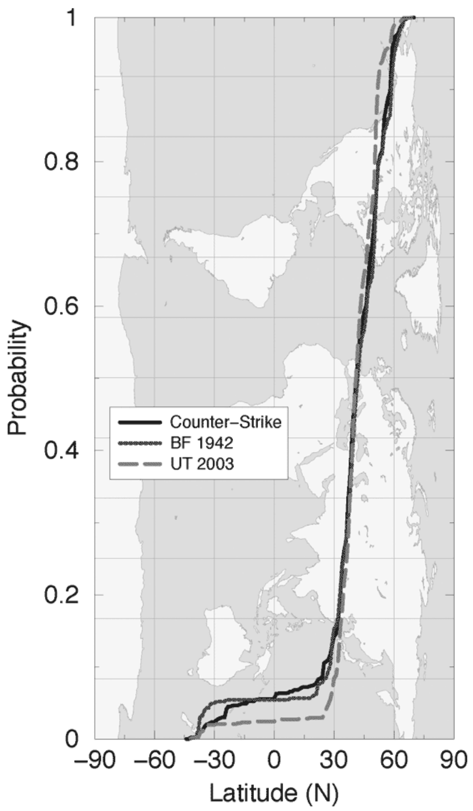

Figure 3: Latitude CDF of game servers

the latitudinal CDF of the same set of game servers over a map of the world. As the figure shows, except a handful of servers located in Australia and New Zealand, almost all of the game servers reside north of the equator. In fact, for the three games studied, more than 90% of the game servers reside in the northern hemisphere.

3.2 Game players

Without access to centralized authentication servers or centralized game server farms, it is impossible to characterize the geographic distribution of global game populations. However, it is possible to examine the geographic distributions of a single game server. Because lag is such a large factor in the quality of game play, it is expected that player populations for a particular server are clustered around the geographic location of the server itself. It is important to note, though, that geographic proximity and network proximity are only somewhat correlated. While the continued build-out of Internet infrastructure (in terms of exchange points and last-mile links) will continue to tighten this correlation, it is possible for players that are geographically close to a server to have considerable network latency to it.

To study the geographic distributions of game players, we examined the Counter-Strike trace from an extremely popular Counter-Strike server [4] located in Beaverton, Oregon (45.4828N, -122.825E). Figure 4 shows the player histogram of all players playing on the server over the week-long trace. The geographic location of the server is designated by a circle in the northwestern part of the United States. As the figure shows, while a large number of players reside close to the server, it is surprising to find a significant portion of the player population connecting from Europe and Asia, as well.

Figure 5 plots the longitudinal CDF of player locations for the mshmro trace. As the figure shows, the CDF shows that only 30% of all players reside within 10 degrees longitude and just over 50% of all players reside in North America. This leaves an astounding 45% of players connecting from across the ocean to play on the server. Without further studying additional servers located throughout the world, it is difficult to understand the reason for this. Some of the possible explanations for this phenomenon could include:

- *Disparity between geographic location and network topology:* As described earlier, geographic proximity does not always correspond to network proximity. For example, before an exchange point was brought up in the Portland metro area, connections between machines within Portland often travelled via San Jose, California.

- *Application server delays dominate network delay:* End-to-end response times contain both application delays and network delays. For the mshmro trace, however, it is unlikely that this was the case due to the speed of the server. Several sessions empirically observed during the week-long trace had single-digit (ms) delays indicating a very small baseline application server delay.

- *Server selection mechanisms for popular games are broken:* Many gamers rely on a “Quick Start” or auto-connecting mechanisms to connect themselves to appropriate game servers. While it would be ideal if such mechanisms connected players to the closest server with the lowest latency, anecdotal experiences with them have shown that they can select poorly at times.

- *The number of players on a server determines desirability over delay:* Many on-line games are simply not fun when only a limited number of players are playing. For example, large maps can be boring with a small number of players since it takes a significant amount of time before players find each other to shoot. Because of this, players often connect to servers that have a reasonable number of players on them. This can make it difficult for an idle server to obtain new players, even if it has been configured properly and is in a prime network location

- *A shortage of servers overseas:* While the data in the previous section indicates a significant number of servers located in Europe, it is possible that during peak times that there simply aren’t enough servers to support the number of players. In addition, the converse may be true with U.S. players connecting to European servers during peak hours of the day. Such a

