How Players Lose Interest in Playing a Game: An Empirical Study Based on Distributions of Total Playing Times

Christian Bauckhage, Kristian Kersting, Rafet Sifa, Christian Thurau, Anders Drachen, and Alessandro Canossa

***Abstract*—Analyzing telemetry data of player behavior in computer games is a topic of increasing interest for industry and research, alike. When applied to game telemetry data, pattern recognition and statistical analysis provide valuable business intelligence tools for game development. An important problem in this area is to characterize how player engagement in a game evolves over time. Reliable models are of pivotal interest since they allow for assessing the long-term success of game products and can provide estimates of how long players may be expected to keep actively playing a game. In this paper, we introduce methods from random process theory into game data mining in order to draw inferences about player engagement. Given large samples (over 250,000 players) of behavioral telemetry data from five different action-adventure and shooter games, we extract information as to how long individual players have played these games and apply techniques from lifetime analysis to identify common patterns. In all five cases, we find that the Weibull distribution gives a good account of the statistics of total playing times. This implies that an average player's interest in playing one of the games considered evolves according to a non-homogeneous Poisson process. Therefore, given data on the initial playtime behavior of the players of a game, it becomes possible to predict when they stop playing.**

I. Introduction

Developing a profitable game is a challenging endeavor. Over 1,500 commercial titles are published every year and compete for players' time and attention. This rivalry on the globalized market and the high costs of producing quality games are reasons for the gaming industry to attempt to streamline their production. In order to develop games more efficiently, a variety of tools can be adopted and techniques ranging from business practices to user testing have noticeably influenced game development in recent years. They have led to a new focus on the analysis of games and gameplay, whether for industrial or research purposes [1].

One of the new sources of business intelligence is *game telemetry*. Data on revenues, technical performance, development processes, and, most importantly, the behavior of players has gained attention as a new source of insights into a number of processes in the game industry [2], [3]. Accordingly, *game data mining* or *game mining* is now widely recognized as a means for analyzing the demographics of populations of players and for making sense of the ways people play games. Being able to answer questions such as "*How do players behave within a game?*", "*How to entice non-paying players to become paying customers?*", or "*How quickly do players lose interest in playing?*" can help the game industry to prevent fraud, improve game design or user-oriented testing procedures and thus may reduce production and marketing costs.

C. Bauckhage and K. Kersting are with Fraunhofer IAIS and the University of Bonn, Bonn, Germany (emails: {christian.bauckhage, kristian.kersting}@iais.fraunhofer.de); R. Sifa is with the University of Bonn, Bonn, Germany (email: thrafet@gmail.com); C. Thurau is with Game Analytics, Copenhagen, Denmark (email: christian@gameanalytics.com); A. Drachen is with Aalborg University, Ballerup, Denmark (email: drachen@id.aau.dk); A. Canossa is with the IT University of Copenhagen, Copenhagen, Denmark (email: alec@itu.dk).

TABLE I
Statistics of five game data sets analyzed in this paper. Note that the number of players in the different data sets are not in any way indicative of sales figures for the five games!

	#players observed	observation period		
		start date	end date	#months
Just Cause 2	5,331	03-23-2010	10-07-2010	7
Tomb Raider: U.	146,233	12-01-2008	01-31-2009	2
Battlefield B.C. 2	87,126	03-14-2010	12-26-2011	21
Crysis 2	4,364	04-13-2011	11-23-2011	6
Medal of Honor	12,328	11-03-2010	12-18-2011	14

Game mining of telemetry data is particularly appealing, because it provides a viable means for analyzing the behavior of populations of thousands of players and substantially adds to the explanatory power of traditional tools for game user research [1], [3]. While other sources on player behavior, such as customer feedback or surveys, are tedious to evaluate and prone to biased results, recordings of game telemetry data provide direct and consistent information for in-depth analysis. Mining this data allows game developers to uncover frequent usage patterns or behaviors and thus helps them to identify popular game elements or flaws in game mechanics. They have thus greater certainty about design decisions and can accelerate development. Therefore, game telemetry analysis has become acknowledged as a way for the gaming industry to increase their return on investment [2], [4].

A. Related Work

The games research community, too, is increasingly paying attention to in-game pattern recognition and game mining. Early work in this context was notably focused on learning from in-game data in order to create more human-like game bots [5], [6], [7], [8], [9], [10], [11]. And while such efforts have noticeably intensified (see [12] for a recent review), game mining also seems to hold the key to many other problems. Recent work considers modeling of player experience, procedural generation of content, and challenges due to the massive scale of game telemetry data. Examples include recognizing player behaviors and motivations [13], [14], [15], tracking the evolution of social groups in MMORPGs [16], identifying patterns in the naming of player characters and guilds [17], detecting causes of player frustration [18], [19], or uncovering social ties among players of FPS games [20].

In this paper, we address yet another problem that –to our knowledge– has not been studied before. Given data as to the behavior of a population of players, we want to understand how engagement in a game evolves over time. Prior work most closely related to our question can be found in [13], [14], [15], [16]. However, in contrast to these contributions, our aim is not to categorize players or teams but to proceed towards a general understanding of how the allure of a game changes with time.

B. Contribution and Main Result

Our focus in this paper is on applying game mining in order to learn about the temporal evolution of people's interest in playing a game. In particular, we analyze large samples of behavioral telemetry data from five different action-adventure or shooter games (see Table I) and mine them for common patterns or deviating characteristics. Our goal is to infer process models that characterize time series of player engagement. Once available, such models might be applied to predict a game's long-term success from only a few measurements of player activity. The implications for game development are obvious, since developers will be able to estimate how long people will keep playing the game in question. Models could be applied to data from alpha- and beta-testing with larger numbers of players, as well as to data recorded immediately after a game's release. The latter is particularly interesting for persistent games such as free-to-play online games which are currently one of the most rapidly growing genres in the industry.

