Entertainment Computing xxx (2011) xxx–xxx

Contents lists available at ScienceDirect

Entertainment Computing

journal homepage: ees.elsevier.com/entcom

Analysis of revisitations in online games

Ruck Thawonmas [a,*], Keisuke Yoshida [a], Jing-Kai Lou [b], Kuan-Ta Chen [b]

[a] Intelligent Computer Entertainment Laboratory, Graduate School of Science and Engineering, Ritsumeikan University, Kusatsu, Shiga 525-8577, Japan
[b] Multimedia Networking and Systems Laboratory, Institute of Information Science, Academia Sinica, Taipei 115, Taiwan

ARTICLE INFO

Article history:
Received 4 June 2010
Revised 24 January 2011
Accepted 25 January 2011
Available online xxxx

Keywords:
Revisitation
Online game
Analysis
Support Vector Machine
Ward hierarchical clustering
Shen Zhou Online
World of Warcraft

ABSTRACT

This paper analyzes revisitations in online games focusing on two types of revisitations: game revisitations and area revisitations. A player revisits a game and areas therein with purposes. For game revisitations, we conduct a large-scale analysis using Shen Zhou Online access log collected for nearly 6 years consisting of 50,000 characters and have succeeded in using in the information on game revisitations, together with the login time and login frequency information, for predicting the players who will be absent from the game. For area revisitations, we conduct yet another large-scale analysis using World of Warcraft access log collected for 2 years consisting of more than 60,000 characters and have discovered four main groups of area revisitation patterns. We also discuss in the paper how our findings can be utilized to support both game developers and players.

1. Introduction

Analyses of revisitations in Web pages [1–3] are actively being conducted by the Web research community for understanding Web navigation, surfing behaviors, and revisitation intents. Inspired by those studies, we hypothesize that, similar to web pages, typical patterns should also exist in online-game players' revisiting to a game of interest and its areas. Such patterns would be useful for supporting both game developers and players.

In order to validate our hypothesis, we conduct two large-scale analyses using access log from Shen Zhou Online (SZO) and World of Warcraft (WoW), both categorized as Massively Multiplayer Online Role-Playing Game (MMORPG). The key component in both analyses is the revisitation curve [2] representing the number of times that the analysis target (a game, an area, or a group of areas) is revisited within each predefined time interval by a single player or a group of players of interest. We exploit the game revisitation curve and the area revisitation curve for predicting the SZO players who will be temporarily or permanently absent from the game and for investigating groups of WoW players who have similar area revisitation patterns, respectively.

The contributions of this work are (a) the first analysis of this kind on MMORPGs, (b) representative SZO game-revisitation patterns for the continuing players and the absenting players, (c) four typical WoW area-revisitation patterns, and (d) our implications to applications of the findings.

2. Revisitation curve and two sets of MMORPG access log

Here we define in Section 2.1 the key component in our work, i.e., the revisitation curve, and describe the SZO access log and the WoW access log in 2.2 and 2.3, respectively. Because our access log contains noisy data, such data must be removed before performing analyses. Different filtering criteria are applied to the two data sets because the analysis targets are different, i.e., revisitations to the game for the former and to in-game areas for the latter.

2.1. Revisitation curve

The revisitation curve of **A** to **B** represents the number of times that **A** revisited **B** within each predefined time interval (bin). Here, **A** can be a single player or a group of players; and **B** a game, an area, or a group of areas. As done in [2], in order to form logarithmic scale as well as the interpretable meaning for the time intervals, the following 13 bins are employed: 32, 64, 98, 136, 212, 424, 848, 1696, 3392, 6784, 13,568, 27,136 mins, and above. These intervals represent approximately, 30 min, 1 h, 1.5 h, 2 h, 3 h, 6 h, 12 h, 1 day, 2 days, 4 days, 8 days, 16 days, and above, respectively.

In our analyses to follow, we use the normalized version of the revisitation curve. This is to facilitate comparisons between the

* Corresponding author. Fax: +81 77 561 5203.
E-mail address: ruck@ci.ritsumei.ac.jp (R. Thawonmas).

1875-9521/$ - see front matter © 2011 International Federation for Information Processing Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.entcom.2011.01.003

players in the case of SZO and between the areas in the case of WoW. Normalization is done as follows.

For SZO,

$$\frac{\text{bin}[j] \text{ of user}[i]}{\text{average of bin}[j] \text{ among all users}},$$

where bin[j] of user[i] is the number of times user[i] revisited SZO within bin[j]'s interval, and, in the SZO analysis, user[i] is represented by a vector of 13 dimensions whose elements are determined according to the above ratio.

