

# World of Warcraft Avatar History Dataset

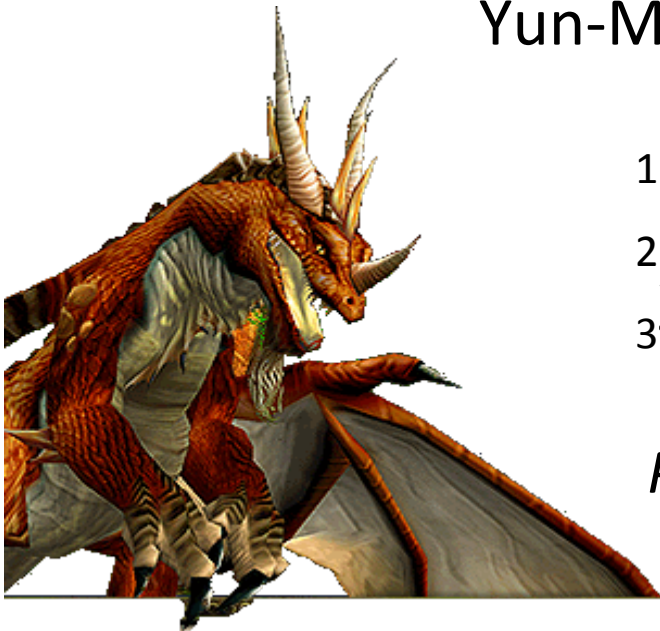
Yeng-Ting Lee<sup>1</sup> Kuan-Ta Chen<sup>1</sup>  
Yun-Maw Cheng<sup>2</sup> Chin-Laung Lei<sup>3</sup>

<sup>1</sup>National Taiwan University

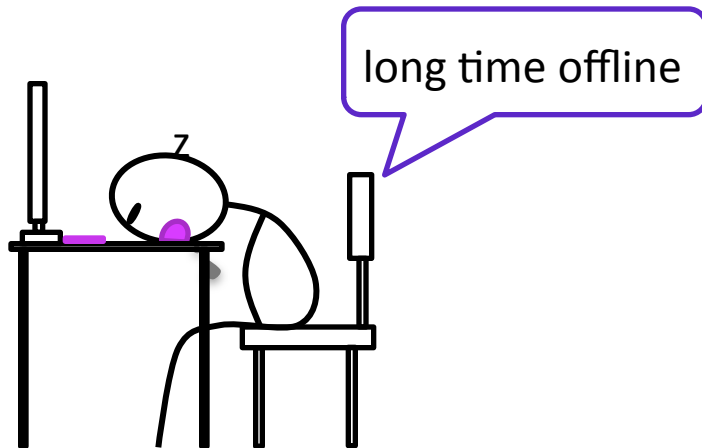
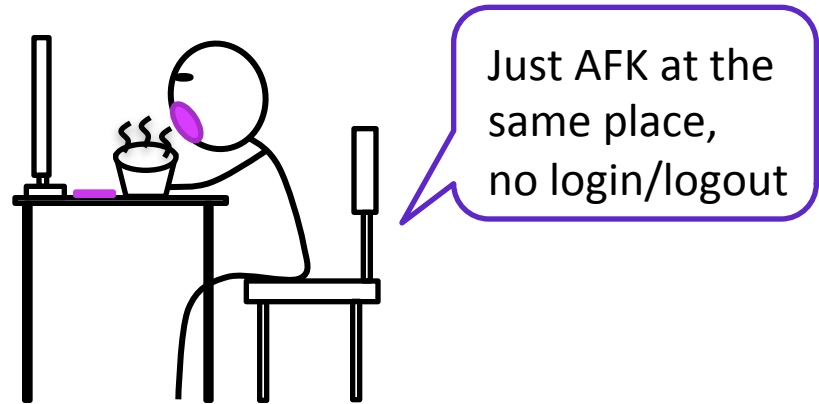
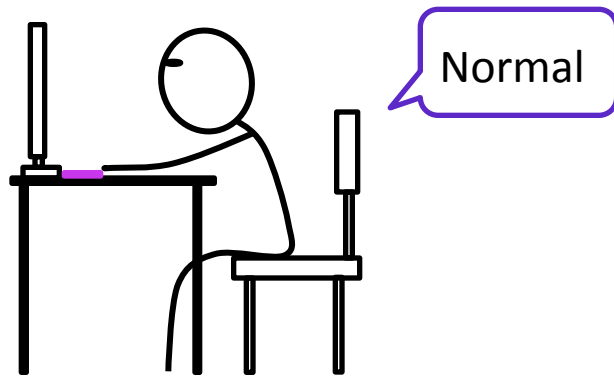
<sup>2</sup>Academia Sinica