There are two fundamental difficulties that need to be addressed in modeling the dynamics of player engagement. First of all, *interest in playing a game* is an abstract quantity. Psychologists or Sociologists would call it a *latent* or *hidden variable* which influences measurable quantities such as the frequency of playing sessions or total playing times but cannot itself be observed directly. We are therefore in need of a principled way of inferring the dynamics of player interest from observable data.

Second of all, we are in need of models that are *physical* in that they represent processes that can take place in the real world. In other words, while it is always possible to produce an abstract mathematical model that fits well a sample of data, not every such model may correspond to physical reality. In particular, time series can easily be modeled using computational intelligence techniques such as Neural Networks or Hidden Markov Models, but the resulting representations will hardly explain the mechanisms that caused the data to appear as observed. Even if data analysts are often tempted to assign meaning to abstract representations, such reification is seldom justified; interpretative claims cannot arise from mathematics but must abide by the general conditions of the application domain. Results from game telemetry analysis need to provide tangible explanations that address actual needs and requirements of game developers in order to be useful for the industry.

Addressing both these difficulties, our contribution in this paper is to introduce methods from random process theory [21], [22] and lifetime analysis [23] to game mining. We assume that a player's interest in playing a game evolves as a random process with drift. This appears reasonable, since, on a day-to-day basis, a player's eagerness to play will depend on numerous unforeseeable events (including private circumstances, releases of new game content or competing games, etc.), but, in the long run, it will most certainly decline, because at some point most players will stop playing the game. Given this assumption, we apply lifetime analysis to infer details about the unobservable process.

Although interest in gameplay cannot itself be measured from game telemetry data, we can estimate how long it takes for it to drop to zero. To this end, we measure the duration of individual total playing times observed among a population of players. Having thus obtained an empirical distribution of total playing times, we fit different lifetime- or first passage time distributions to the data. In particular, we consider four well established distributions, namely the Weibull-, Gamma-, Log-normal- and Inverse Gaussian distribution, because they are known to be in one-to-one correspondence to four distinct random processes which frequently occur in nature and are very well understood (cf. e.g. [21], [22], [24], [25]).

Our extensive empirical tests reveal that, overall, the Weibull distribution gives a very good account of the distribution of total playing times. Observing first passage times (to a level of zero interest) to be Weibull distributed implies that the underlying latent process is a particular non-homogeneous Poisson process, namely a power law process. For the data sets at hand, we can thus infer that an average player's interest in playing action games drops according to a power law.

This is indeed an interesting finding, because power law statistics have been observed in a wide variety of contexts and are known to describe different aspects of human behavior. Examples include the dynamics of economic decisions [26] or politics and conflicts [27], [28], email communication patterns [29], [30], but also the frequency of choices of character names in online games [17]. Our result is therefore well in line with with a vast body of literature from fields such as sociometrics or mathematical psychology.

C. Organization of the Remainder of the Paper

Next, we briefly introduce the particularities and characteristics of the five game telemetry data sets we considered in our analysis. Then, in Section III, we present details on random process models of interest in gameplay and how they are related to first passage time distributions. In Section IV, we report and discuss the results of our analysis and Section V once again summarizes our goals and findings. For details as to our methodology for maximum likelihood fitting of lifetime distributions, we refer to the appendix.

II. Data Sets and Characteristics

For the study presented in this paper, we analyzed in-game data from five recent action-adventure or shooter games. These include two single-player games (*Just Cause 2*, *Tomb Raider: Underworld*) as well as three multi-player games (*Battlefield Bad Company 2*, *Medal of Honor*, *Crysis 2*).

Data from Tomb Raider: Underworld (TRU) were drawn from the Square Enix metrics suite, which contains data from over 1.5 million players. The sample drawn covers all data collected in a two month period (from Dec 1st, 2008 to Jan 31st, 2009), providing records from approximately 203,000 players (around 100 GB). The game was launched in November 2008, so that the data represent a time period where the game was recently released to the public. The data in the sample were extracted in a series of tables, cleaned and transposed to a single table. Records with missing information and other irregularities were removed (e.g. only the first playthrough for players who played the game more than once was retained) which provides us with a final sample size of 146,233 player observations.

Data from Just Cause 2 (JC2) were drawn from the Square Enix metrics suite, which contain data from all players of the game. A sample of over 5,000 players was drawn using simple random sampling, and processed similar to the TRU data, resulting in a final dataset containing 5,331 records.

The data for the three multi-player games Battelefield Bad Company 2 (BF2), Crysis 2 (CR2) and Medal of Honor (MOH) were collected from the p-stats network[1]. We note that this website only collects information from players who have installed a specific add-on for their games. While extremely casual players may not have installed such an add-on, many regular players use it to meticulously keep track of their in-game activities and achievements. These data, too, were preprocessed as above to mitigate effects due to incomplete information.

After preprocessing, each data set contains anonymized observations of several thousand individual players and their in-game activities over different periods of time (see Tab. I). In particular, the data describes when, i.e. on which dates, and for how long, i.e. for how many seconds, each player was actively playing the corresponding game.

Accordingly, the data allows for various comparisons of the games with respect to temporal preferences of their players. As an example, Fig. 1 shows histograms of player activity per weekday. For the two single-player games, there seems to be a tendency towards higher activity on the weekend. While for *Tomb Raider: Underworld* this trend is noticeable but minuscule, it is well pronounced for *Just Cause 2*. Other than this, the average daily playing activities in each of the five games considered here appear to be evenly distributed over the days of the week. As we shall see in later sections, other temporal statistics of player behavior, too, are in close agreement for the five different games.

