MMORPG Player Behavior Model based on Player Action Categories

Mirko Suznjevic, Ivana Stupar, Maja Matijasevic
Faculty of Electrical Engineering and Computing, University of Zagreb, Zagreb, Croatia

mirko.suznjevic@fer.hr, ivana.stupar@fer.hr, maja.matijasevic@fer.hr

Abstract— In this paper we present the modeling of player behavior for Massively Multiplayer Online Role-Playing Games (MMORPGs). We have performed action specific measurements of player sessions in terms of the previously defined action categories for MMORPGs (*Trading, Questing, Dungeons, Raiding, and Player versus Player Combat*). We explore the hourly trends in user behavior and form models based on observed patterns. We explore the session duration as well as lengths and probability of session segments (i.e., parts of the session consisting of only one category of player actions). Our aim is to create the user behavior model and to combine it with previously established traffic models in order to be able to explain, predict, and generate network traffic of MMORPGs more accurately. As a case study we use World of Warcraft.

I. INTRODUCTION

Massively Multiplayer Online Role-Playing Games (MMORPGs) are in the recent years becoming significant sources of Internet traffic. Millions of people flock to MMORPGs on a daily basis. While inside networked virtual worlds, players perform various activities such as exploring the world, gathering virtual items, fighting with computer controlled entities or other players, etc. The player behavior tends to be complex, but nevertheless certain patterns exists. Models of player behavior are invaluable tools for more accurate traffic estimations and load balancing algorithms.

In this paper we are exploring the behavior of MMORPG players using *World of Warcraft* (WoW) by *Activision Blizzard* as a case study. We model the player behavior based on categories of player actions (*Trading, Questing, Dungeons, Raiding, and Player versus Player Combat*). In our previous works we have defined these action categories and performed measurements of player behavior [1], [2]. While those measurements were performed at the start of the 1st WoW expansion *The Burning Crusade*, we performed additional measurements during the 2nd expansion *Wrath of the Lich King* [3], but in time frame in which there was no recent addition of new content to the game (i.e., all present content in the game was more than a month old). Modeling performed in this paper is based on those measurements.

The proposed player behavior model is based on *session segments*. As session segment is identified as a part of the playing session in which the player performs only actions from one specific category. We model the segment duration, examine how often session segments of specific type appear, and determine the transition probabilities between categories.

We have also modeled the length of a session, while the data about the player population across the day has been taken from the literature. A 24 hour period is modeled in order to demonstrate hourly patterns. Our goal is to create a player behavior aware traffic generator. We aim to achieve this through combining the player behavioral model based on action categories with previously identified traffic models for each of the categories [4].

The remainder of this paper is structured as follows: in Section 2 we present the related work, followed by a brief description of WoW and our action categories in Section 3. Measurement methodology and data preparing process is explained in Section 4. Section 5 contains the presentation of the behavior models, followed by the implementation and results in Section 6, and conclusion of the paper in Section 7.

II. RELATED WORK

We now briefly summarize the findings in the related research work in the area of player behavior in MMORPGs. Player behavior is usually characterized by using session lengths, distribution of players through the virtual space, and number of active players.

Tarng, Chen, and Huang [5] analyze WoW player's game hours, on a trace sample gathered for almost two years' time with the use of an add-on for WoW client which polled all active players on one server. They investigate subscription time, daily and weekly patterns, and consecutive play days. Additionally, they try to predict player behavior for both short and long term. Lee and Chen [6] use the same strategy to gather data. They study the hourly, daily, and weekly patterns of player numbers on one WoW server in Taiwan. They propose a server consolidation strategy which uses the number of players and spatial locality to lower the hardware requirements and energy usage. Additionally, they present a dataset of measurements in their following work [7]. Those measurements were taken in a 1107-day period between Jan. 2006 and Jan. 2009. The dataset includes the avatars game play times and a number of attributes, such as avatars's in-game race, profession, current level, and in-game locations.