<sup>3</sup>Tatung University

*Presented by **Cheng-Chun Tu**  
(Stony Brook University)*



# Diversity in Game Play Behavior



# 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
  - A three-year period dataset from Jan. 2006 to Jan. 2009.
- Data collection methodology
- Sample use of the dataset
  - Player unsubscription prediction
  - Feasibility of server consolidation

## Summary of WOWAH

Realm	TW-Light's Hope
Faction	Horde
Start date	2006-01-01
End date	2009-01-10
Duration	1,107 days
Sampling rate	144 samples per day
# of samples	159,408
# of missing samples	21,324
# of avatars	91,065
# of sessions	667,032

# World of Warcraft

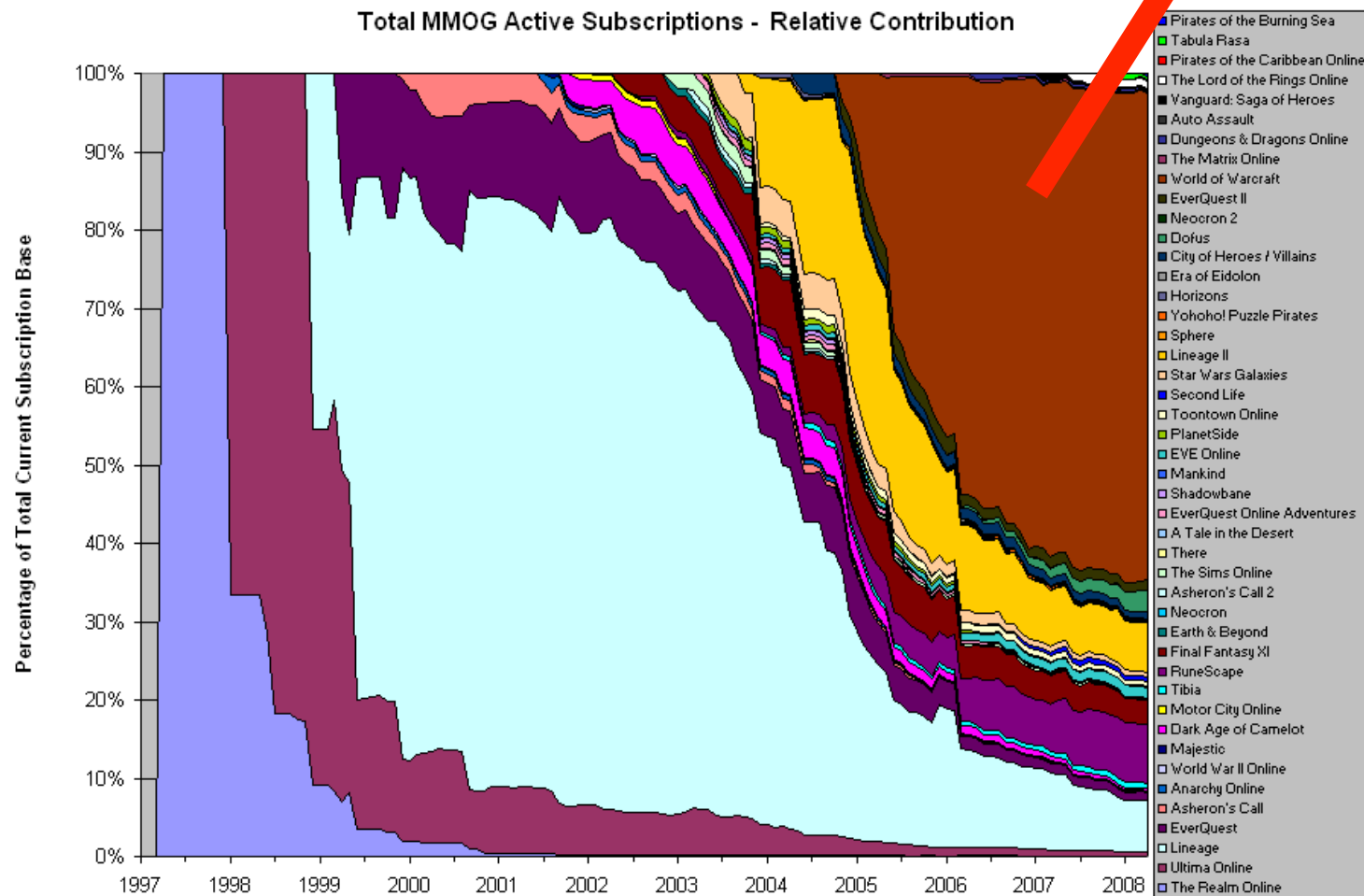


- WOW is the most popular MMORPG in the world, designed by BLIZZARD Entertainment.
- There are many players/avatars interacting, combating, chatting and dying within the battle field.