[1]http://p-stats.com

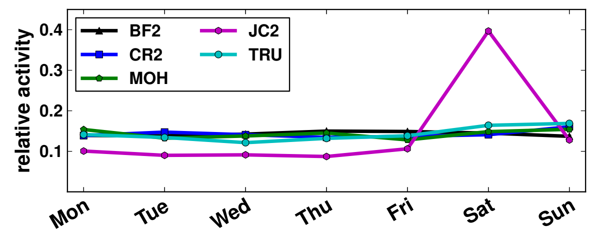


Fig. 1. Normalized histograms of player activity per weekday (averaged over population) for Battle Field Bad Company 2 (BF2), Crysis 2 (CR2), Medal of Honor (MOH), Just Cause 2 (JC2), and Tomb Raider: Underworld (TRU). Except for Just Cause 2, for which player activity noticeably spikes on Saturday, average daily playing times for the different games appear fairly evenly distributed over the week.

III. Modeling Interest in Game Playing

In this section, we motivate our approach to modeling the dynamics of player engagement and provide a brief introduction to random process theory and first passage time distributions.

A. Random Processes Models for Player Interest

Our main goal with this paper is to attempt a formalization of the temporal evolution of people's interest in playing a game. This is not a straightforward endeavor, because *interest in playing a game* is a rather abstract quantity. It cannot be measured directly but is a *latent variable* that has to be inferred from observable data. Moreover, it appears reasonable that a player's interest in playing a game does not evolve in a deterministic fashion that could be described by means of simple equations. Rather, everyday experience tells us that, on any given day, a player's eagerness to play a game will be influenced by different unforeseeable events ranging from private circumstances or schedules to the release of new game content or competing games.

Given both these considerations, it makes sense to model the evolution of a player's interest in playing a game as a stochastic time series or random process. That is, at any time t, a player's interest I_t in playing a game is a random quantity that may or may not depend on previous values and whose future values cannot be predicted exactly.

Random process models are frequently applied to characterize the dynamics of systems or events observed in nature. Familiar examples include the frequency of incoming telephone calls, fluctuations of financial markets, biomedical measurements such as a patient's blood pressure, or the motion of particles. There are various types of random process models which differ with respect to their mathematical complexity and application domain (cf. e.g. [21], [22], [25]). Since our interest is in understanding the temporal dynamics of a one-dimensional quantity, we restrict our discussion and experimental verification to the four most common models for this purpose. Their mathematical form is shown in the right column of Tab. II; their names and properties are as follows:

TABLE II
FIRST PASSAGE TIME DISTRIBUTIONS AND UNDERLYING PROCESSES

first passage time distribution	underlying latent process
Gamma distribution $p(t) = \frac{1}{\alpha^\kappa \Gamma(\kappa)} t^{\kappa-1}\, e^{-\frac{t}{\alpha}}$	Poisson process $p(I_t = \kappa) = \frac{(\lambda t)^\kappa}{\Gamma(\kappa+1)}\, e^{-\lambda t}$
Weibull distribution $p(t) = \frac{\kappa}{\alpha}\left(\frac{t}{\alpha}\right)^{\kappa-1}\, e^{-\left(\frac{t}{\alpha}\right)^\kappa}$	non-homogeneous Poisson process $p(I_t = \kappa) = \frac{\lambda(t)^\kappa}{\Gamma(\kappa+1)}\, e^{-\lambda(t)}$
Inverse Gaussian distribution $p(t) = \sqrt{\frac{\alpha}{2\pi t^3}}\, e^{-\frac{\alpha(t-\mu)^2}{2\mu^2 t}}$	Wiener process with drift $dI_t = \nu t + \beta dW_t$
Log-normal distribution $p(t) = \frac{1}{t\sigma\sqrt{2\pi}}\, e^{-\frac{(\log t - \mu)^2}{2\sigma^2}}$	Fokker-Planck diffusion $\frac{\partial}{\partial t} f = \frac{\partial^2}{\partial I^2} I^2 f - \frac{\partial}{\partial I} I f$

Posisson processes: If a player's interest in playing a game would follow a Poisson process, we could picture this model by imagining that, at certain random points in time, the player feels an urge to play. On average, this urge to play would set in according to a constant rate λ. This parameter does not change with time and fully characterizes the process.

Non-homogeneous Poisson processes: this model class generalizes simple Poisson processes in that the rate parameter $\lambda(t)$ is now assumed to be a function of time.

Bwownian motion / Wiener processes: while the two previous models do not assume that, at time t, the quantity of interest I_t depends on its history, in a Wiener process, the current value of I_t would result from random Gaussian perturbations dW_t of its previous value. In a Wiener process with drift ν, individual updates are still random, yet, over time, the quantity of interest will tend towards decreasing values.

Fokker-Planck diffusion: this model provides the most general description of the temporal evolution of a one-dimensional stochastic variable. It assumes that there is a probability density $f(I, t)$ that characterizes the random variable I_t. A stochastic differential equation then describes trends and speeds with which the density of the quantity of interest progresses.

As these four standard models cover the whole range from fairly simple to fairly complex stochastic processes, we can reasonably expect that one or several of them may explain the dynamics of player interests in a statistically meaningful way. However, since a player's interests in playing a game is not immediately apparent in our game telemetry data, we are in need of a method that would bridge the gap between our data and the proposed models. In the next subsection, we discuss how lifetime analysis accomplishes this feat.

B. First passage Time Distributions

First passage time distributions are used to characterize how long it takes for a random process to reach a certain value. They form the backbone of lifetime analysis which is concerned with investigating, say, how long complex (mechanical or biological) systems that consist of several components or factors work properly on average [23].