For WoW,

$$\frac{\text{bin}[j] \text{ of area}[i]}{\text{average of bin}[j] \text{ among all areas}},$$

where bin[j] of area[i] is the number of times area [i] was revisited by the WoW users within bin[j]'s interval, and, in the WoW analysis, area [i] is represented by a vector of 13 dimensions whose elements are determined according to the above ratio.

2.2. Shen Zhou Online

We use the access log of SZO provided to us by the SZO staff. SZO is popular in Taiwan [4], where there are thousands of players online at any time. A player can engage in a fight both against other players or creatures, train themselves for special skills, participate in marketplaces, or take on a quest. The SZO log contains the players' login and logout timestamps for 50,000 characters during the period of nearly 6 years. In the analysis section, we show how the revisitation curve derived from the access log of a given period can be used for predicting whether the player will be absent from the game.

In order to remove data of the players who only played during the 30-day free trial period, we filtered out all characters having the interval between the first logout and the last login less than 30 days. After this filtering, 3596 characters remained. For each remaining character, we divided its log length into multiple partitions of the length of 30 days such that any two consecutive partitions overlap for 15 days. We further removed every partition containing less than two logins. For each remaining partition, the intervals between each logout and the subsequent login in that partition are derived to form the revisitation curve of the corresponding character for the period that the partition covers.

2.3. Word of Warcraft

We use access log from a Taiwanese server of WoW. At present, WoW is arguably the most popular MMORPG [5]. WoW players must choose to belong to either Alliance or Horde faction from the beginning. We obtained access log of the Horde characters playing at the server by executing /who, at our client every 10 min during 2 years from January 2006 to December 2007. Before filtering, our log contains more than 60,000 characters of all Horde races: Blood Elf, Orc, Tauren, Troll, and Undead. An execution of /who gives us information on all currently-logging-in characters and their playing areas at the execution time; other information includes the race, job, and level of the corresponding character.

In order to increase the reliability of the data, we removed the characters with the total play time below 40 h, a typical RPG length. We further removed the characters with at least one of the four attributes falling below 10th percentile against other characters. These four attributes are (a) the total number of area transitions, (b) the total number of unique areas visited, (c) the total number of areas revisited, and (d) the area-transition ratio. The last one is the ratio between (a) and the log length of the character.

Table 1
Example of simplified WoW access log.

Time	Area
10/08/06 02:12:37	A
10/08/06 02:22:28	A
10/08/06 02:32:37	B
10/08/06 02:42:38	C
10/08/06 03:02:30	D
10/08/06 03:12:37	D
10/08/06 03:22:35	B
10/09/06 00:52:37	C
10/09/06 01:02:37	A
10/12/06 23:36:37	A

Table 1 shows an example WoW access log of a character. For this character, the aforementioned attributes are 6, 4, 3, and 0.6, respectively. Note that because this character consecutively visited area D, its total number of areas revisited is 3, not 4.

In order to exclude outliers, such as bots [6] and extremely hardcore players who spent an extraordinary amount of time in the game, we further filtered out those characters with the log length of 99th percentile against other characters. As a result, the number of remaining characters becomes 6491. For area filtering, we removed the areas with the integral of the revisitation curve of 1st percentile against other areas. This results in 107 areas remaining after filtering.

3. SZO analysis

Our hypothesis is that the game revisitation curve of a player before absent from the game is different from that of a continuing player. Therefore, the revisitation curve information can be used for predicting whether or not a player of interest will be absent from the game. We consider a period of *the* $(i+1)$*th login – the ith logout* of a given character, for any integer *i*, beyond 1 month an absent period of the character. For each character, we label its partitions, obtained in 2.2, into two types: *absent* and *normal*. A partition is defined as absent if its last logout timestamp is the beginning of an absent period; otherwise, defined as normal. As a result, we obtain 6101 absent partitions and 70,953 normal partitions; in this analysis, every character has both absent partitions and normal partitions.

In order to prove the above hypothesis, we first selected the representative revisitation curve of the absent partition and that of the normal partition from each character. The generalized median is used for each type; here it is the revisitation curve whose 13-dimensional vector has the smallest sum of the distances to the others of the same type. Fig. 1 shows the average representative curves for the absent and normal ones. This figure confirms that the game revisitation curve of a player before absent from the game is different from that of a continuing player.

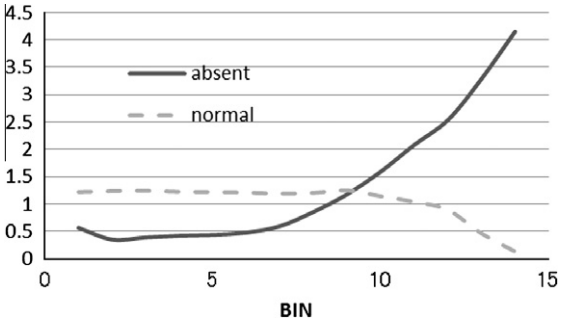


Fig. 1. Average representative SZO revisitation curves for normal and absent partitions.

