

A Key Risk Indicator for the Game Usage Lifecycle

Luiz Bernardo Martins Kummer, Júlio Cesar Nievola, Emerson Cabrera Paraiso

Graduate Program in Informatics, Pontificia Universidade Católica do Paraná, Rua Imaculada Conceicao, 1155, Curitiba-PR, Brazil
{luiz.kummer, nievola, paraiso}@ppgia.pucpr.br

Abstract

Among the different types of digital games, Massively Multiplayer Online Role-Playing Games (MMORPGs) are one of the most popular. Game producers use usage data to compute metrics to analyze their game lifecycles. The most popular is the MAU (Monthly Active Users), which indicates the number of active players in each timestamp. MAU only describes how many players played a game. It cannot show how players are motivated for that game. We support that motivation is a key factor to follow and game producers should be aware of it. In this way, we are proposing in this article a new game independent Key Risk Indicator. The risk indicator is based on commitment and, in our context, can be defined as the attachment of a player to a given game. The game industry may use commitment to evaluate how attractive a game is during its usage lifecycle. A machine learning approach is presented to predict the risk indicator. We applied the method to a real dataset. The proposed approach identified risky situations where the classical approach did not.

Introduction

The digital game industry has many needs that Artificial Intelligence (AI) can supply, being a hot topic in research nowadays (Silver et al. 2016). As games evolve, becoming increasingly complex, new challenges arise, leaving space for AI researchers, especially the ones interested in Machine Learning (Galway et al. 2008). Among the different types of digital games, Massively Multiplayer Online Role-Playing Games (MMORPGs) are one of the most popular. There are several elements of MMORPGs suitable for observation. As players play a game, usage data are generated opening an opportunity for machine learning. Game producers use usage data to calculate metrics to analyze their game lifecycles. The most popular is the MAU (Monthly Active Users), which indicates the number of active players in each timestamp (Speller 2012). Usage data are any data which contain information about usage (e.g., login records or a list of online players). Game pro-

ducers use metrics like MAU to support decisions, such as: to identify the best moment to release a new version, to start an advertisement campaign, to stop the operations when a bad situation is identified or to evaluate the new release acceptance (Zhu et al. 2010 ; Speller 2012 ; Sheffield and Alexander 2008). The problem is that MAU only describes how many players played a game. It cannot show how players are motivated for that game (a subjective analysis). Finally, it is important to highlight that game producers aim in extending the game lifecycle to the maximum, maximizing profit. According to Speller (2012), motivational factors (i.e. motivation that keeps a player active) are empirically interpreted. We support that motivation is a key factor to follow and game producers should be aware of it. In this way, we are proposing in this article a new Key Risk Indicator (KRI). In order to deal with the subjective aspect, we applied machine learning techniques to identify tendencies on player behavior, proposing a metric called Commitment. The final KRI is computed based on commitment. Our goal with KRI is to strengthen the risk management in a game lifecycle.

For Rusbult et al. (1998), commitment represents a variable to understand interpersonal relationships, showing why some persist and others do not. Different from that perspective, in our context, commitment represents how attached a player is to a given game. Players are voluntary users and have a motivational factor associated with their actions (Zhu et al. 2010 ; Cook 2007). We advocate that the motivational usage can be expressed in the usage data. The commitment is mainly related to two aspects of the use: the time spent playing a game and the score achieved. To predict the commitment of a given player, we developed a machine learning approach based on two stages: an unsupervised (clustering) stage and a supervised (classification) stage. The final KRI value is computed from the commitment variation over time.

The proposed method was applied to one of the most popular MMORPGs: World of Warcraft (WoW). We used a dataset that contains three years of usage data from 91,065 unique players (Lee et al. 2011).

The remainder of this paper is organized as follows. Next, we present the relevant background details regarding game usage lifecycle. Next, some related works are presented. Then we present the risk prediction method. Experimental results are discussed. Finally, we give some conclusions and discuss future research.

Game Usage Lifecycle: Basic Concepts

Speller (2012) studied how game producers manage their games' lifecycles. After a game is released on the market, some usage data are observed aiming in identifying good or bad situations. Some examples are the rate of new players, abandonment players and profitability. The most usual metric to follow is the MAU (Monthly Active Users), which illustrates the quantity of active players play a game in one month (Figure 1 shows an example).

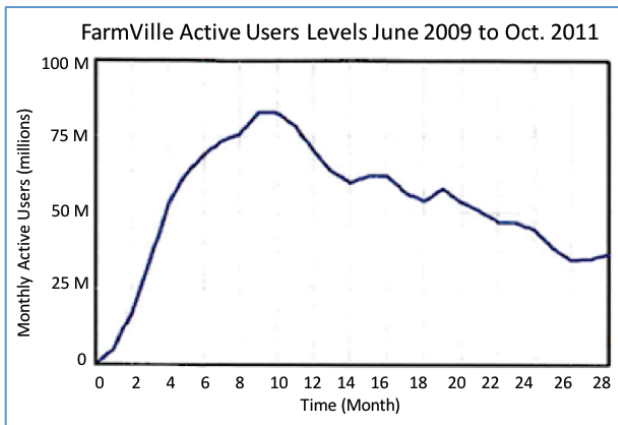


Figure 1. An example of MAU (extracted from (Speller 2012)).