Zhuang et al. [8] gather data with help of a similar WoW add-on. Dataset created in five months of measurements is used to calculate player count, availability, downtime, inter-arrival times, and geographical distribution of players in the virtual world. They note that the stay time of the player in

certain area of the virtual world depends on the nature of the selected area, which they classify into *Transit*, *City*, and *Quest*.

Pittman and GauthierDickey [9] try to model the virtual populations and behavior of WoW players. They identify four facets needed for a complete model: population changes over time, arrival rates and session duration of players, spatial distribution of players over the virtual world, and movements of players over time. They state that artificial workloads using uniform distribution of players differ significantly from the situation which is observed in their measurements. In their following work [10] they develop behavior models for WoW and *Warhammer Online* (WaR). They prove that models for two MMORPGs with different play styles can be modeled by using a unified set of functions, through modeling session lengths (Weibull distribution), player distribution through zones (Weibull distribution), number of zones visited related to session lengths (linear dependence), time spent in zones (Weibull distribution), and player movement (Log-normal distribution).

Not only WoW is examined, but also other MMORPGs. Feng, Brandt, and Saha [11] study the trace of MMORPG *EVE Online* which was provided to them by *CCP* - the publisher of the game. They focus on examining daily and weekly patterns of player load, growth of player population over time, and impact of adding new content on play time. Also, they try to perform a prediction of the overall workload and prediction of players disinterest in the game.

Social interaction in MMORPGs has been proven to have significant impact on player behavior. Chen and Lei [12] studied player interaction in MMORPG *ShenZhou Online* and stated that social interaction has strong relationship with players' game time.

Kawale, Pal and Srivastava [13] study the problem of player churn in *EverQuest 2*. They propose a modified diffusion model for the social influence on the player, which takes into account player engagement, based on activity patters. They show that their model increases churn prediction accuracy for their dataset when compared with convential diffusion model for the player engagement.

We approach examination of player behavior somewhat differently, as we determine what exactly players are doing in the virtual world and for how long, as opposed to examining the characteristics of overall session time. Our measurements have been taken on the client side of each participating player through our WoW add-on deployed on players' computers [14]. This approach has certain advantages compared to remote monitoring. First, our precision is the highest as compared to the works of others (i.e., 1 second as opposed to 10 and 5 minutes). Second, accuracy is improved since we can exactly determine the actions taken by a specific player regardless of the in-game virtual character.

III. ACTION CATEGORIES IN WOW

WoW uses a client–server architecture, in which the virtual world is replicated on multiple "shards". A shard is a copy of a part or the whole virtual world, which resides on a specific server. In our modeling we focus on the behavior of players located on a single shard. Based on several key characteristics (e.g., number of actively participating players, dynamics of player input, number of active Non-Player Characters (NPCs) etc.), we have defined following action categories for MMORPGs:

Trading: Exchange and creation of virtual goods, between two players directly or through auction system. Single player non combat activity.

Questing: Performing different tasks set by NPCs for specific rewards (e.g., experience and virtual goods). Mostly single player activity, but there are also quests which require a group effort. This activity often involves combat.

Dungeons: Combat between a small player group and hostile NPCs, in specific instances (i.e., isolated portions of the virtual world which are replicated for each group of players). Instances do not allow interruption or help from players outside the specific group. This is a primary activity for small groups.

Raiding: Fighting among large, organized group of players and more difficult and complex NPCs. This category is similar to *Dungeons*, but it is larger on all scales. As the complexity of the task increases, so does the value of the prizes, therefore *Raiding* yields the best rewards.

PvP combat: Combat between players with very low number or in complete absence of NPCs. Player count may vary significantly from just a few to tens of players.

For more details on action categories specification and properties an interested reader is referred to [1], [2].

IV. METHODOLOGY

A. Data gathering

The measurements were done within a student project in which the participating students got extra credit for the course. The task of each student was to involve five or more active WoW players who agreed to participate in this research and install our WoW add-on named WoW Session Activity Logger (WSA-Logger). The total number of players in the survey was 104. No personal data about the participating players other than their age was gathered. This approach for gathering participating players resulted in an (on average) younger player sample than the general population of MMORPG found by Williams et al. [15] (average age 24 compared to average age 33, respectively). Our monitoring period was from May 5, 2009 to June 21, 2009, in a 3.x version of WoW (i.e., expansion "Wrath of the Lich King"). The biggest obstacle met with respect to the use of the with distributing the add-on was the trust of the players who feared for the safety of their accounts.