Please cite this article in press as: R. Thawonmas et al., Analysis of revisitations in online games, Entertainm. Comput. (2011), doi:10.1016/j.entcom.2011.01.003

Now we show how the revisitation curve derived from the access log of a given period, 30 days in this paper, can be used for predicting whether or not the player will be absent from the game. For this task, Support Vector Machine (SVM) [7] is used. SVM is trained by a training data set consisting of both absent partitions and normal partitions. For comparisons, we represent each partition by the following three methods:

Method 1 – the 13 dimensional vector discussed in 2.1.
Method 2 – the 15 dimensional vector constructed by the 13 elements of the above vector and the element indicating the login time in that partition, normalized (divided) by the average login time for all partitions, as well as the element indicating the login frequency in that partition, also normalized by the average login frequency for all partitions.
Method 3 – the two dimensional vector consisting of the last two elements in Method 2.

After training, SVM is evaluated with a testing data set consisting also of both partition types. For evaluation, we use the following three performance indices:

$$\text{the recognition rate} = \frac{\text{the number of correctly classified partitions}}{\text{the total number of tested partitions}},$$

$$\text{the precision rate} = \frac{\text{the number of correctly classified absent partitions}}{\text{the total number of partitions classified as absent}},$$

and

$$\text{the recall rate} = \frac{\text{the number of correctly classified absent partitions}}{\text{the total number of tested absent partitions}}$$

Because the number of the absent partitions is about 12 times less than that of the normal partitions, we adopt a technique called undersampling which uses only a randomly selected subset for the majority class, in our case the normal one. We summarize the performance evaluation procedure as follows:

Step 1 Randomly select a subset of 6101 partitions from the whole normal set.
Step 2 Use the above normal subset together with the whole absent set in 10-fold cross validation executing the SVM tool [8] with the RBF-kernel.
Step 3 Repeat Steps 1 and 2 twelve times in order to obtain the average recognition, precision, and recall rates.

In the above procedure, the RBF-kernel was selected for the SVM tool in [8] because it gave the best recognition rate, compared with the other available kernels, in our preliminary evaluation using only the representative normal and absent partitions of all characters.

4. SZO results and discussion

Table 2 shows the performance indices of the three methods in our SZO analysis. Method 2 has the highest recognition rate and the precision rate. Its recall rate is slightly lower than Methods 1 and 3. Method 3, which does not use the revisitation curve information, has the best recall rate, but the lowest recognition rate and precision rate. Method 1, which uses only the revisitation curve information, has its performance indices between those of Methods 2 and 3. From these results, if the bottom line is set to 75%, we argue that the combination of the revisitation curve, the login time, and the login frequency, as used in Method 2, provides useful information for predicting players' future login trend.

We investigated two more cases where the characters were filtered out if the interval between the first logout and the last login is less than 90 days and 150 days, respectively. In the former case, the number of remaining characters is 2342; the number of absent partitions and the number of normal partitions are 5083 and 68,693, respectively. In the latter case, the number of remaining characters is 1798; the number of absent partitions and the number of normal partitions are 3954 and 64,485, respectively. For these two cases, we applied the evaluation procedure in the end of Section 3, with the size of the normal subset set to 5053 and repeating Steps 1 and 2 fourteen times for the former case and with the size of the normal subset set to 3954 and repeating Steps 1 and 2 sixteen times for the latter case. The same trends among Methods 1, 2, and 3 are also seen for these two cases (see Tables 3 and 4).

5. WoW analysis

Our analysis approach follows the recipe in [2]. First, we use the revisitation curve to characterize an area of interest by its 13-dimensional vector discussed in 2.1. Then we cluster these areas, according to their vectors into related groups using a hierarchical clustering algorithm.

To cluster areas into multiple groups, each having similar revisitation characteristic, we adopt commonly used Ward hierarchical clustering [9], available in R. Because shape information is important, the cosine distance is selected in Ward clustering. The index in use for automatically deciding the number of clusters is maximization of

$$\frac{\text{interclass variance}/(\text{number of clusters} - 1)}{\text{intraclass variance}/(\text{number of areas} - \text{number of clusters})}$$

[10]. Because we anticipate that there should also be four main groups as in [2], the minimum number of clusters is set to four.

6. WoW results and discussion

Fig. 2 shows the clustering results where all areas are divided into 15 clusters. Each cluster is labeled to one of the four groups, i.e., Fast, Medium, Slow, and Hybrid, based on the revisitation curve's shape of the cluster centroid. Table 5 summarizes the cluster information for Fast, Medium, and Slow; and Table 6 summa-

Table 2
Performance indices of the three methods in the SZO analysis for the characters with the log length above 1 month.