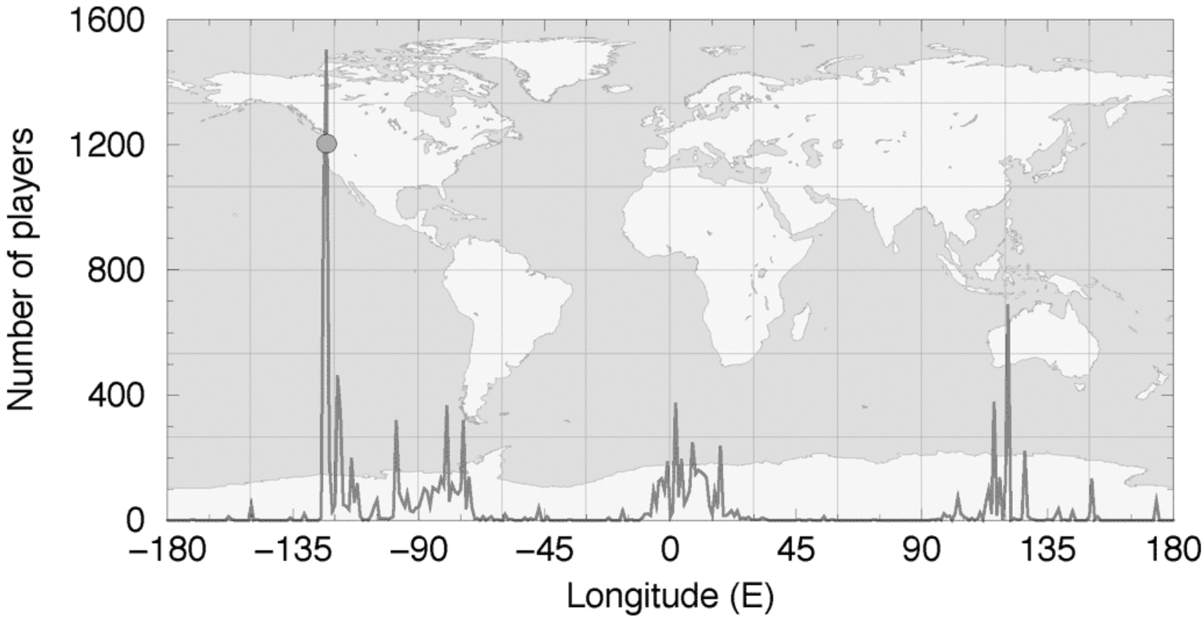

Figure 4: Longitude histogram for players in mshmro trace

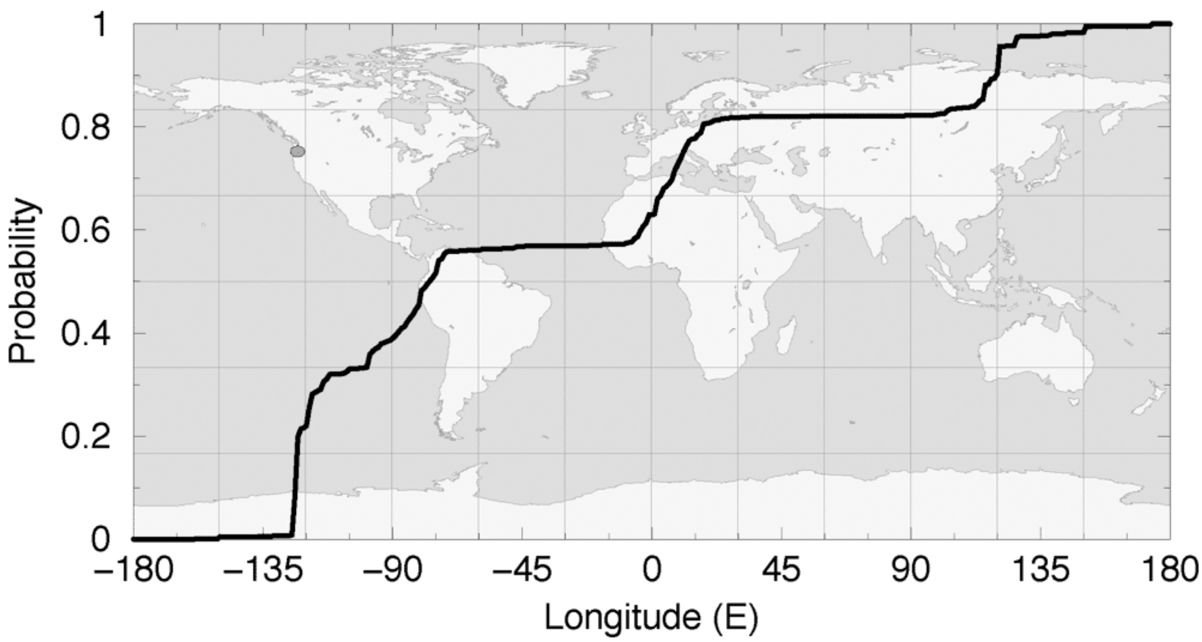

Figure 5: Longitude CDF for players in mshmro trace

(a) midnight to 4am	(b) 4am to 8am
(c) 8am to noon	(d) noon to 4pm
(e) 4pm to 8pm	(f) 8pm to midnight