In our context of estimating the dynamics of a population's interest in a game, we might ask for how long individual players have played the game in total. From the point of view of random processes, analyzing distributions of total playing times is tantamount to estimating how long it takes until an average player's interest in playing that game drops to a value of zero.

What is interesting is that specific random process models are in a one-to-one relation with certain first passage time distributions. Table II lists the correspondences for the four random processes considered here.

For Poisson processes, first passage times are distributed according to a Gamma distribution [24]. That is, if an empirical distribution of total individual playing times can be modeled by a Gamma distribution with scale parameter α and shape parameter κ, we can infer that an average player's interest in playing consisted of κ independent urges to play which occurred at a rate of $\lambda = \frac{1}{\alpha}$.

If an empirical lifetime distribution can be modeled as a Weibull distribution, we can infer that the underlying random process is a non-homogeneous Poisson process [24]. In particular, Weibull distributed lifetimes indicate the presence of a power law process where $\lambda(t) = \frac{\kappa}{\alpha}(\frac{t}{\alpha})^{\kappa-1}$.

For Wiener processes, first passage times follow Inverse Gaussian distributions [24] and for processes characterized through Fokker-Planck diffusion, first passage times follow Log-normal distributions [31].

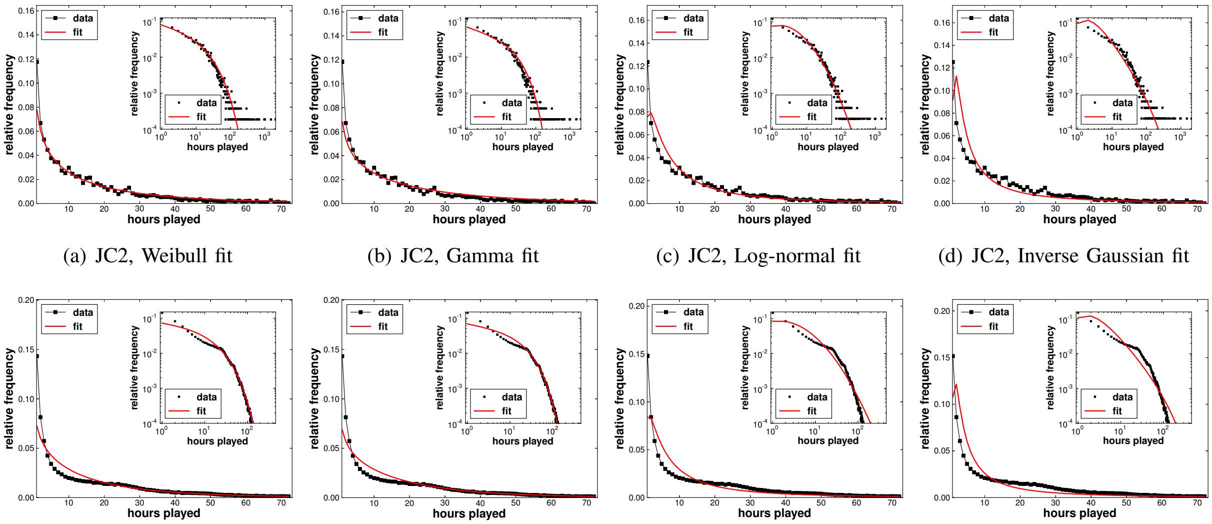
Obviously, the fortunate existence of one-to-one relations between random processes and first passage time distribution can help us to infer the dynamics of player engagement. Even though a player's interest in playing a game is a latent variable that cannot be read from game telemetry data, an analysis based on first passage time distributions allows us to make claims as to the dynamics of player engagement. Next, we present results obtained from fitting lifetime distributions to observed total playing times and discuss implications of our findings.

IV. RESULTS AND DISCUSSION

Given the game telemetry data introduced in section II, we extract the total individual playing times of each observed player for each of the five games by accumulating the durations of all of their observed playing sessions. This provides us with total playing times measured in seconds which we round to hours in order to accomplish statistical smoothing for robustness against over-fitting as well as to facilitate graphical representation.

Figures 2 and 3 show the resulting empirical distributions of individual playing times (plotted in black) for the two single-player and the three multi-player games, respectively. Note that, for better visibility, the time axes in the body of each plot are truncated at 72 hours; the insets in each plot display the data in a doubly logarithmic representation and cover the whole range of total playing times observed for each game.

Looking at the figures, we immediately note that the empirical distributions of total playing times per player are



(a) JC2, Weibull fit (b) JC2, Gamma fit (c) JC2, Log-normal fit (d) JC2, Inverse Gaussian fit

(e) TRU, Weibull fit (f) TRU, Gamma fit (g) TRU, Log-normal fit (h) TRU, Inverse Gaussian fit

Fig. 2. Observed empirical distributions of playing times per player (in hours) and fitted first passage time distributions for the two single-player games Just Cause 2 (JC2) and Tomb Raider: Underworld (TRU). Shape and scale parameters of the fits were obtained from from maximum likelihood estimation. For better visibility, the times axes in the figures are truncated at 72 hours; the insets show data and fitted functions on a logarithmic scale and cover the whole range of observed playing times. For both games, the observed empirical distribution are heavily skewed to the right: while most of the observed players played for only a few hours, the long tails of the distributions indicate that a noticeable fraction players played for more than 100 hours in total.

similar across the five games. For each game considered here, we observe playing time distributions that are highly skewed to the right. That is to say that, for each game, there is a surprisingly high percentage of players who have played the game for only a few hours. At the same time, since the tails of the observed distributions are long, we find that a noticeable fraction of players have played the game for a considerably longer time. It is important to note that this general behavior is independent of the duration of the corresponding observation period. Regardless of whether our data covers a period of only two months or almost two years, the distributions of total playing times assume their maxima at the ordinate, decrease rather smoothly, and are long-tailed.