# World of Warcraft





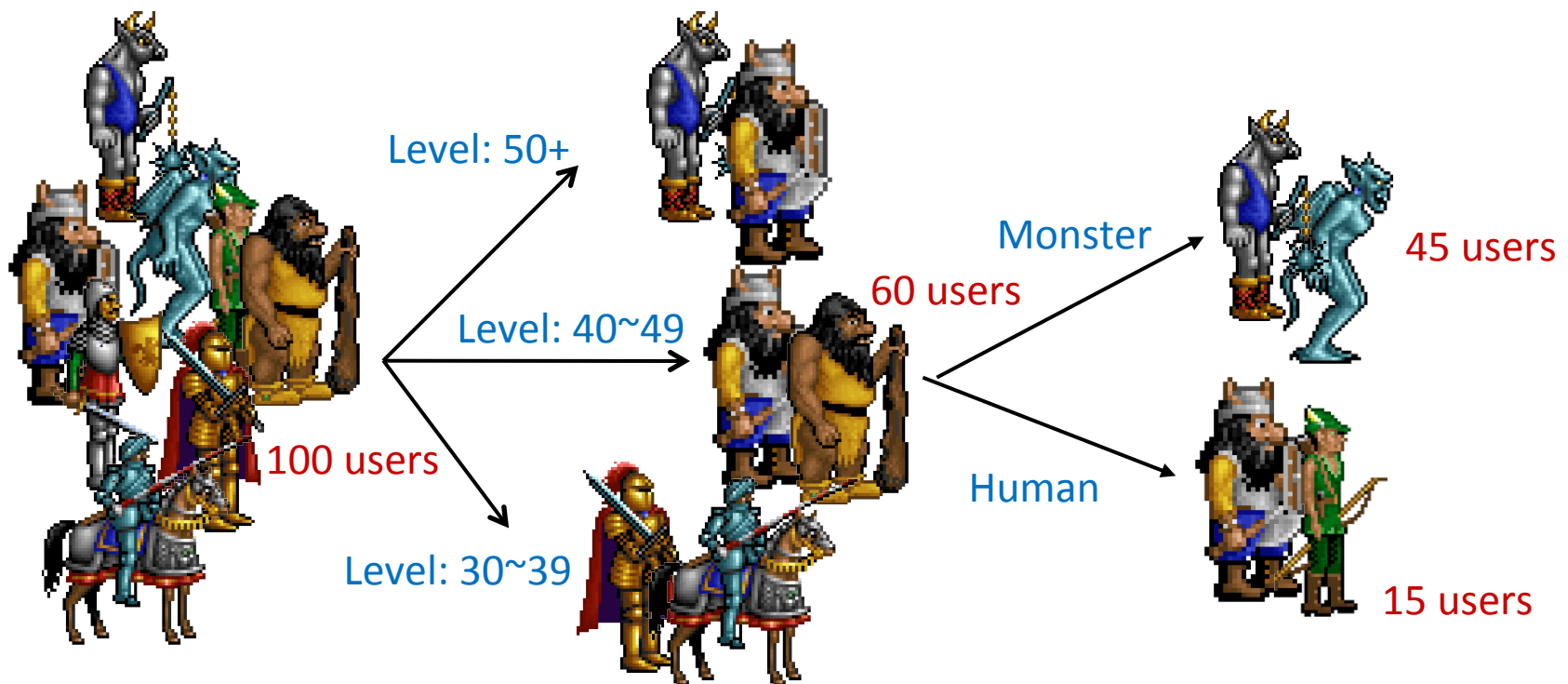
# Data Collection Methodology

- 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



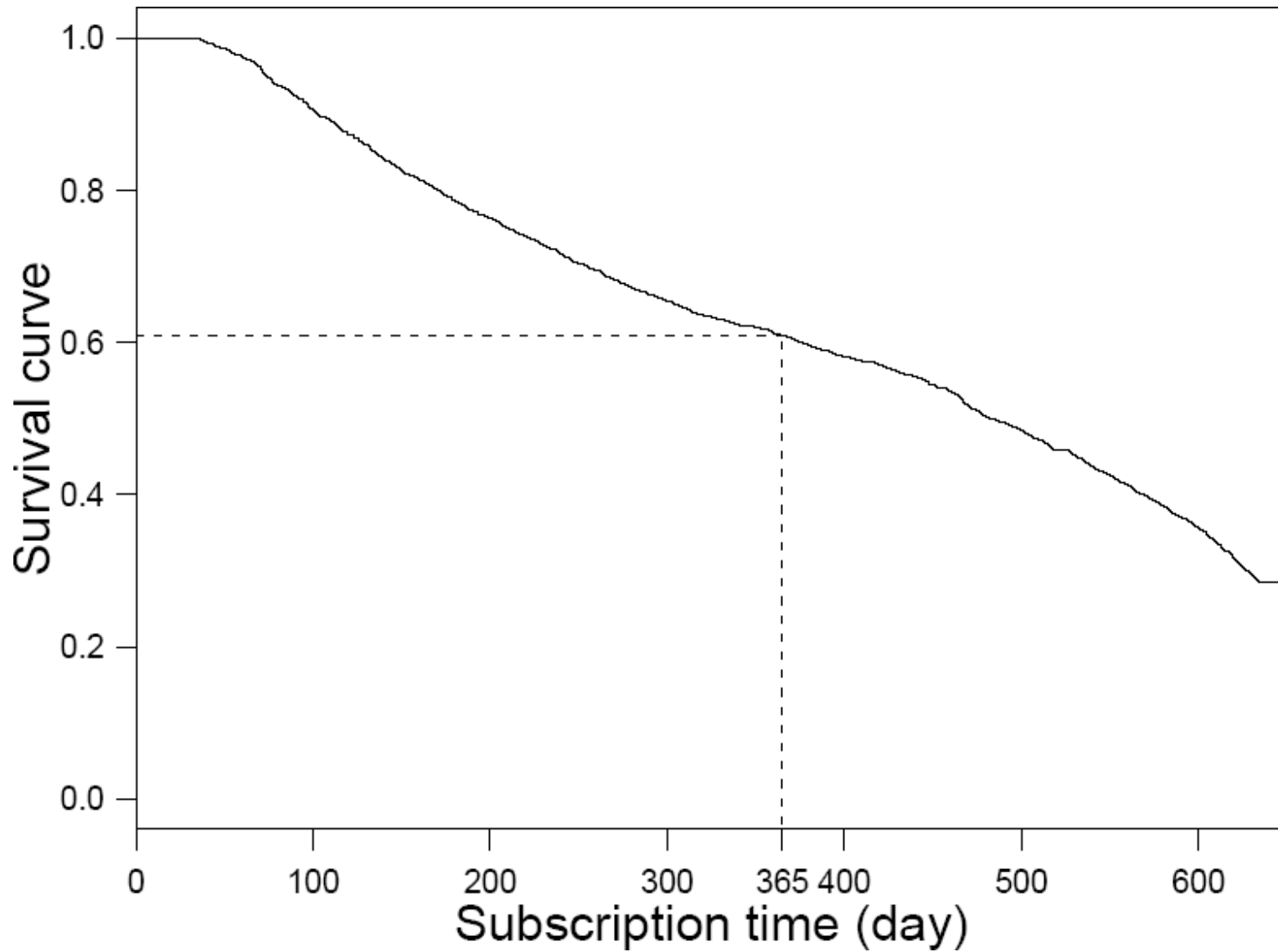