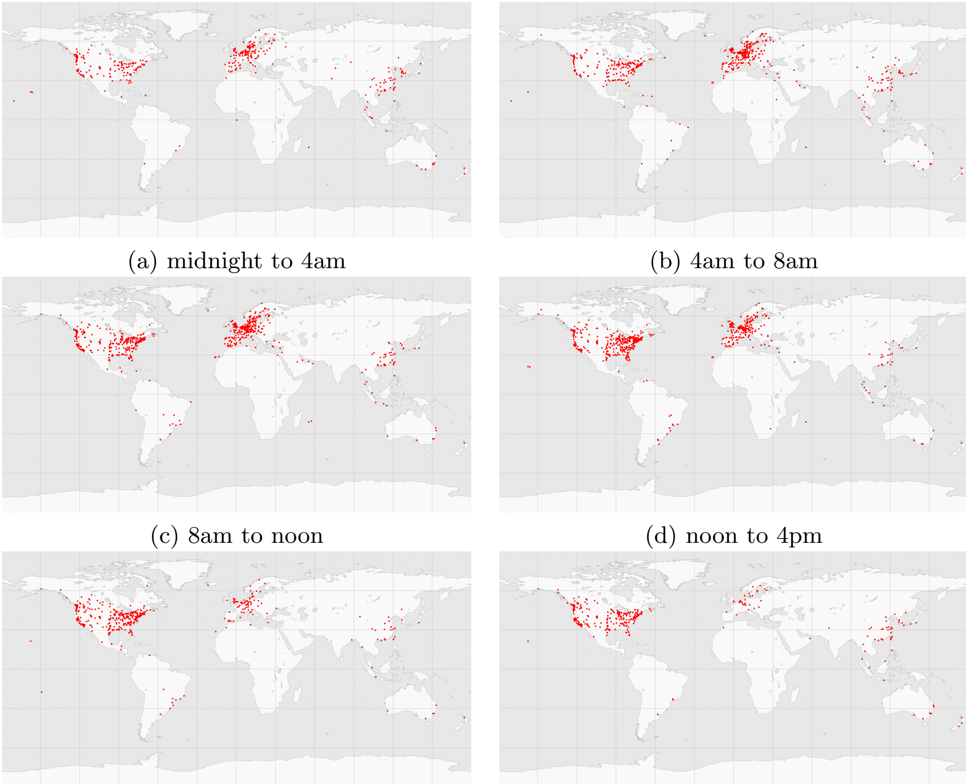

Figure 6: mshmro player locations over time

phenomenon could be verified with traces from European servers.

Although, by no means definitive, Figure 6 provides a partial indication that the determining factor in the geographic distribution of players is, in fact, the time of day. The figure plots player connections during 6 different 4-hour blocks of the day during the week-long trace. The times are given in the server's local time (Pacific Standard Time). As the figure shows, the locations of players is driven by the time of day, with most of the connections originating from Europe and Asia occuring during early morning and afternoon hours of the day. This is an interesting phenomenon as it means that for grid-based, computing on-demand, gaming services such as Butterfly.net, that a global positioning and repositioning of resources over time is required to match the usage patterns of players. For example, it would be desirable to shift game servers and computing capacity to European data centers during early morning hours (PST) to match where the demand is coming from.

Figure 7 quantifies the time of day phenomenon by plotting the great-circle distance (defined as the shortest distance along a sphere between two points on the sphere) between the mshmro server and its players over the duration of the trace. As the figure shows, the average distance shows regular periodic behavior over time, with the average distance peaking during early morning hours and dipping during late afternoon and evening hours. This result is consistent with previous observations on the variation of client latencies to a number of game servers across different times of the day [12, 6]. The high correlation between the behavior of network latencies and geographic distances from the server, seem to indicate a fairly strong correlation between the two.