As one can see in Figure 1, the number of players playing an online game changes over time. Initially, there is an excitement, just after the game release. Then, the usage slowly decreases until the end of the usage lifecycle. Zhu and colleagues (2010) studied the changes on players' motivation over time. They identified four stages: Try, Tasting, Retention, and Abandonment. The Try stage is related to the curiosity a player has for the new game. Once the game is approved the Tasting stage starts: the player spends more and more time playing the game, improving his abilities. After completing all the game's challenges, the player enters on the Retention stage. The players continue to play the game to keep their friends (social network). The game content is less important. Finally, comes the Abandonment stage, where there is no motivation for the player to play the game, leading to the abandonment. When this stage happens, the game producer has a problem, because it can represent the end of a rentable usage (MMORPGs have costs to maintain). It is important to the

game producer to identify this lack of motivation as early as possible to provide new content for the game.

Cook (2007), studied game genre lifecycles and defined a set of stages. Cook defined a genre as games with similar mechanisms of risk and award (e.g., war games, sports games, etc.). The Intro stage represents the appearance of a new genre, few games are developed and players are curious about the new challenge. After that, on the Growth stage, more games are produced if the genre is accepted, and now some players are genre fans. On the Maturity stage the genre is a great success, and great game producers start to develop games of that genre. On the Decline stage, players do not have the pleasure to play as they had before, less games are produced and then the Niche stage starts. In the Niche stage, the active players are very attached to the game, hardly ever leaving it, but unfortunately, the game is usually no more profitable.

Comparing the studies of Cook (2007) and Zhu et al. (2010), it is possible to identify some similarities between player behaviors on both models. The Intro stage of Cook represents the first behavior stage, similar to the try stage of Zhu and colleagues. The Growth stage of Cook represents the acceptance of a genre, similar to the Tasting stage, which also represent an acceptance. The Maturity and Decline stages of Cook are similar to the Retention stage, because they represent the beginning of a disgust behavior, due to the consume of the game content. At last, the Niche stage represents the same aspect of the Abandonment stage, because it is the stage where several players leave the game, only remaining very attached players.

We advocate that the game genre behavior fits with the game behavior, because initially players are interested, spend a lot of time playing and improving their abilities and then leave the game when it is not funny anymore (lack of new game content). Thus, game producer must follow usage data to early detect risky situations, such as:

- initial frustration (a new game which cannot keep players playing) (Zhu et al. 2010);
- lack of motivation over the lifecycle (Zhu et al. 2010);
- abandonment rate greater than new players rate (Speller 2012);

The aim of our study is to define and to calculate a key risk indicator, to be used with the MAU, allowing game producers to better detect the motivation of players.

Related Works

We search for studies involving machine learning and game usage data. A few works were published motivated by understanding what happens during the game usage lifecycle. One main reason for that is the lack of available real usage data. We found some studies ranging from strat-

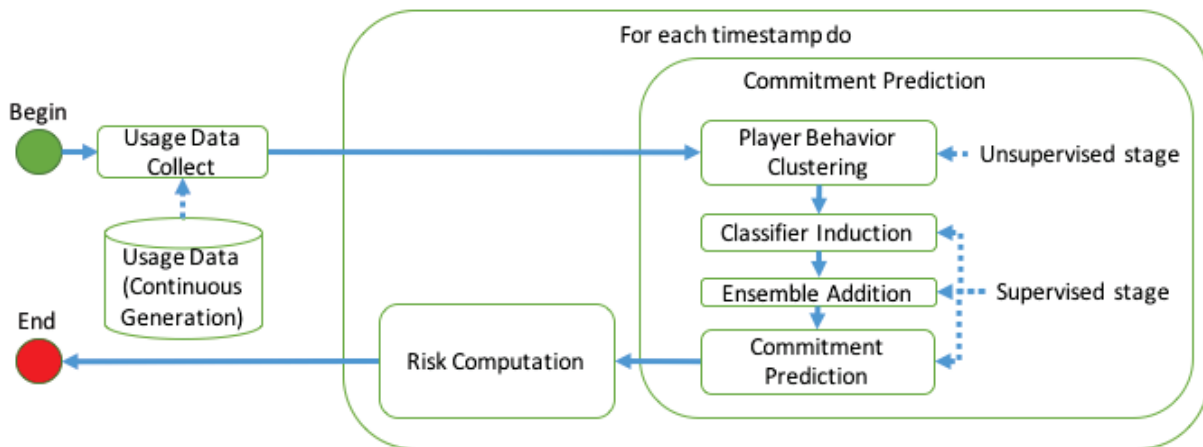


Figure 2. Risk prediction method overview.

egy prediction (Pingen and Geert 2014) to virtual player detection (Kang et al. 2013).

