World of Warcraft Avatar History Dataset

Yeng-Ting Lee¹  Kuan-Ta Chen¹
Yun-Maw Cheng²  Chin-Laung Lei³

¹National Taiwan University
²Academia Sinica
³Tatung University
Diversity in Game Play Behavior

- Normal
- Just AFK at the same place, no login/logout
- long time offline
- long consecutive game play time

World of Warcraft Avatar History Dataset / ACM Multimedia Systems 2011
Motivation

- Understanding users’ game player behavior
  - Login/logout (i.e., game sessions)
  - Movement
  - Involvement (i.e., level-up)

- Understanding users’ interaction
  - Same game player behavior in a gang

- Understanding game systems’ workload
We Present WOWAH Dataset

Overview

From the perspective of game system designers, players' behavior is one of the most important factors they must consider when designing game systems. To gain a fundamental understanding of the game play behavior of online gamers, exploring users' game play time provides a good starting point. This is because the concept of game play time is applicable to all genres of games and it enables us to model the system workload as well as the impact of system and network QoS on users' behavior. It can even help us predict players' loyalty to specific games.

We present the World of Warcraft Avatar History (WoWAH) dataset, which comprises the records of 91,065
Contributions

- WOWAH Dataset
  - The most popular MMORPG, World of Warcraft

- Data collection methodology

- Sample use of the dataset
  - Player unsubscription prediction
  - Feasibility of server consolidation
## Summary of WOWAH

<table>
<thead>
<tr>
<th></th>
<th>TW-Light’s Hope</th>
</tr>
</thead>
<tbody>
<tr>
<td>Realm</td>
<td>Horde</td>
</tr>
<tr>
<td>Faction</td>
<td>Horde</td>
</tr>
<tr>
<td>Start date</td>
<td>2006-01-01</td>
</tr>
<tr>
<td>End date</td>
<td>2009-01-10</td>
</tr>
<tr>
<td>Duration</td>
<td>1,107 days</td>
</tr>
<tr>
<td>Sampling rate</td>
<td>144 samples per day</td>
</tr>
<tr>
<td># of samples</td>
<td>159,408</td>
</tr>
<tr>
<td># of missing samples</td>
<td>21,324</td>
</tr>
<tr>
<td># of avatars</td>
<td>91,065</td>
</tr>
<tr>
<td># of sessions</td>
<td>667,032</td>
</tr>
</tbody>
</table>
World of Warcraft
Data Collection Methodology

- Create a game character
- Use the command `\who`
- The command asks the game server to reply with a list of players who are currently online

- Write a specialized data-collection program (using C#, VBScript, and Lua)
The Limitation of WoW API

- WoW returns at most 50 users in one query
- We narrow down our query ranges by dividing all the users into different races, professions, and levels

- Level: 50+
  - 100 users

- Level: 40~49
  - 60 users

- Level: 30~39
  - 45 users
  - Monster
  - 15 users
  - Human
# Data Format

<table>
<thead>
<tr>
<th>Query Time</th>
<th>Seq. #</th>
<th>Avatar ID</th>
<th>Guild</th>
<th>Level</th>
<th>Race</th>
<th>Class</th>
<th>Zone</th>
</tr>
</thead>
<tbody>
<tr>
<td>01/01/06 23:59:39</td>
<td>1</td>
<td>467</td>
<td>1</td>
<td>1</td>
<td>Orc</td>
<td>Warrior</td>
<td>Orgrimmar</td>
</tr>
<tr>
<td>01/01/06 23:59:39</td>
<td>1</td>
<td>921</td>
<td>19</td>
<td>1</td>
<td>Orc</td>
<td>Shaman</td>
<td>Orgrimmar</td>
</tr>
<tr>
<td>01/02/06 00:03:31</td>
<td>45</td>
<td>1367</td>
<td>8</td>
<td>60</td>
<td>Undead</td>
<td>Warrior</td>
<td>Arashi Mountain</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Field</th>
<th>Valid Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Query Seq. #</td>
<td>An integer ≥ 1</td>
</tr>
<tr>
<td>Avatar ID</td>
<td>An integer ≥ 1</td>
</tr>
<tr>
<td>Guild</td>
<td>An integer within [1, 513]</td>
</tr>
<tr>
<td>Level</td>
<td>An integer within [1, 80]</td>
</tr>
<tr>
<td>Race</td>
<td>Blood Elf, Orc, Tauren, Troll, Undead</td>
</tr>
<tr>
<td>Class</td>
<td>Death Knight, Druid, Hunter, Mage, Paladin, Priest, Rogue, Shaman, Warlock, Warrior</td>
</tr>
<tr>
<td>Zone</td>
<td>One of the 229 zones in WoW world</td>
</tr>
</tbody>
</table>
Basic Statistics
Daily Game Play Activities