WSA-Logger creates a log file consisting of the events fired by the WoW API when a certain action is performed in the virtual world. For example, the *Auction_House_Close* event is fired once the auction interface (used for trading) is closed in the virtual world and the WSA-Logger notes the date, time, and player action type as *Trading*. Timestamps of the events are created with a 1 second precision.

B. Data analysis, processing, and filtering

For the analysis of log files created by WSA-Logger, we developed a log file parser in Java. This was needed in order to keep the add-on which runs on players computers, as simple as possible. The timeouts (i.e., time periods in logs with no events fired) longer than 5 minutes are labeled as *Uncategorized*. After the logs were parsed, the data about each session segment, comprising of its start time, end time, player designation, type, and session number were stored into a MySQL database. For the purpose of visualization of the gathered data, an Ajax based web tool was developed, which displays the data characteristics. This tool was used for visual data filtering, for easier identification of any anomalous behavior and outliers in the data. Detailed information about the functionality of WSA-Logger, data gathering, and filtering processes can be found in our previous works [1], [3].

V. MODEL

We model the player behavior by using Markov chains, distributions of action specific segments durations, distributions of session duration, and number of active players (described through initial player count, and arrival and departure processes). Parameters are calculated for each of the 24 hours in the day to capture a dominant hourly behavior pattern.

A. Player count

Number of active players on the WoW server has been studied in several related works [8], [6], [9], [10]. We do not repeat the measurement process of the player population during the course of the day, as this subject is well covered in the literature. We model the arrival and departure processes for new players as a *Homogeneous Poisson Process* (HPP) based on results of Chen et. al. [16] who stated that arrival process can be modeled as HPP for each 1-hour interval. While their analysis is done on a different game – *ShenZhou Online*, the authors stated that the results are applicable for other MMORPGs. Also, we model arrival and departure processes with rates for each 10 minutes as the estimation of an hourly rates resulted in higher underestimation of the number of users. For estimating the values of player load, arrival and departure processes we use the dataset provided by Lee et al. [7]. It should be noted that only one faction in WoW is monitored in that dataset (while there are two factions, namely *Alliance* and *Horde*) so for the full server model the rate of arrivals and starting population should be doubled (assuming equal representations of the factions on a given server).

B. Session length

Session length in MMORPGs can be described in two manners, as a *character session* (i.e., duration of presence of specific character in the virtual world), and as a *player session* (i.e., one player's subsequent use of multiple characters). This classification is not always emphasized in the literature which leads to quite different results regarding session duration. Most of the measurements use polling active number of players online which results in description of character sessions [7],
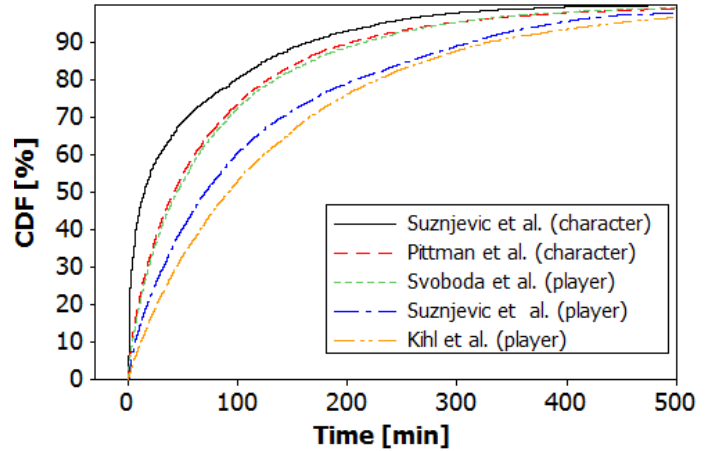