# Data Format

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

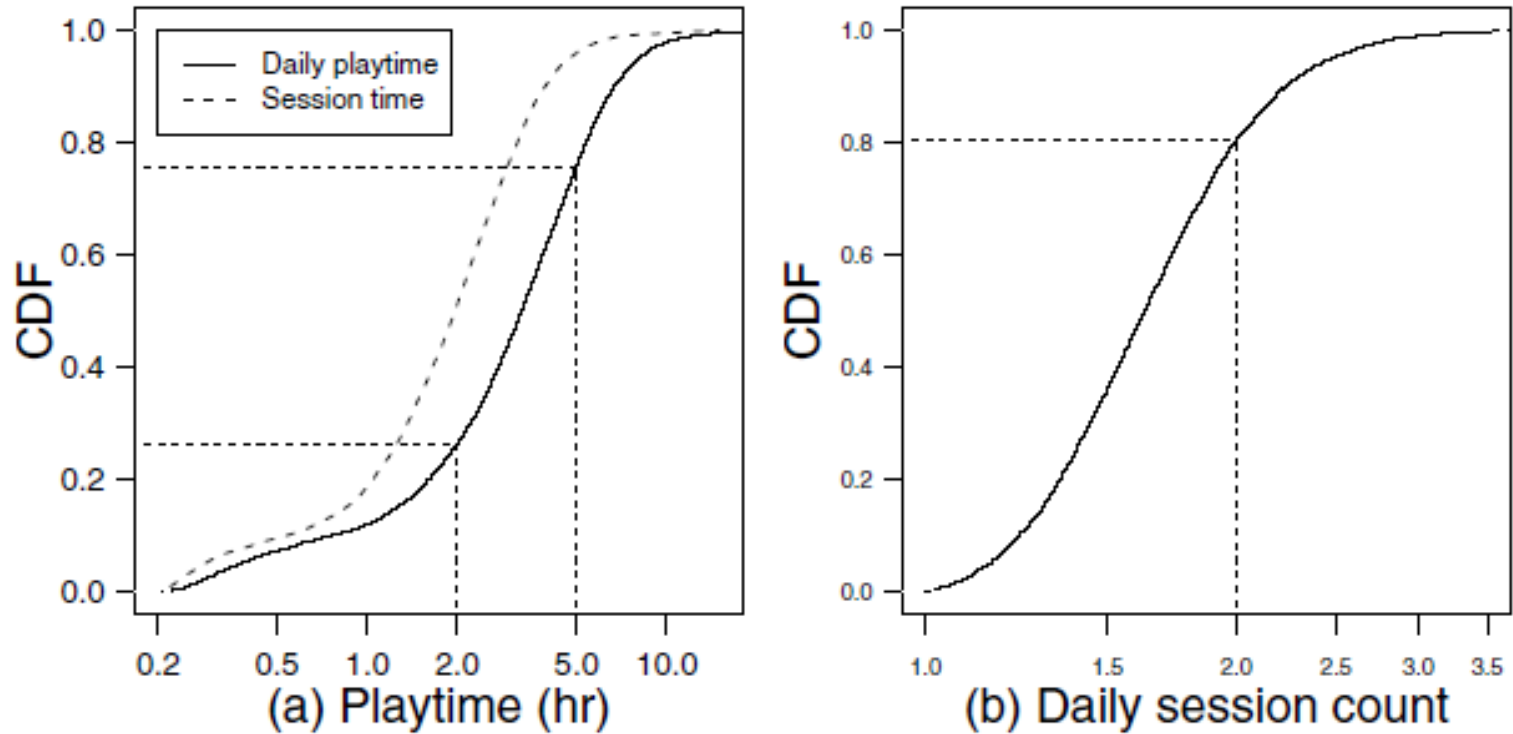
Field	Valid Values
Query Time	Between Jan. 2006 and Jan. 2009
Query Seq. #	An integer $\geq 1$
Avatar ID	An integer $\geq 1$
Guild	An integer within $[1, 513]$
Level	An integer within $[1, 80]$
Race	Blood Elf, Orc, Tauren, Troll, Undead
Class	Death Knight, Druid, Hunter, Mage, Paladin, Priest, Rogue, Shaman, Warlock, Warrior
Zone	One of the 229 zones in WoW world