![CDF Graphs](a) Playtime (hr) (b) Daily session count

<table>
<thead>
<tr>
<th>Activity</th>
<th>Mean, SD</th>
<th>Quantiles (5%, 25%, 50%, 75%, 95%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Session time (hr)</td>
<td>(2.8, 1.8)</td>
<td>(0.4, 1.0, 1.8, 3.0, 5.5)</td>
</tr>
<tr>
<td>Daily session count</td>
<td>(1.7, 0.9)</td>
<td>(1.0, 1.1, 1.4, 2.1, 3.3)</td>
</tr>
<tr>
<td>Daily play time (hr)</td>
<td>(3.7, 2.8)</td>
<td>(0.5, 1.6, 3.1, 5.1, 8.8)</td>
</tr>
</tbody>
</table>
Workload Analysis

Variability

- The number of players constantly fluctuates between 0 – 600 in each day.
- **High daily variability**

Regularity

- **Strong weekly and daily periodicity**

Predictability

- **Highly predictable based on the last hour**
- Prediction power is still high over adjacent weeks
Sample study #1

Player Unsubscription Prediction

Pin-Yun Tarng, Kuan-Ta Chen, and Polly Huang

ACM NetGames 2009 Poster
Unsubscription Prediction: Our Proposal

- Rationale: players’ satisfaction / enthusiasm / addiction to a game is embedded in her game play history
Prediction Accuracy

Accuracy (%) vs. d value (days)

- Longer than 2 years
- Longer than 1 year
- Longer than 0.5 year
- Shorter than 0.5 year

World of Warcraft Avatar History Dataset / ACM Multimedia Systems 2011
Sample study #2

Is Server Consolidation Beneficial to MMOG?

Yeng-Ting Lee and Kuan-Ta Chen

IEEE CLOUD 2010
Server Consolidation

- ... is an approach to the **efficient usage of computer server resources** in order to **reduce the total number of servers**.
- ... in response to the problem of **server sprawl**, a situation in which multiple, under-utilized servers take up more space and consume more resources than can be justified by their workload.

![Diagram showing server consolidation](image)

- Workloads without Virtualization:
  - Servers poorly utilized at average of 4% to 7% capacity
  - Limited in failover capability
  - Prone to hardware failure

- Workloads migrated to Virtual Machines Using Virtualization:
  - Each workload is now encapsulated stacking its workload for better hardware utilization – around 80%
  - Inherit virtualization capabilities include:
    - Dynamic resource pools
    - High availability without complicated clustering
    - Provision new servers in minutes
  - Virtual Machines are hardware independent
Motivations

- **Cost down**
  - Hardware investment (servers, network devices, cooling, space)
  - Administration (labor) costs
  - Energy saving

- **Elasticity & Agility**
  - Equipment sharing among different game shards and even game titles
Server Consolidation is good to MMORPGs: 3 Reasons

- Spatial locality property in players’ interaction
  - naturally partitionable

- Workload is highly variable but predictable
  - potential to aggregate workload in off-peak periods

- Operators normally run multiple games
  - possibility to share infrastructure
Simulation Setup

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simulated period</td>
<td>2 months</td>
</tr>
<tr>
<td>Server capacity</td>
<td>7500 avatars</td>
</tr>
<tr>
<td>Realm number ((r))</td>
<td>100</td>
</tr>
<tr>
<td>Server number ((s))</td>
<td>100</td>
</tr>
<tr>
<td>Zone per realm ((z))</td>
<td>83</td>
</tr>
<tr>
<td>Avatar number per realm (\mathcal{N}(2640, 1500^2))</td>
<td></td>
</tr>
<tr>
<td>Power consumption per server</td>
<td>300 W (idle), 750 W (peak)</td>
</tr>
<tr>
<td>VM allocation algorithm</td>
<td>First-Fit Decreasing</td>
</tr>
</tbody>
</table>

- An operator owns \(s = 100\) servers hosting \(r = 100\) realms of a game
- Each realm contains \(r = 83\) zones
- Supposing that a server is capable of serving 7,500 avatars
- Modeling the avatar number in a realm as a normal distribution with mean 2640 and standard deviation 1500, which is derived from the data set on Warcraft Census and Wow Database
- The avatar number in a zone is normalized using our traces from the TW-Light’s Hope realm
### Summary of Performance

<table>
<thead>
<tr>
<th></th>
<th>Single game</th>
<th>Multiple games</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$D_d$</td>
<td>$D_h$</td>
</tr>
<tr>
<td><strong>Server Investment</strong></td>
<td>84%</td>
<td>53%</td>
</tr>
<tr>
<td><strong>Energy Consumption</strong></td>
<td>89%</td>
<td>43%</td>
</tr>
</tbody>
</table>

†: $D_d$ denotes the Dynamic (day) strategy.
‡: $D_h$ denotes the Dynamic (hour) strategy.

- Server investment can be saved up to 52% while energy consumption can be reduced by 62%
- **Server consolidation is beneficial to MMOGs**
Use the WOWAH dataset, Luke!

Thank You!

Thanks to LucasArts