Method	Recognition rate	Precision rate	Recall rate
1	0.75	0.73	0.79
2	0.77	0.77	0.78
3	0.72	0.69	0.80

Table 3
Performance indices of the three methods in the SZO analysis for the characters with the log length above 3 months.

Method	Recognition rate	Precision rate	Recall rate
1	0.75	0.73	0.78
2	0.77	0.76	0.77
3	0.73	0.71	0.78

Please cite this article in press as: R. Thawonmas et al., Analysis of revisitations in online games, Entertainm. Comput. (2011), doi:10.1016/j.entcom.2011.01.003

Table 4
Performance indices of the three methods in the SZO analysis for the characters with the log length above 5 months.

Method	Recognition rate	Precision rate	Recall rate
1	0.75	0.73	0.78
2	0.76	0.76	0.77
3	0.73	0.71	0.78

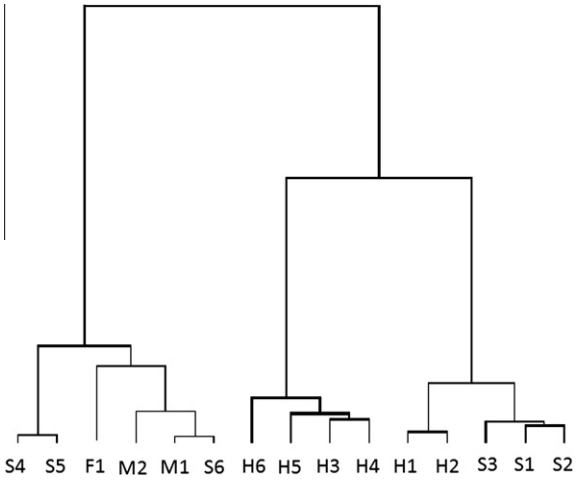


Fig. 2. WoW hierarchical clustering results.

rizes the cluster information for Hybrid. Each group is discussed in detail in the following.

6.1. Fast

The areas in this group are often revisited within a short interval, say, less than hour. They can be classified into two categories: battlegrounds and starting areas.

6.1.1. Battleground

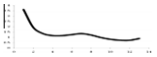
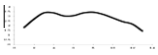
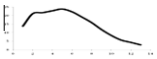
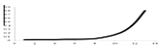
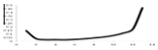
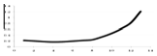
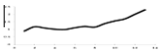
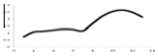
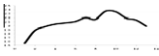
One of the main features of the Warcraft universe is the raging conflict between Horde and the Alliance. Battlegrounds are instanced areas solely created for Player-versus-Player (PvP) combats. The two initial Battleground areas, Alterac Valley and Warsong Gulch, went online on June 7th, 2005. Arathi Basin was introduced in Patch 1.7 on September 13th, 2005. The fourth battleground, named Eye of the Storm, was released with the Burning Crusade extension pack.

Battlegrounds are instances and have a limit on the number of players who can enter at the same time, where the limit varies in different battlegrounds. The game provides Battleground Queuing System to assist players register for joining battlegrounds given the player number limits. Players can wait to enter a battleground anywhere in the game world via the game's user interface. Once it is the turn for a player to enter the battleground, he or she will be teleported into the battleground, and teleported back after the battle finishes. All the battlegrounds are classified as fast-revisiting areas, which indicates that a battle is usually finished in 30 min and PvP players often reenter a battleground immediately for the next match.

6.1.2. Starting area

During our data collection period (January 2006 to December 2007), an extension pack called "The Burning Crusade" was released in March 2007. In this extension pack, there are two new races introduced: The Draenei for the Alliance and the Blood Elves for the Horde. New Blood Elf characters start their journey in the Eversong Woods, the northern section of Quel'Thalas. Farther south is the Ghostlands, a level 10–20 zone. Low-level Blood Elves gain experience points and reward items by pursuing the quests in Eversong Woods and Ghostlands. However, most of the quest givers are in the capital city, Silvermoon (S6), which is located in the northeastern part of the Eversong Woods. Therefore, low-level Blood Elves have to regularly return to Silvermoon to complete their quests or to take new quests. After that, they tend to go back to the starting areas and carry on their new quests. The fast revisitation of starting areas confirms that the quests for low-level players are relative easy and can be done within a short time.

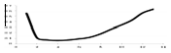
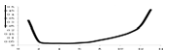
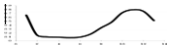
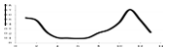
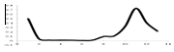
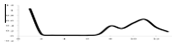
Table 5
Summary of the WoW cluster information for Fast, Medium, and Slow.