Figure 1: Probability of specific session segments through the hours of the day

[5], [9], while session duration estimates from network traffic measurements [17], [18], as well as our measurements taken on players' computers [3] measure the duration of player session. Additionally, polling measurement approaches can introduce an error as sampling a player base each 10 minutes intervals results in the shortest session time at 10 minutes. The difference resulting from different measurement techniques and different definition of the session is illustrated by Figure 1. The session duration is shorter in our first work [1] as measurements had 1 minute precision (30% of the session are shorter than 5 minutes) and were character based. It should be noted that the player sample in those measurements is very small. Pitman measures character based sessions [10] with a 15 minute sample time, while Svoboda [18] models player sessions based on a traffic trace of a mobile network which have shorter session times compared to wired networks. Longer session times are reported in our previous work [3], as we measure player session with 1 second measurement precision with additional parsing to ignore disconecting. The longest sessions are reported in the work by Kihl et al. [17] in which authors report player session length based on the results of traffic measurements in the wired accsess network. It is clear that character based estimates of session duration are significantly shorter than the real session duration.

We define the (player) session length as the time passed between player login and his logout from the game. If the time between a given players logout and a new login is less than 5 minutes we treat those two subsequent sessions as one session. In this way we identified 5872 player sessions, which comprised 11775 character based sessions.

We assume that the session length also follows a daily pattern, in particular, the sessions in the morning and over the day are lasting shorter than the sessions in the evening. This assumption is based on the measurements performed in our previous work [3], in which we noted that *Raiding*, an action category with the longest average duration shows a strong daily pattern with high rise in the evening, corresponding

with the availability of the average player. As it was noted by Williams et al. [15], average age of the MMORPG player is 33, as opposed to the common misconception that mostly teenagers play MMORPGs, so most of the players follow a daily routine of work or school over day which results in their availability in the late afternoon/evening.

In order to identify the underlying distribution we used the least-squares method. Previous works [10], [18], identified that session lengths conform to Weibull distribution, and we confirm those results. We also prove that session lengths are heavily dependent on the time of the day. In Figure 2 the Cumulative Distribution Functions (CDFs) of the sessions started between 1:00-2:00 and 19:00-20:00 are shown. Due to the space constraints, we do not present models for every hour, but only for two selected time frames between 1:00-2:00 and 19:00-20:00. Sessions started in the evening are on average a lot longer than sessions started late in the night. Also, sessions started in the evening do not conform so well to the Weibull distribution as those started later in the night.

The models for session length are as follows.

- For sessions started between 1:00 and 2:00:
$$f(x) = 61.8 * (-ln(x)^{(1/0.9327)})$$

- For sessions started between 19:00 and 20:00:
$$f(x) = 155.7 * (-ln(x)^{(1/0.9414)})$$

where x is uniformly distributed random variable $0 =< x =< 1$.

C. Session segment probability

Player behavior in terms of specific action categories shows a strong daily pattern due to availability of an average player [1], [3]. Probability of a segment of specific action category appearing during the day presented in Figure 3. The obvious trend is that group based activities (*Raiding* and *Dungeons*) have a raising trend from morning towards the evening, while *Questing* as a main single player activity has a decreasing trend. *Trading* and *PvP Combat* have more or less the same
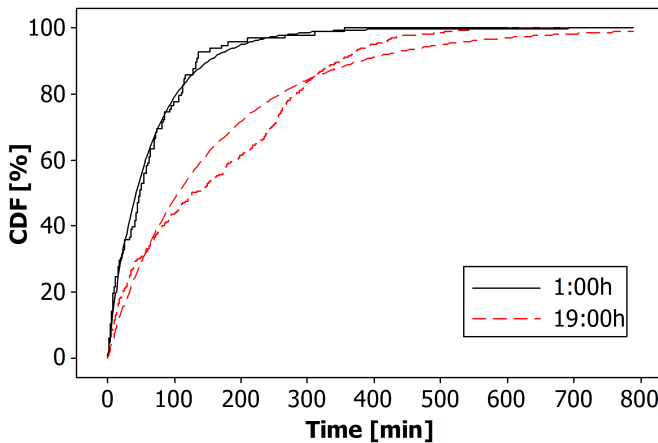
percentage of segments during the day. To explain this further we must understand the "battlegrounds" system in WoW. Players enlist for the battle and the game system automatically creates a player group consisting of random players which might even not be from the same server. In this way battles are always accessible, and players do not have to organize themselves as they have to for *Raiding*. A similar system was introduced recently for the *Dungeons* category, (after our measurements were performed). As we plan to perform additional measurements in order to validate our model, we expect that due to this change, the *Dungeons* will show the trend similar to that of *PvP Combat*. Also, our estimate is that *PvP Combat* will show a trend similar to *Raiding* as recently rated battlegrounds have been introduced, in which organized (i.e., not randomly assembled) groups of players compete in order to achieve higher ranks on the player scoreboards.