# Basic Statistics

# Subscription Time



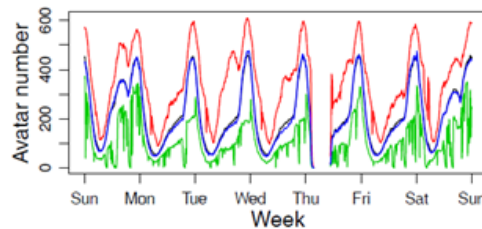
# Daily Game Play Activities



	(Mean, SD)	Quantiles (5%, 25%, 50%, 75%, 95%)
Session time (hr)	(2.8, 1.8)	(0.4, 1.0, 1.8, 3.0, 5.5)
Daily session count	(1.7, 0.9)	(1.0, 1.1, 1.4, 2.1, 3.3)
Daily play time (hr)	(3.7, 2.8)	(0.5, 1.6, 3.1, 5.1, 8.8)

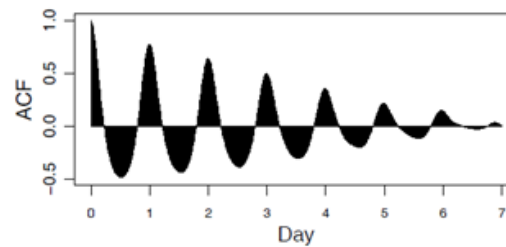
# Workload Analysis

## Variability



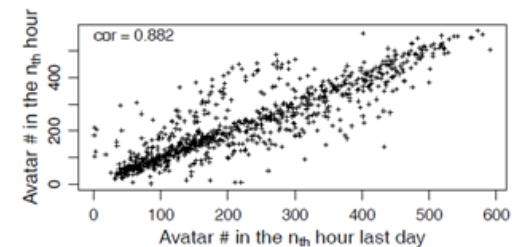
- The number of players constantly fluctuates between 0 – 600 in each day.
- **High daily variability**