Group	Cluster	Centroid curve	Number of areas	Area names
Fast	F1		7	Alterac Valley, Arathi Basin, Eversong Woods. Eye of the Storm, Ghostlands. Maraudon. Warsong Gulch
Medium	M1		8	Blade's Edge Mountains, Hellfire Peninsula, Nagrand, Netherstorm, Shadowmoon Valley, Terokkar Forest, Undercity, Zangarmarsh
	M2		2	Orgrimmar, Shattrath City
Slow	S1		6	Duskwood, Elwynn Forest, Loch Modan, Redridge Mountains, Twisting Nether, Wetlands
	S2		5	Blackfathom Deeps, Ragefire Chasm, Razorfen Downs. Razorfen Kraul, Shadowfang keep
	S3		11	Blasted Lands, Burning Steppes, Darkshore, Dustwallow Marsh, Hall of legends, Moonglade, Mulgore, Scarlet Monastery, Swamp of Sorrows, The Temple of Alal'Hakkar, Wailing Caverns
	S4		14	Alterac Mountains, Arathi Highlands, Ashenvale, Azshara, Badlands, Desolaee, Feralas, Hillsbrad Foothills, Silverpine Forest. Tanaris, The Hinterlands, Thousand Needles. Thunder Bluff. Un'Goro Crater
	S5		6	Blackrock Mountain, Deadwind Pass, Durotar. Gates of Aln'Qiraj, Searing Gorge, Tirisfal Glades
	S6		9	Eastern Plaguelands, Felwood, Silithus, Silvermoon City, Stonetalon Mountains, Stranglethorn Vale, The Barrens, Western Plaguelands, Winterspring

6.2. Medium

Areas in this cluster have a peak in the left middle of the curve, i.e., between 1 h and 1 day. We find that the areas in this cluster are due to two possible reasons as follows.

Table 6
Summary of the WoW cluster information for Hybrid.

Group	Cluster	Centroid curve	Number of areas	Area names
Hybrid	H1		11	Dire Maul, Hellfire Ramparts, Scholomanee, Shadow Labyrinth, Stratholme, The Areatraz, The Black Morass. The Botanica, The Shattered Halls, The Steamvault, The Underbog
	H2		10	Auchenai Crypts, Blackrock Depths, Gnomeregan, Ma na-Tombs, Old Hillsbrad Foothills, Sethekk Halls, The Blood Furnace, The Slave Pens, Uldaman, Zul'Farrak
	H3		5	Blackrock Spire, Naxxramas, Ruins of Ahn'Qiraj. The Mechanar. Zul'Gurub
	H4		3	Blade's Edge Arena, Nagrand Arena, Ruins of Lordaeron
	H5		5	Gruul's Lair, Magtheridon's Lair, Molten Core, Onyxia's Lair, Tempest Keep
	H6		5	Ahn'Qiraj, Blackwing Lair, Karazhan, Serpentshrine Cavern, Zul'Aman

6.2.1. Daily quest

Daily quests (often called "dailies") are repeatable quests firstly introduced in Patch 2.1. Areas for daily quests are included in M1, such as Etherstorm and Shadowmoon Valley. These quests provide significantly better rewards than regular quests. A player can complete up to 25 daily quests in each day. After a player solves 25 daily requests, he or she must wait for the "daily quest reset" event. The exact time for such reset depends on the configuration of each realm, which is often chosen in the early morning. Thus, it is reasonable for players to revisit the areas which provide daily quests between 1 h and 1 day.

6.2.2. Bank and auction house

The other areas in the Medium cluster are capital cities, including Undercity (M1) for the Undead, Orgrimmar (M2) for the Orc and Troll, and Shattrath city (M2) for the united force. The cities serve as the main cultural, political, and economical center of each race. To players, one of the main functionalities of the capital cities is their economical functions, i.e., banking and auction. Banks and auction houses are usually located only in major cities or towns, especially in capital cities. A bank is a building with teller NPCs which allow a player to access his own private storage. A player can freely deposit and withdraw his items and currency in the storage via a bank or its branches. An auction house is a place for players to trade goods, weapons, and equipments. As of the release time of Patch 1.9, each capital city has its own auction house. The medium revisitation patterns of capital cities imply that players tend to regularly come back to their respective capital cities to access their private storage or trade goods with others.

6.3. Slow

The revisitation curves in this cluster have a slow peak, i.e., revisitation intervals are longer than 1 day. The level 30–70 areas are included. The revisitation curves of growth areas for moderate and upper level (30–70) characters, included in S4, S5, S6, have higher amplitudes than those of lower-level growth areas.