In order to model the player behavior throughout the day we construct a first order Markov chains for each hour of the day as we assume that the next action is only dependent on the previous action (e.g., after a dungeon or a raid, it is typical for a player to go sell items he/she does not need, or upgrade the newly obtained ones). In Figure 4 transition probabilities between states for hour 20:00-21:00 are depicted. Each action category is modeled as a single state of the Markov chain and is marked with its starting letter. Transitions with probabilities lower than 2% are not shown in order to simplify the figure.

D. Session segment length

In the next step we investigate to which distributions the length of action specific segments conforms, and whether the time of the day has an influence onto the length of the specific segment. Through inspecting CDFs of hour specific segments we conclude that *Questing* and *Trading* segment durations are not significantly dependent of the time of the day (i.e., differences between values of the CDF for different parts of the day were within 10%). *PvP Combat* has more differentiated CDFs, similar to *Dungeons*, but still the differences amongst hours of the day are rather small (i.e., different values of CDF within 20%). Based on this information we decided to
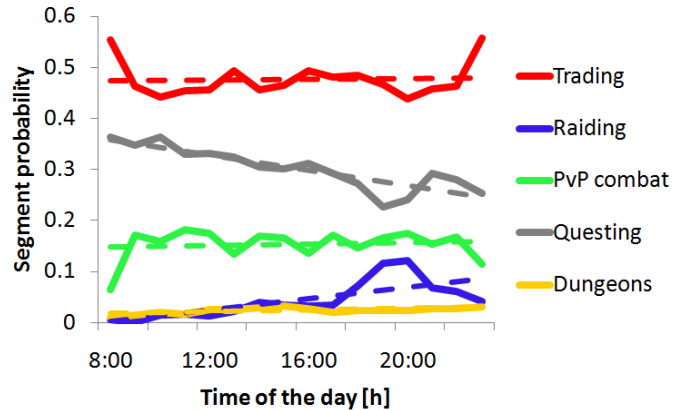

Figure 2: CDFs with Weibull fits of the session length started at specific hour (19:00-20:00 and 1:00-2:00)

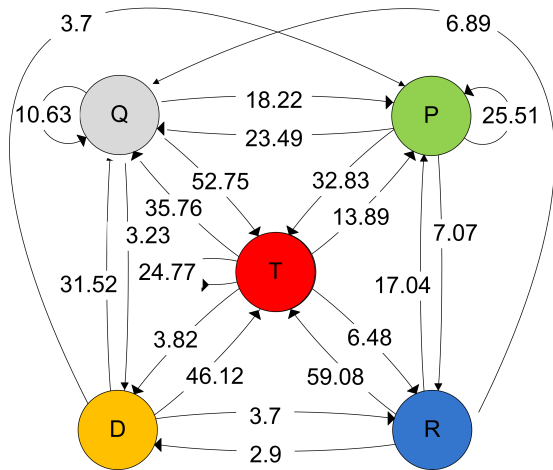Figure 3: Probability of specific session segments though hours of the day