Figures 2 and 3 also show the results (plotted in red) of maximum likelihood fits of the four previously discussed first passage time distributions to the empirical data.

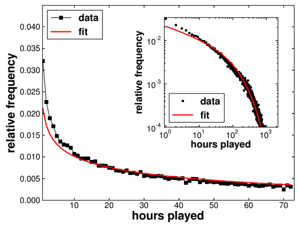
Already by means of visual inspection it appears that the Inverse Gaussian distribution generally gives a rather poor account of the empirically observed distributions of playing times. Based on what we discussed above, this immediately rules out the hypothesis that an average player's interest in playing a game develops according to a Wiener process.

For the other three first passage time distributions, the general picture is less clear from visual inspection only. The Weibull and the Gamma distribution apparently provide rather good approximations to our data, in particular, they seem to model the tails of the empirical distributions well. But for some of the games, the Log-normal distribution, too, appears to provide a good fit. In order to quantify this qualitative observation, we resort to goodness of fit tests.

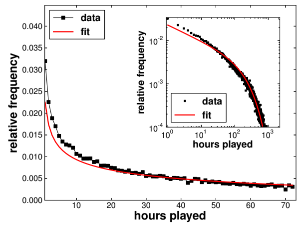
Table III summarizes results we obtain from running Kolmogorov-Smirnov tests. It presents the corresponding p-values where higher p-values indicate better fits. Looking at the table, three observations stick out: first of all, for the game of Tomb Raider: Underworld, none of the tested lifetime distributions achieves a p-value that would indicate a good (if any) fit to the data. However, this is a consequence of the large amount of data for this game (more than 100,000 observed players) rather than a failure on the part of the models. Note that statistical tests such as the Komogorov-Smirnov test were devised at a time when data samples were small; it is common lore among statisticians that for very large sample sizes such tests will indicate departure from the tested distribution even if actual deviations are minor. The fact that traditional statistical tests are indeed tailored towards smaller samples is corroborated by the observation that, for the two games for which we have less than 10,000 data, we obtain the highest the p-values. This ineptitude of statistical testing leaves us with the dilemma of either having to down-sample the data (thus introducing other difficulties) or of having to rely on qualitative impressions. Adopting the latter strategy to determine which lifetime distribution provides the better fit to the data obtained for

TABLE III
RESULTS OF GOODNESS OF FIT TESTS

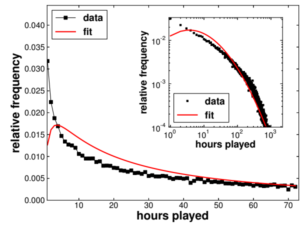
	p-value according to Kolmogorov-Smirnov test			
	Weibull	Gamma	Log-normal	Inv. Gaussian
Just Cause 2	**0.055**	0.018	0.018	0.002
Tomb Raider: U.	0.000	0.000	0.000	0.000
Battlefield B.C. 2	**0.008**	0.003	0.002	0.000
Crysis 2	**0.051**	0.022	0.037	0.003
Medal of Honor	0.014	0.005	**0.030**	0.001