6.3.1. Low-level instance

These instances are special areas where a group of players can interact privately within a dungeon. Instance dungeons tend to feature the most difficult and rewarding content. Since the majority of the players have achieved the 70th level in our traces, the low-level (20–30 level) instances, such as Blackfathom Deeps, Ragefire Chasm, Razorfen Downs, Razorfen Kraul, and Shadowfang Keep (all in S2), can no longer offer satisfying rewards. Such low-level instances thus become less popular. Therefore, players do not revisit the areas frequently.

6.3.2. Monthly event

Another reason for slow revisitations is due to a monthly event, the Darkmoon Faire (the Faire), which is a regular event firstly introduced in Patch 1.6. The Faires always start on the first Monday of a month, and the location rotates among three places, namely, Elwynn Forest (S1), Mulgore (S3), and Terokkar Forest (M1). The event lasts for 1 week and appears in turn in the three places each month. Because some special goods can only be purchased in the Fairs, players may manage to attend such events, even the location is far away from their homelands. For example, when the Faire takes place in the Elwynn Forest near the Alliance land, we find that some Horde players traveled through Dust Wood, Loch Modan, Redridge Mountains, Twisting Nether, and Wetland (all in S1) in order to get to Elwynn Forest. Because these in-between areas are unpopular and may only be visited due to such rare events, those areas exhibit slow revisitation patterns and are included in the slow revisitation cluster.

6.4. Hybrid

6.4.1. High-level instance

High-level instances feature the most difficult and rewarding content, thus such instances are very popular. For example, Stratholme, Old Hillsbrad Foothills, Blackrock Spire, Molten Core, and Ahn'Qiraj which associate with H1, H2, H3, H5, and H6, respectively, are in this category. Due to the challenge in solving quests offered in these areas, very often the entire party of players are eliminated by the enemies and transferred to the nearest land outside the instance. This kind of situation is called "wipeout". After a wipeout, players usually gather and enter the same instance within 10 min, which causes the left peak in the revisitation curve.

On the other hand, in order to control the number of valuable items from inflating, an instance may have a "lockout" period after it is conquered. During the lockout period, the players who completed the instance are not allowed to re-enter the instance. The lockout period is often set to 1 day or 1 week depending on the scarcity of the reward. This design leads to the right peak in the revisitation curves for these areas.

6.4.2. Arena

The remaining areas in the Hybrid cluster are mostly arenas, including Blade's Edege Arena, Nagrand Arena, and Ruins of Lordaeron (all in H4). An arena is a place in which teams of players compete against each other to gain the Arena Points in deathmatch-like player-vs-player combats. Since such matches often conclude in 10 min, players tend to reenter an arena frequently. This is the reason for the left peak.

However, the Arena Points are not awarded after each match. For example, one requirement to gain the Arena Points is to attend at least ten matches within a week, and the Points is awarded on Thursdays regularly. Therefore, players tend to reenter an arena

Please cite this article in press as: R. Thawonmas et al., Analysis of revisitations in online games, Entertainm. Comput. (2011), doi:10.1016/j.entcom.2011.01.003

to gain rewarding points weekly. We consider this design the cause to the right peak in the revisitation curve.

7. Design implications

Design implications are given in this section. Although SZO and WoW are used in the analyses in this paper, we note here that our implications are applicable to other MMORPGs.

7.1. Player support

For the players who are predicted that they will be absent from the game, the game publisher should perform some measures to prevent them from actually doing so. Possible measures include (i) sending to those players a promoting message stating that they can earn special points if they login within a specific time, (ii) inviting them to join an in-game event or an out-of-game event where only players who share similar game play preference and strategies participate, and (iii) asking them their opinions about the game and fixing the located problems. How to find similar players, mentioned in ii), based on the area revisitation information is discussed in the next subsection.

Individual player support can also be provided based on area revisitation characteristic of a player of interest. For example, assume that the revisitation curves of a character of interest to battle areas exhibit Slow, but those to growth areas are Fast. This indicates that the player does not favor fighting against other players, but rather prefers to raise his/her character's level. The game system can thus give to this player personal supports such as a direct invitation to an event at a growth area. In addition, the game system can provide to a player a kind of reminder service that reminds the player to visit an area whose interval with no login since the last visit has passed the peak time in the revistation curve.

If an area search function is available in the game, information on area revisitation can be used in area retrieval computation. For example, assigning more preferences to Slow areas, than to Fast or Medium areas, in the search result might give players higher satisfaction. This is because we think there is a higher possibility that players use the area search function to locate less known areas, such as those in Slow group.