Figure 4: Transition probabilities between states in hour 20:00-21:00 (Q–Questing, P–PvP Combat, R–Raiding etc.)
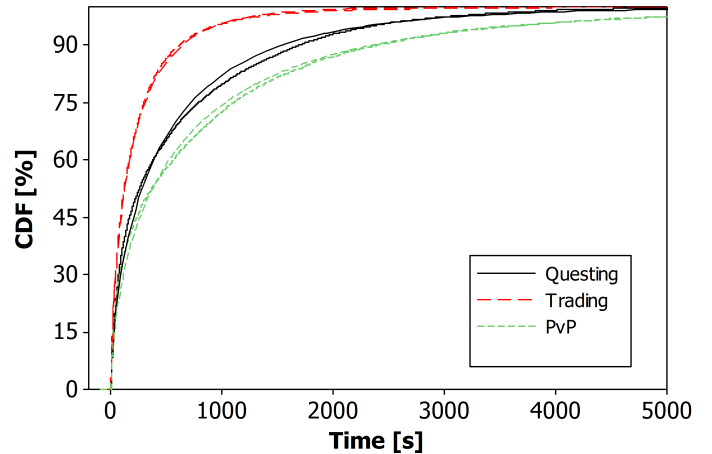

Figure 5: Fits of Questing, Trading, and PvP Combat

model session segments of those four action categories as independent of the time of the day. The segments durations models are provided in Table I with the values goodness of fit tests (Anderson Darling statistic and the P-value). The P-Values are showing that even the best fits are not closely following the empirical distribution but it should be noted that these tests are biased for large messy datasets [19]. To graphically describe goodness of fit we plot the empirical dataset and the fit on Figure 5.

On the other hand, *Raiding* shows significant differences in the CDF based on the time of the day. It is noted that the longest sessions are in the evening, when players can afford long uninterrupted time periods (e.g., some *Raiding* sessions span few hours). We decided to model *Raiding* segments' duration with five different distributions based on the time, one for the hours 18:00-19:00, 19:00-20:00, 20:00-21:00, 21:00-22:00, and one distribution for the remainder of the day. Also, those distributions proved to be harder to model, so modeling with split distributions is applied with a limit at 7500 seconds. Models for *Raiding* segment duration are depicted in Table II.

VI. IMPLEMENTATION AND RESULTS

We developed a tool in Java for simulating user behavior. Input parameters are initial player count, arrival and departure processes, session segment length, and session segment probabilities. These parameters are specified in the input XML file. The simulation may run on daily and weekly time basis based on the provided input data. Simulation can be run in real time, or in simulated time which can be accelerated up to 1500 times. Each player is represented by an individual thread,

so the arrival of new players to the system is simulated by creating new threads according to defined HPP. The sequence of user's activities during a session is a result of the Markov chain model described in Section 5. The progress of the simulation can be observed on a graphical user interface through a dynamically plotted graph showing the number of active players (y axis) in the specific moment (x axis) of the simulation. The output of the simulator consists of three log files containing data about player session lengths, the number of active players in the system (taken every ten minutes in simulated time) and data about simulation progress (timestamps for every action of arriving and departing user, as well as timestamps marking a transition of user to a new action category). The simulator also concurrently outputs player identifier, action category the player is entering, and the duration of the new segment. These output parameters are generated during the simulation after users change their state (i.e., enter a new session segment). Thus, the simulator can be used to feed data to other tools (e.g., for traffic generation, or server load simulation).

In order to compare the results of the simulation with the dataset by Lee et al. [7] on which the model is based, we plot the average number of users in the dataset over the results of the simulation for 1 day (in Figure 6). The simulation tends to slightly underestimate the number of users, but captures the main trends such as high decline in the late night, slow increase of players in the morning, and a steep increase in the evening.