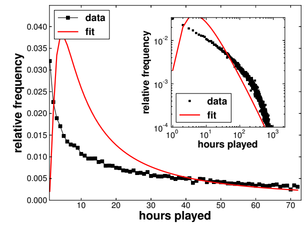
(a) BF2, Weibull fit



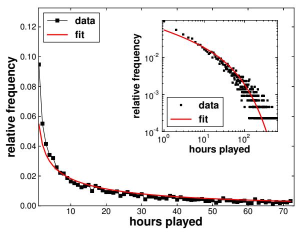
(b) BF2,Gamma fit



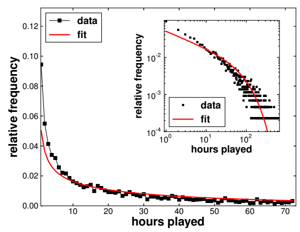
(c) BF2,Log-normal fit



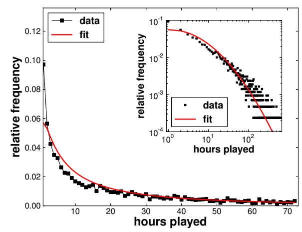
(d) BF2,Inverse Gaussian fit



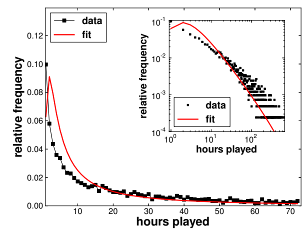
(e) CR2, Weibull fit



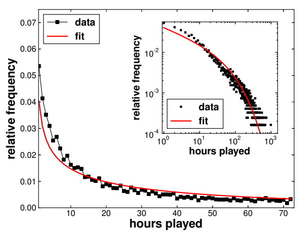
(f) CR2,Gamma fit



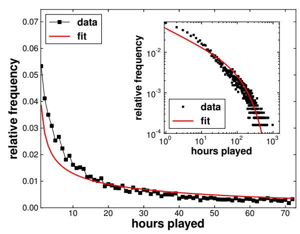
(g) CR2,Log-normal fit



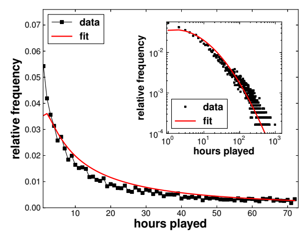
(h) CR2,Inverse Gaussian fit



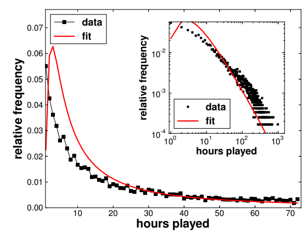
(i) MOH, Weibull fit



(j) MOH, Gamma fit



(k) MOH, Log-normal fit



(l) MOH, Inverse Gaussian fit

Fig. 3. Observed empirical distributions of playing times per player (in hours) and fitted first passage time distributions for the three multi-player games Battlefield Bad Company 2 (BF2), Crysis 2 (CR2), and Medal of Honor (MOH). The parameters of the fitted statistical distributions were obtained from from maximum likelihood estimation. For better visibility, the times axes in the figures are truncated at 72 hours; the insets show data and fitted functions on a logarithmic scale and cover the whole range of observed playing times. Just as with the single-player, for the multi-player games, too, the observed empirical distributions are skewed to the right; while many players played only a few hours, a considerable fraction played much longer than 100 hours.

TABLE IV
DETAILS ON WEIBULL FITS TO THE EMPIRICAL DATA

	sample mean	Weibull mean	Weibull parameters α	Weibull parameters κ
Just Cause 2	21.37	20.33	17.26	0.76
Tomb Raider: U.	18.00	17.93	16.42	0.85
Battlefield B.C. 2	143.69	143.67	118.93	0.74
Crysis 2	51.35	50.02	37.48	0.66
Medal of Honor	84.09	82.54	61.79	0.66

Tomb Raider: Underworld suggests that the Weibull and the Gamma distribution outperform the Log-normal model.

Second of all, for three out of the remaining four games, Weibull fits to the empirical playing time distributions yield the highest p-values. Third of all, for the two games for which there are less then 10,000 player observations, the Weibull distribution actually yields statistically significant fits ($p > 0.05$). Since one of these games is a single-player game (Just Cause 2) and the other is a multi-player game (Crysis 2), it appears that the Weibull distribution as a suitable model for the lifetimes of interest in a game regardless of whether or not that game is played online.

Finally, we note that while for Medal of Honor the Log-normal distribution seems to provide the best fit to our data, its behavior at the ordinate diverges from the empirical distribution (see Fig. 3). On the other hand, the Weibull distribution gives a reasonable account, too, and also provides a better explanation of the lower tail.

Summarizing our findings for the five games under consideration, it appears that suitably parameterized Weibull distributions provide convincing abstractions of empirically determined distributions of total playing times per player. The figures in Tab. IV further stress this conclusion. For each of the five games, the table compares *average* total playing time per player either computed from the data (sample mean) or predicted by the best fitting Weibull distribution (Weibull mean). In each case we find the difference between sample- and predicted mean to be two orders of magnitude less than the value of the sample mean which further confirms the adequacy of the Weibull model.

Table IV also lists the shape and scale parameters κ and α of the best fitting Weibull models. The corresponding Weibull distributions are plotted logarithmically in Fig. 4. It is noticeable that the curves for the two single-player games and the curves for the three multi-player games cluster together; also, total playing times for multi-player games

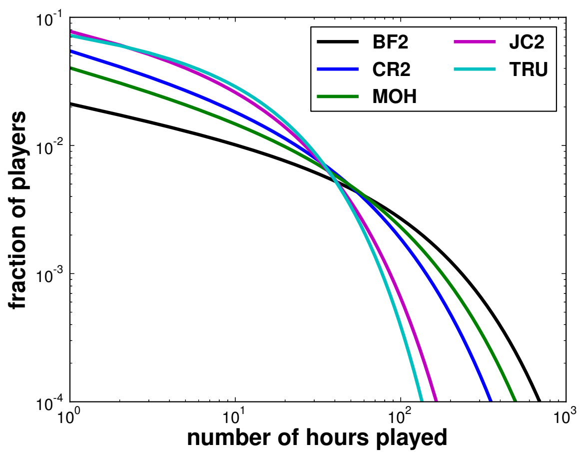


Fig. 4. Comparison of the Weibull fits to the playing times per player (averaged over population) for Battle Field Bad Company 2 (BF2), Crysis 2 (CR2), Medal of Honor (MOH), Just Cause 2 (JC2), and Tomb Raider: Underworld (TRU). Two effects are clearly visible: for the single-player games (JC2 and TRU), the largest amount of time spent playing (between 100 and 200 hours) is less than for the multi-player games (between 300 and 700 hours). Correspondingly, for the single-player games, the fraction of players investing less than 30 hours is larger than for the multi-player games.

exceed those of single-player games. Both observations make perfect sense according to everyday experience and therefore once more underline the appropriateness of the Weibull distribution.

In light of what we discussed in Section III, it is therefore justified to assume that an average player's interest in playing either of the games considered here declines according to a non-homogeneous Poisson process. Since the time-dependent rate of this process is given by $\lambda(t) = \frac{\kappa}{\alpha}(\frac{t}{\alpha})^{\kappa-1}$, where κ and α are the shape and scale parameters of the Weibull distribution that fits playing times, we recognize a power law process. Looking again at Tab. IV, we find that the exponent $\kappa - 1$ is negative for each of the five games. This establishes that the rate, at which an average player's urges to play set in, decreases over time.

V. Conclusion

Game data mining, or game mining, is a topic of rapidly increasing interest [1], [2], [3], [4]. When applied to game telemetry data, techniques from data analysis and statistical pattern recognition are valuable tools for customer analytics. In particular the analysis of player behavior can help improve game design and playability.

An important open problem in this context is to understand and model how a players interest in playing a game evolves over time. Once available, such models will enable game developers to assess the longterm success of their products early on and thus to correct design problems and to adapt their products or marketing strategies correspondingly.