## 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

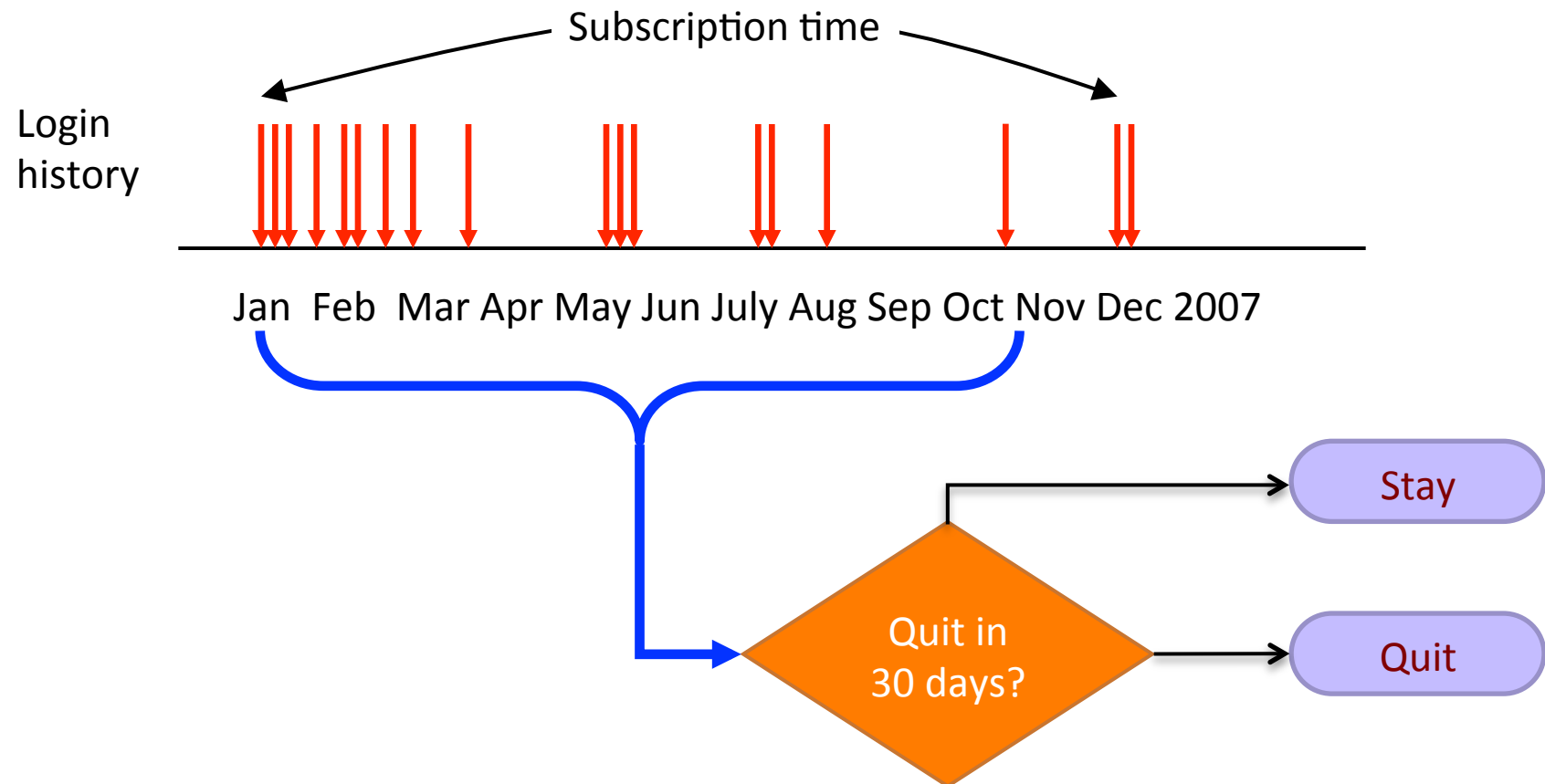
# Player Unsubscription Prediction

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

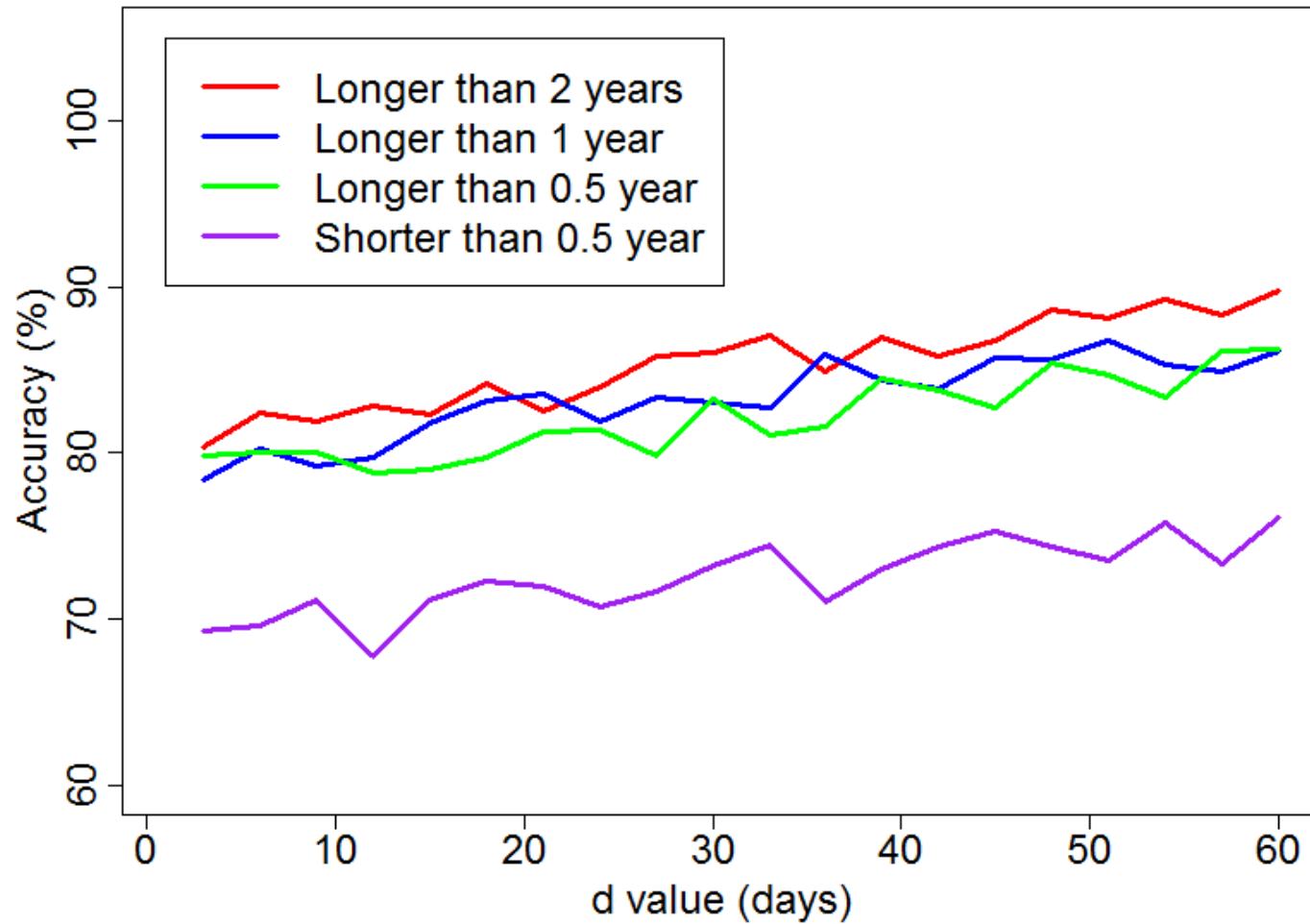
ACM NetGames 2009 Poster

# Unsubscription Prediction: Our Proposal

- Rationale: players' satisfaction / enthusiasm / addiction to a game is embedded in her **game play history**



# Prediction Accuracy



## Sample study #2

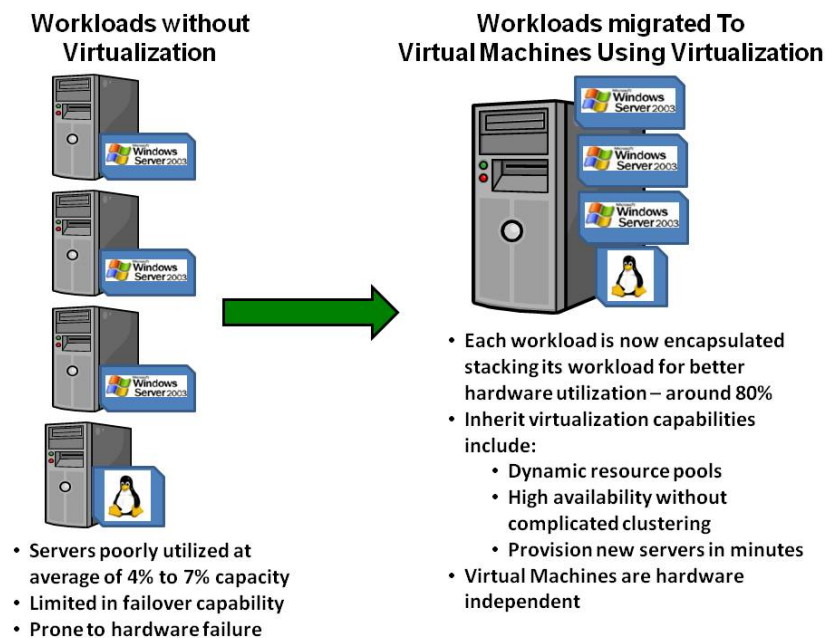
# Is Server Consolidation Beneficial to MMOG?

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.



# 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

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

- 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

	Single game		Multiple games	
	$D_d^\dagger$	$D_h^\ddagger$	$D_d$	$D_h$
Server Investment	84%	53%	91%	48%
Energy Consumption	89%	43%	84%	38%

$\dagger$ :  $D_d$  denotes the Dynamic (day) strategy.

$\ddagger$ :  $D_h$  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!



Thanks to LucasArts

***Thank You!***