TABLE I: Session segment duration models

	Fit	Shape/Loc.	Scale	AD	P
Trading	Weibull	0.69	190.98	14.70	<0.01
Questing	Weibull	0.66	440.01	59.93	<0.01
PvP Com.	Weibull	0.61	603.57	28.23	<0.01
Dungeons	LEV	1321.42	1116.00	4.80	<0.01

TABLE II: Raiding models for different hours of the days

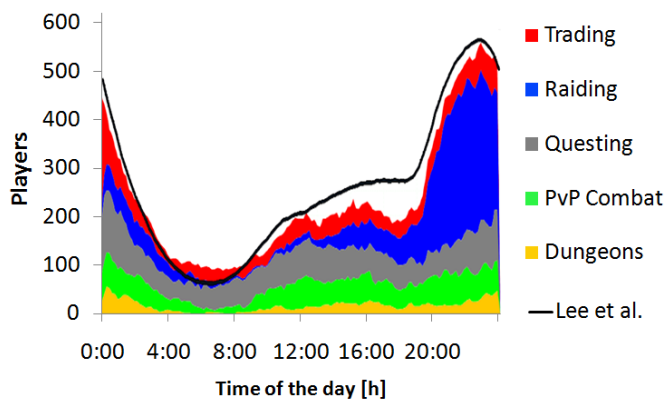| Hour	Data	Fit	Shape/Loc.	Scale	AD	P-value
18-19	42%	Weibull	1.175	2513	0.405	>0.250
	58%	Weibull	6.110	15306.5	1.722	<0.010
19-20	44%	Weibull	1.121	2583.76	0.663	0.085
	56%	Lognorm.	9.4173	0.227	1.016	0.011
20-21	42%	Weibull	1.279	3303.08	1.18	<0.010
	58%	Lognorm.	9.287	0.19762	0.768	0.045
21-22	100%	Weibull	1.368	6476.70	0.625	0.102
22-18	100%	Weibull	1.046	2786.5	5.209	<0.010

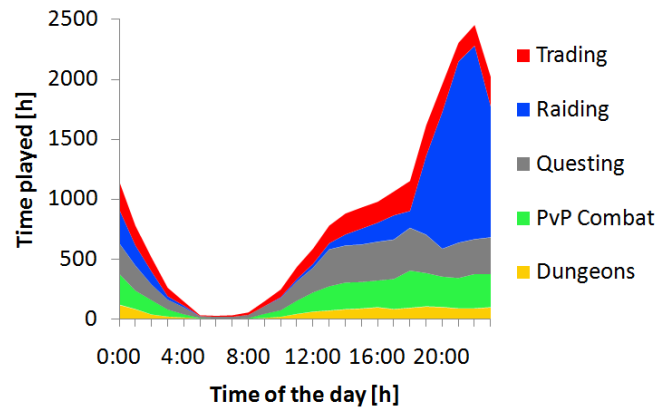Figure 6: Active players per action category (simulated)

Figure 7: Time played per action category (measured)

We evaluate the goodness of fit of the player behavior model through inspection of share of each action category through hours of the day. Comparing results of the simulation with the data from our empirical measurements depicted in Figure 7, it can be observed that the hourly behavioral pattern is captured for all categories.

VII. CONCLUSION

In this paper we have investigated characteristics of player sessions in MMORPGs based on player action categories. We have identified and modeled daily trends in session duration, as well as the action specific session segment duration. We determined how frequent the session segments pertaining to given action categories occur in specific times of the day, and created a model based on the first order Markov chain. We implement the proposed model of player behavior in Java–based tool. In future work we aim to validate the model with additional measurements. Also, we plan to use the behavior model in conjunction with network traffic models for each of the action categories to develop a network traffic generator based on player behavior.

ACKNOWLEDGMENT

REFERENCES

[1] M. Suznjevic, O. Dobrijevic, and M. Matijasevic, "MMORPG player actions: Network performance, session patterns and latency requirements analysis," *Multimedia Tools Applications*, vol. 45, no. 1-3, pp. 191–241, 2009.

[2] M. Suznjevic, M. Matijasevic, and O. Dobrijevic, "Action specific Massive Multiplayer Online Role Playing Games traffic analysis: Case study of World of Warcraft," in *Proceedings of the 7th ACM SIGCOMM Workshop on Network and System Support for Games*, 2008, pp. 106–107.