Given large game telemetry data sets from five games covering two different genres (action-adventure and shooter games), we extracted information as to how long individual players have played these games and applied techniques from lifetime analysis in order to identify distinctive patterns. In all five cases, we found that the Weibull distribution gives a good account of the empirical distribution of total playing times. This, in turn, implies that an average player's interest in playing these games evolves according to a non-homogeneous Poisson process whose intensity function is given by a power law.

Given suitable game telemetry data from, say, alpha- or beta-testing, the techniques presented in this paper make it possible to predict how long a player will stay engaged even before a game is released. Moreover, generalizations are possible that would allow for monitoring and understanding the effects of updates to a game. However, while we have performed statistical analysis on a solid empirical basis of five different games, we cannot guarantee that the Weibull distribution will also be found to fit playtime data from other types of games such as, for instance, MMOGs. Corresponding investigations are planned for future work.

Appendix

To obtain the results reported throughout this paper, we applied maximum likelihood estimation (MLE) to determine the parameters of the best fitting lifetime distributions. For the Inverse Gaussian- and the Log-normal distribution, maximum likelihood estimators of their parameters are readily available in closed form; for the Gamma- and the Weibull distribution, such closed form solutions do not exist. In order for this paper to be self-contained, this appendix therefore provides details as to our methodology.

A. *MLE for the Inverse Gaussian distribution*

Given a sample $D = \{t_i\}_{i=1}^N$ of observed first passage times, MLEs of the location and scale parameter of the Inverse Gaussian distribution are given by

$$\mu = \frac{1}{N}\sum_i t_i = \bar{t} \quad \text{and} \quad \alpha = \frac{N}{\sum_i \big(\frac{1}{t_i} - \frac{1}{\mu}\big)},$$

respectively.

B. *MLE for the Log-normal distribution*

Given a sample of N observed first passage times, MLEs of the location and scale parameter of the Log-normal distribution are given by

$$\mu = \frac{1}{N}\sum_i \log t_i = \overline{\log t} \quad \text{and} \quad \sigma = \frac{1}{N}\sum_i \big(\log t_i - \mu\big)^2,$$

respectively.

C. *MLE for the Gamma distribution*

Given a data sample $D = \{t_i\}_{i=1}^N$, the log-likelihood for the parameters of the Gamma distribution is

$$L(\alpha, \kappa \mid D) = N\big((\kappa - 1)\overline{\log t} - \log\Gamma(\kappa) - \kappa\log\alpha - \bar{t}/\alpha\big)$$

where overlines indicate averaging. The MLE for α is readily determined as $\alpha = \bar{t}/\kappa$. However, differentiating the log-likelihood with respect to κ and equating to zero does not permit for a closed form solution. We therefore follow Minka's proposal [32] and apply Newton's method using

$$\frac{1}{\kappa^{\text{new}}} = \frac{1}{\kappa} + \frac{\overline{\log t} - \log \bar{t} + \log \kappa - \Psi(\kappa)}{\kappa^2 \big(1/\kappa - \Psi'(\kappa)\big)}$$

where Ψ is the digamma function.

D. MLE for the Weibull distribution

Given a data sample $D = \{t_i\}_{i=1}^{N}$, the log-likelihood for the parameters of the Weibull distribution is

$$L(\alpha, \kappa \mid D) = N\big(\log \kappa - \kappa \log \alpha\big) - (\kappa - 1)\overline{\log t} - \sum_i (t_i/\alpha)^\kappa.$$

Deriving L with respect to α and κ leads to a coupled system of partial differential equations for which there is no closed form solution. Consequently, we apply Newton's method for simultaneous equations and compute

$$\begin{bmatrix} \kappa^{\text{new}} \\ \alpha^{\text{new}} \end{bmatrix} = \begin{bmatrix} \kappa \\ \alpha \end{bmatrix} + \begin{bmatrix} \frac{\partial^2 L}{\partial \kappa^2} & \frac{\partial^2 L}{\partial \kappa \partial \alpha} \\ \frac{\partial^2 L}{\partial \kappa \partial \alpha} & \frac{\partial^2 L}{\partial \alpha^2} \end{bmatrix}^{-1} \begin{bmatrix} -\frac{\partial L}{\partial \kappa} \\ -\frac{\partial L}{\partial \alpha} \end{bmatrix}$$

where the entries of the Hessian matrix and the gradient vector are given by

$$\begin{aligned}
\frac{\partial L}{\partial \kappa} &= N/\kappa - N \log \alpha + \overline{\log t} - \sum_i (t_i/\alpha)^\kappa \log(t_i/\alpha) \\
\frac{\partial L}{\partial \alpha} &= \kappa/\alpha \Big(\sum_i (t_i/\alpha)^\kappa - N \Big) \\
\frac{\partial^2 L}{\partial \kappa^2} &= N/\kappa^2 - \sum_i (t_i/\alpha)^\kappa \big(\log(t_i/\alpha)\big)^2 \\
\frac{\partial^2 L}{\partial \alpha^2} &= \kappa/\alpha^2 \Big(N - (\kappa + 1) \sum_i (t_i/\alpha)^\kappa \Big) \\
\frac{\partial^2 L}{\partial \kappa \partial \alpha} &= 1/\alpha \sum_i (t_i/\alpha)^\kappa + \kappa/\alpha \sum_i (t_i/\alpha)^\kappa \log(t_i/\alpha) - N/\alpha.
\end{aligned}$$

Acknowledgments

This work would not have been possible without the continued collaboration and help from our colleagues at Square Enix, IO Interactive and Crystal Dynamics. The authors would like to direct special thanks to Jim Blackhurst and Tim Ward from Square Enix/Eidos team in London for their support, insights, and advice. The authors also thank Janus Rau Sørensen and the IOI User-Research Team, Anders Nielsen, company secretary at IO Interactive, Chris Glover, Phil Elliot, and the management of all three companies.