7.2. Players' social network solidification

Players are encouraged or even restricted to visit the areas suitable for their characters. Because the contents in the areas designed for high-level characters are mostly guarded by powerful monsters, it is far too risky for low-level characters to take benefit from these areas. For that reason, the higher capability of characters the more areas they can explore. Consequently, the revisitation patterns of the characters with different capability can be very distinct due to the large number of areas in a game (107 areas in our WoW dataset).

Although the characters share the similar areas to visit, their revisitation patterns can still be various due to players' game strategies or preference. For example, a group of players may prefer to boost the game levels through slashing monsters in caves rather than simply chatting with other players in towns. On the other hand, some other players may prefer accumulating valuables by manufacturing goods, providing services, or trading; thus they tend to spend a lot of time in market places and rarely get into battlefield areas. Via the clustering of area revisitation patterns, we are able to perform player clustering, where a cluster contains players whose preference and strategies mentioned above are similar.

Game operators can exploit this kind of information to solidify the social network between gamers by designing systems like "friend recommendation" or by holding social events that encourage players to know about each other. For players who share similar game play strategies, they are more likely to connect to and share with each other, which will establish social relationships between those players and further solidify the overall social network of the game's participants. Consequently, this will increase the stickiness to the game and furthermore prolong the game's lifetime.

7.3. Area design reconsideration

Through the analysis of players' area revisitation patterns, game designers can check whether the usage of areas conforms to their expectation. Specifically, some areas may be unreasonably popular where players revisit them very frequently, while some others may be unexpectedly unpopular where players seldom revisit. The unexpectedly popular areas might be because players have found some tricks that can defeat "boss" monsters or gain a great quantity of reward, whereas the unexpectedly unpopular areas might be because the game missions provided are too difficult or the reward is not worth enough. Thus, we consider area revisitation patterns could be a useful summarization tool of game players' passion about certain areas, and, at the same time, a useful analytical tool for game designers to reconsider the design of game areas.

In addition, to maintain high user satisfaction, the contents of Fast and Medium areas should be updated with higher frequencies than those of Slow. Fast and Medium areas are more popular ones where players more frequently visit. Neglect in updating them may lead to player retirement.

8. Related work

Since an early report [11] on a large amount of re-visit information in web navigation, many studies [1–3,11–15] have been conducted to understand webpage revisitation and browsing behavior. Compared to those studies, the work in [2] that we base on is the largest study in terms of the number of participants. More recently, the same group of authors has furthered their work and summarized their recent findings on the association between change in content and revisitation in [3].

SZO was the target game platform for the work in [16], which is the first formal analysis of MMORPG server traces. The work in [17] uses the information on average daily playtime and playing density for predicting of SZO player departure. The method therein is similar to Method 3 in this paper and thus does not use the revisitation curve information.

WoW has been used as a research platform in many studies because of its popularity and availability of WoW API [18], allowing interaction with and modification of the game. In [19,20], automatically collected data were explored to understand gaming experiences and social dynamics, respectively. The predictability of online-game players' subscription time and the effect of changes in game features to social interaction were studied in [21,22], respectively. More recently, a system for automatically generating comics from WoW gaming experiences [23] and a zone-based server consolidation strategy [24], evaluated using a subset of extensive WoW traces dicusssed in [25], have been developed.

9. Conclusions

Inspired by an existing work on analysis of web revisitation patterns, in this paper, we analyzed revisitations in SZO and WoW. For SZO, we focused on the use of the game revisitation curve in order

Please cite this article in press as: R. Thawonmas et al., Analysis of revisitations in online games, Entertainm. Comput. (2011), doi:10.1016/j.entcom.2011.01.003

to predict players who will be absent from the game and succeed in doing this task, with the additional information on the login time and the login frequency. For WoW, we focused on the use of the area revisitation curve in order to understand player groups of similar area revisitation patterns and discussed four area groups: Fast, Medium, Slow and Hybrid, divided based on their revisitation curves. A number of design implications were given that support not only the game-developer side but also the game-player side. Extensive tests of these implications are left as our future work.

Acknowledgments

This work was supported in part by Japan Society for Promotion of Science (JSPS) under the Grant-in-Aid for Scientific Research (C) 20500146 as well as by the National Science Council under the grant NSC97-2221-E-001-009 (Taiwan).