One of the questions that is interesting to pose is how different the geographic distribution of players for a game server is from the geographic distribution of clients for a web server. Figure 8 plots the average great circle distance of clients versus the time of day for the month of September 2002 on three different servers: the mshmro Counter-Strike server, our department web server, and the mshmro web server. The mshmro web server contains game statistics and a bulletin board for players to post and read messages from other players. As the figure shows, compared to the departmental web server, the Counter-Strike server tends to attract clients that are consistently closer geographically than the web server itself. This is can be attributed to the slight preference that clients have to play on servers that are close-by. The figure also shows that the mshmro web site that hosts the player community shows highly localized client connects. This indicates that there is a strong geographic correlation among active gaming communities and the servers that they play on.

4. CONCLUSION

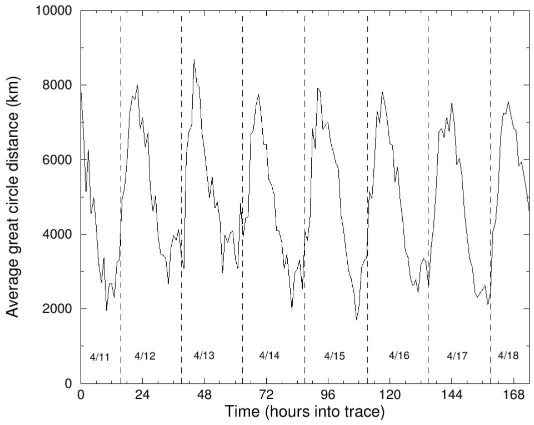

Figure 7: Average great-circle distance of players over time

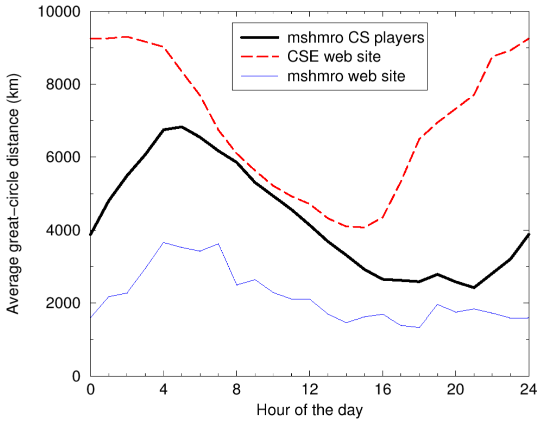

Figure 8: Comparison of average great circle distance versus time-of-day for various servers

This paper describes our preliminary work on characterizing the geographic distribution of game servers and game player populations. Our results show that game servers are distributed mainly across the northern hemispheres in the United States, Europe, and Eastern Asia. In addition, the results indicate that while there is some geographic preference between players and servers, that the dominant factor in the geographic distribution of game players is the time-of-day. We are continuosly looking to improve the accuracy of our results and expand our characterization to more completely capture geographic distributions of servers and players. In addition, we plan on studying the evolution of populations over time as gaming applications change. Finally, we are also interested in global game player populations and how they are distributed across the world over time. Such a characterization will more aid in the provisioning of global server resources for distributed gaming services such as Butterfly.net [3].

5. REFERENCES

[1] Electronic Arts, Inc., "EA.com," `http://www.ea.com/`, 2003.

[2] Microsoft Corporation, "Xbox Live," `http://www.xbox.com/live`, 2003.

[3] Butterfly.net, Inc., "Butterfly Grid Solution for Online Games," `http://www.butterfly.net/`, 2003.

[4] W. Feng, F. Chang, W. Feng, and J. Walpole, "Provisioning On-line Games: A Traffic Analysis of a Busy Counter-Strike Server," in *Proc. of the Internet Measurement Workshop*, November 2002.

[5] I. S. MacKenzie and S. Ware, "Lag as a Determinant of Human Performance in Interactive Systems," *Proceedings of the ACM Conference on Human Factors in Computing Systems - INTERCHI*, pp. 488–493, 1993.

[6] T. Henderson and S. Bhatti, "Modelling User Behavior in Networked Games," in *ACM Multimedia*, 2001, pp. 212–220.

[7] T. Henderson, "Latency and User Behaviour on a Multiplayer Game Server," in *Networked Group Communication*, 2001, pp. 1–13.

[8] Counter-Strike, "Counter-Strike: A Counter-Terrorism Half-Life Modification," `http://www.counter-strike.net/`.

[9] Electronic Arts, Inc., "Battlefield 1942," `http://www.ea.com/eagames/official/battlefield1942/home.jsp`, 2003.

[10] Epic Games, "Unreal Tournament 2003," `http://www.unrealtournament2003.com/`.

[11] Geobytes, Inc., "Geobytes Home Page," `http://www.geobytes.com/`, 2003.

[12] G. Armitage, "Sensitivity of Quake3 Players To Network Latency," in *Internet Measurement Workshop (Poster Session)*, November 2001.