References

[1] K. Isbister and N. Schaffer, *Game Usability: Advancing the Player Experience*. Morgan Kaufman, 2008.
[2] L. Mellon, "Applying Metrics Driven Development to MMO Costs and Risks," Versant Corporation, Tech. Rep., 2009.
[3] A. Drachen and A. Canossa, "Evaluating Motion: Spatial user Behaviour in Virtual Environments," *Int. J. of Arts and Technology*, vol. 4, no. 3, pp. 294–314, 2011.
[4] J. Kim, D. Gunn, E. Schuh, B. Phillips, R. Pagulayan, and D. Wixon, "Tracking Real-Time User Experience (TRUE): A Comprehensive Instrumentation Solution for Complex Systems," in *Proc. ACM CHI*, 2008.
[5] E. Sklar, A. Blair, P. Funes, and J. Pollack, "Training Intelligent Agents Using Human Internet Data," in *Proc. Asia-Pacific Conf. on Intelligent Agent Technology*, 1999.
[6] P. Spronck, I. Sprinkhuizen-Kuyper, and E. Postma, "Improving Opponent Intelligence through Machine Learning," in *Proc. GAME-ON*, 2002.
[7] C. Bauckhage, C. Thurau, and G. Sagerer, "Learning Human-like Opponent Behavior for Interactive Computer Games," in *Pattern Recognition*, ser. LNCS, vol. 2781. Springer, 2003.
[8] S. Zanetti and A. El Rhalibi, "Machine Learning Techniques for FPS in Q3," in *Proc. ACM ACE*, 2004.
[9] R. Le Hy, A. Arrigioni, P. Bessière, and O. Lebeltel, "Teaching Bayesian Behaviours to Video Game Characters," *Robotics and Autonomous Systems*, vol. 47, no. 2–3, pp. 177–185, 2004.
[10] C. Bauckhage and C. Thurau, "Towards a Fair 'n Square Aimbot – Using Mixtures of Experts to Learn Context Aware Weapon Handling," in *Proc. GAME-ON*, 2004.
[11] B. Gorman, C. Thurau, C. Bauckhage, and M. Humphrys, "Bayesian Imitation of Human Behavior in Interactive Computer Games," in *Proc. IEEE ICPR*, 2006.
[12] G. Yannakakis, "Game AI Revisited," in *Proc. Conf. Computing Frontiers*, 2012.
[13] N. Yee, "Motivations for Play in Online Games," *Cyber Psychology & Behavior*, vol. 9, no. 6, pp. 772–775, 2006.
[14] A. Drachen, A. Canossa, and G. Yannakakis, "Player Modeling using Self-Organization in Tomb Raider: Underworld," in *Proc. IEEE CIG*, 2009.
[15] B. Weber and M. Mateas, "A Data Mining Approach to Strategy Prediction," in *Proc. IEEE CIG*, 2009.
[16] C. Thurau and C. Bauckhage, "Analyzing the Evolution of Social Groups in World of Warcraft," in *Proc. IEEE CIG*, 2010.
[17] C. Thurau and A. Drachen, "Naming Virtual Identities: Patterns and Inspirations for Character Names in World of Warcraft." in *Proc. FDG*, 2011.
[18] E. Andersen, Y.-E. Liu, R. Snider, R. Szeto, S. Cooper, and Z. Popović, "On the Harmfulness of Secondary Game Objectives," in *Proc. FDG*, 2011.
[19] A. Canossa, A. Drachen, and J. Rau Møller Sørensen, "Arrrgghh !!! - Blending Quantitative and Qualitative Methods to Detect Player Frustration," in *Proc. FDG*, 2011.
[20] Y. Xu, X. Cao, A. Sellen, R. Herbrich, and T. Graepel, "Sociable Killers: Understanding Social Relationships in an Online First-Person Shooter Game," in *Proc. ACM CSCW*, 2011.
[21] N. Krylov, *Intoroduction to the Theory of Random Processes*. AMS, 2002.
[22] A. Papoulis and S. Pillai, *Probability, Random Variables and Stochastic Processes*. McGraw-Hill, 2002.
[23] J. Lawless, *Statistical Models and Methods for Lifetime Data*. Wiley, 2003.
[24] H. Rinne, *The Weibull Distribution*. Chapman & Hall / CRC, 2008.
[25] B. Øksendal, *Stochastic Differential Equations*. Springer, 2010.
[26] B. Mandelbrot and R. Hudson, *The (Mis)Behavior of Markets*. Profile Books, 2005.
[27] H. Jensen, *Self-Organized Criticality: Emergent Complex Behavior in Physical and Biological Systems*. Cambridge University Press, 1998.
[28] L.-E. Cederman, "Endogenizing geopolitical boundaries with agent-based modeling," *Proc. Nat. Acad. Sci.*, vol. 99, no. 3, pp. 7296–7303, 2002.
[29] J. Iribarren and E. Moro, "Impact of Human Activity Patterns on the Dynamics of Information Diffusion," *Phys. Review Letters*, vol. 103, no. 3, p. 038702, 2009.
[30] S. Uddin, S. Murshed, and L. Hossain, "Power-law Behavior in Complex Organizational Communication Networks During Crisis," *Physica A*, vol. 390, no. 15, pp. 2845–2853, 2011.
[31] R. Capocelli and L. Riccardi, "On the Inverse of the First Passage Time Probability Problem," *J. Applied Probability*, vol. 9, no. 2, pp. 270–287, 1972.
[32] T. Minka, "Estimating a Gamma Distribution," Microsoft Research, Cambridge, UK, Tech. Rep., 2002.