[3] M. Suznjevic, O. Dobrijevic, and M. Matijasevic, "Hack, slash, and chat: A study of players' behavior and communication in MMORPGs," in *Proceedings of the 8th Workshop on Network and System Support for Games*, 2009, p. 6.

[4] M. Suznjevic, I. Stupar, and M. Matijasevic, "Traffic modeling of player action categories in a MMORPG," in *Proceedings of the 2nd DIstributed SImulation and Online gaming (DISIO)*, 2011, p. 8.

[5] P.-Y. Tarng, K.-T. Chen, and P. Huang, "An analysis of WoW players' game hours," in *Proceedings of the 7th ACM SIGCOMM Workshop on Network and System Support for Games*, 2008, pp. 47–52.

[6] Y.-T. Lee and K.-T. Chen, "Is server consolidation beneficial to MMORPG? a case study of World of Warcraft," *Cloud Computing, IEEE International Conference on*, vol. 0, pp. 435–442, 2010.

[7] Y.-T. Lee, K.-T. Chen, Y.-M. Cheng, and C.-L. Lei, "World of Warcraft avatar history dataset," in *Proceedings of the second annual ACM conference on Multimedia systems*, 2011, pp. 123–128.

[8] X. Zhuang, A. Bharambe, J. Pang, and S. Seshan, "Player Dynamics in Massively Multiplayer Online Games," School of Computer Science, Carnegie Mellon University, Pittsburgh, Tech. Rep. CMU-CS-07-158, 2007.

[9] D. Pittman and C. GauthierDickey, "A Measurement Study of Virtual Populations in Massively Multiplayer Online Games," in *Proceedings of the 6th ACM SIGCOMM Workshop on Network and System Support for Games*, 2007, pp. 25–30.

[10] D. Pittman and C. G. Dickey, "Characterizing Virtual Populations in Massively Multiplayer Online Role-Playing Games," in *Advances in Multimedia Modeling*, ser. Lecture Notes in Computer Science. Springer Berlin / Heidelberg, 2010, vol. 5916, pp. 87–97.

[11] W.-C. Feng, D. Brandt, and D. Saha, "A long-term study of popular MMORPG," in *Proceedings of the 6th ACM SIGCOMM Workshop on Network and System Support for Games*, 2007, pp. 19–24.

[12] K.-T. Chen and C.-L. Lei, "Network game design: Hints and implications of player interaction," in *Proceedings of the 5th ACM SIGCOMM Workshop on Network and System Support for Games*, 2006, pp. 1–9.

[13] J. Kawale, A. Pal, and J. Srivastava, "Churn prediction in MMORPGs: A social influence based approach," in *Proceedings of the International Conference on Computational Science and Engineering*, 2009, pp. 423–428.

[14] M. Suznjevic. (2011, mar) WoW session activity logger. [Online]. Available: http://www.fer.hr/tel/en/research/research_groups/netmedia/software

[15] D. Williams, N. Yee, and S. E. Caplan, "Who plays, how much, and why? Debunking the stereotypical gamer profile," *Journal of Computer-Mediated communication*, vol. 13, no. 1, pp. 993–1018, November 2008.

[16] K.-T. Chen, P. Huang, and C.-L. Lei, "Game Traffic Analysis: An MMORPG Perspective," *Computer Networks*, vol. 51, no. 3, pp. 19–24, 2007.

[17] M. Kihl, A. Aurelius, and C. Lagerstedt, "Analysis of World of Warcraft Traffic Patterns and User behavior," in *International Congress on Ultra Modern Telecommunications and Control Systems and Workshops*, 2010, pp. 218–223.

[18] P. Svoboda and W. K. M. Rupp, "Traffic analysis and modeling for World of Warcraft," in *Communications, 2007. ICC '07. IEEE International Conference on*, 2007, pp. 1612–1617.

[19] L. J. Gleser and D. S. Moore, " The Effect of Dependence on Chi-Squared and Empiric Distribution Tests of Fit," *The Annals of Statistics*, vol. 11, no. 4, pp. 1100–1108, 1983.