References

[1] H. Obendorf, H. Weinreich, E. Herder, M. Mayer, Web page revisitation revisited: implications of a long-term click-stream study of a Long-term Click-stream Study of Browser Usage, in: Proc. of the 25th Annual SIGCHI Conference on Human Factors in Computing Systems (CHI'07), 2007, pp. 597–606.
[2] E. Adar, J. Teevan, S. Dumais, Large scale analysis of web revisitation patterns, in: Proc. of the 26th Annual SIGCHI Conference on Human Factors in Computing Systems (CHI'08), 2008, pp. 1197–1206.
[3] E. Adar, J. Teevan, S. Dumais, Resonance on the web: web dynamics and revisitation patterns, in: Proc. of the 27th Annual SIGCHI Conference on Human Factors in Computing Systems (CHI'09), 2009, pp. 1381–1390.
[4] UserJoy Technology Co., Ltd.: Shen Zhou Online, <http://www.ewsoft.com.tw/index.php>.
[5] Wikipedia: World of Warcraft, <http://en.wikipedia.org/wiki/World_of_Warcraft>.
[6] R. Thawonmas, Y. Kashifuji, K.T. Chen, Detection of MMORPG Bots based on behavior analysis, in: Proc. of ACM SIGCHI International Conference on Advances in Entertainment Technology 2008 (ACE 2008), 2008, pp. 91–94.
[7] B. Scholkopf, A.J. Smola, Learning with Kernels, MIT Press, Cambridge, MA, 2002.
[8] E. Dimitriadou, K. Hornik, F. Leisch, D. Meyer, A. Weingessel, CRAN – Package e1071, <http://cran.r-project.org/web/packages/e1071/>.
[9] J.H. Ward, Hierarchical Grouping to optimize an objective function, Journal of American Statistical Association 58 (301) (1963) 236–244.
[10] T. Calinski, J. Harabasz, A dendrite method for cluster analysis, Communications in Statistics 3 (1) (1974) 1–27.
[11] L.D. Catledge, J.E. Pitkow, Characterizing browsing strategies in the World-Wide Web, in: Proc. of the Third International World-Wide Web conference on Technology, Tools and Applications (WWW'95), 1995, pp. 1065–1073.
[12] L. Tauscher, S. Greenberg, How people revisit Web pages: empirical findings and implications for the design of history systems, Interntional Journal of Human–Computer Studies 47 (1) (1997) 97–137.
[13] A. Cockburn, B. McKenzie, What do Web users do? An empirical analysis of Web use, International Journal of Human–Computer Studies 54 (6) (2001) 903–922.
[14] E. Herder, Characterizations of user Web revisit behavior, in: Proc. of Workshop on Adaptivity and User Modeling in Interactive Systems (ABIS05), 2005, pp. 32–37.
[15] M. Kellar, C. Watters, M. Shepherd, A goal-based classification of Web information tasks, in: Proc. of the Annual Meeting of the American Society for Information Science and Technology (ASIS&T 2006) (CD-ROM Proceedings, 2006).
[16] K.-T. Chen, P. Huang, C.-L. Lei, Game traffic analysis: an MMORPG perspective, Computer Networks 50 (16) (2006) 3002–3023.
[17] P.-Y. Tarng, K.-T. Chen, P. Huang, On prophesying Online Gamer Departure, in: Proc. of ACM NetGames, 2009 (poster).
[18] WoWWiki: World of Warcraft API, <http://www.wowwiki.com/World_of_Warcraft_API>.
[19] N. Ducheneaut, N. Yee, E. Nickell, R.J. Moore, Building a MMO with mass appeal: a look at gameplay in World of Warcraft, Games and Culture 1 (4) (2006) 281–317.
[20] D. Williams, N. Ducheneaut, L. Xiong, Y. Zhang, N. Yee, E. Nickell, From tree house to barracks: the social life of guilds in World of Warcraft, Games and Culture 1 (4) (2006) 338–361.
[21] P.-Y. Tarng, K.-T. Chen, P. Huang, An analysis of WoW players' game hours, in: Proc. of the Seventh Annual Workshop on Network and Systems Support for Games (NetGames 2008), 2008, pp. 47–52.
[22] V. Chen, H. Duh, R. Hong, The changing dynamic of social interaction in World of Warcraft: the impacts of game feature change, in: Proc. of ACM SIGCHI International Conference on Advances in Entertainment Technology 2008 (ACE 2008), 2008, pp. 356–359.
[23] C.-J. Chan, R. Thawonmas, and K.-T. Chen, Automatic storytelling in comics: a case study on World of Warcraft, in: CHI Extended Abstracts 2009, 2009, pp. 3589–3594.
[24] Y.-T. Lee, K.-T. Chen, Is server consolidation beneficial to MMORPG? A case study of World of Warcraft, in: Proc. of 2010 IEEE Third International Conference on Cloud Computing (CLOUD), 2010, pp. 435–442.
[25] Y.-T. Lee, K.-T. Chen, Y.-M. Cheng, C.-L. Lei, World of Warcraft avatar history dataset, in: Proc. of ACM Multimedia Systems 2011 (MMSys 2011), 2011.

Please cite this article in press as: R. Thawonmas et al., Analysis of revisitations in online games, Entertainm. Comput. (2011), doi:10.1016/j.entcom.2011.01.003