Speller (2012) created a dynamic system based on usage variables to predict bad, normal, or good future MAU behaviors. Tarng and colleagues (2009) created a model (similar to Support Vector Machine (SVM)) to predict player departures based on the quantity of time spent by players playing the game.

Lee and Chen (2010) used usage data to determine the quantity of hardware resource needed to maintain a game platform operating stable. The authors related usage data with energy consumption and created metrics to evaluate the operation as good or bad. Once a bad situation is identified, some actions are suggested to improve the situation.

The player behavior analysis was the goal of Drachen and colleagues (2012). First, the authors identified clusters of players' behavior based on player telemetry data (usage data). They identified very specific behaviors for FPS (First Person Shooter) and RPG (Role-playing game) games. To identify those behaviors, they used the K-means and Simplex Volume Maximization algorithms over specific game features as for example: number of kills, number of death and the number of objectives conquered.

Chiang et al. (2015) were interested in investigating the correlation between the player performance in FPS games and the network. They proved a correlation between the player score and the network quality of service.

Ghali et al. (2016) were interested in detecting if a player of a serious game is engaged or not. In their study, they collected data from three types of sensors (electroencephalography, eye tracking, and automatic facial expression recognition) to build a user adaptation system. Their work differs from ours since we use usage data from an MMORPG.

To the best of our knowledge, the study presented in this article is the first interested in proposing a key risk predictor related to motivation in MMORPGs. The works from Speller (2012), Tarng and colleagues (2009) were based on

the same object of study: the analyzes over the game lifecycle behavior. The player behavior studies of Cook (2007), Zhu and colleagues (2010) showed the motivational factor changing over time, and that changes we want to represent in the KRI through a systematic method. The next section presents the method to obtain it.

Risk Prediction Method

The main idea behind this method is that commitment can be gathered during the usage lifecycle. We define commitment as the attachment of a player to a given game. Commitment is predicted according to a two-stage machine learning approach: an unsupervised stage and a supervised stage.

The proposed method has the following assumptions:

- be game independent;
- access to available usage data containing: the player identification, the instant of time when the player played and its score.

Score is a common game feature which represents the player ability. It can vary according to the game genre (e.g., a level in an RPG, the final score of a soccer match, the final time of a race, etc.).

Figure 2 presents the pipeline of the method: commitment prediction and the risk computation. At the end, the method outputs an indicator (KRI) ranging from 0 to 1, where 1 means "the best commitment condition" for players of a game. The KRI is computed from the commitment variation over time.

The first step is the prediction of commitment. To do so, the method needs the player's scores in each timestamp. The idea behind commitment is simple: if players like the game, they will spend more time playing and improving their abilities (score), being in that way, more committed to a game. After the supervised stage, a player is evaluated as *low*, *average*, or *high* commitment. The commitment of a player *i* may be computed any time in the usage lifecycle.

To predict commitment, the vector v_i is built for each player in every timestamp:

$$v_i = \{id_i, d_i, s_{i-min}, s_{i-max}, \Delta s_i\}$$

where id_i is the player identification, d_i is the number of days the player i played the game in a given period of time, s_{i-min} is the minimum score achieved in d_i , s_{i-max} is the maximum score achieved in d_i and Δs_i is $s_{i-max} - s_{i-min}$. The model assumes that the ability (Δi) of a player i is the score (s_{i-max}) in a given timestamp t . The player changes his abilities over time, improving or reducing it, and that variation must be represented in the score feature (an assumption). Time t is incremented by 1 (one month or one day).

It is not easy to find publicly available usage datasets related to MMORPGs. The ones available do not categorize users according to their commitment. This leads us to add an unsupervised stage to the method. This first stage clusters users using their scores and the variation of that score over time. A k-means algorithm (MacQueen 1967) was used to form three groups ($k = 3$) of users (all instances were used to set the centroids' position). K-means was chosen because it is less susceptible to outliers. To label each group, we assumed that, $s_{-max-low} < s_{-max-avg} < s_{-max-high}$. Figure 3 shows the three clusters assignment.

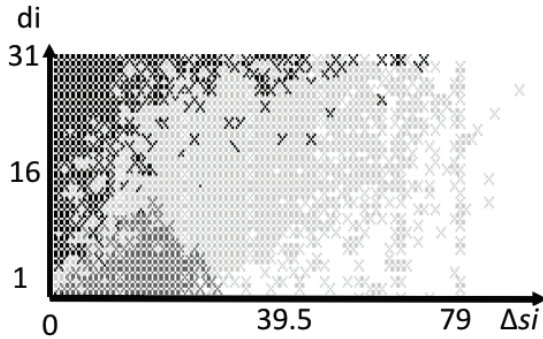


Figure 3. Commitment clusters assignment (in the case of WoW dataset (Lee et al., 2011)).

Low committed players (medium gray points in Figure 3) are players who play less time and have the lowest score. Average players (light gray points in Figure 3) play for more time and have a better score than low committed ones (they could evolve very fast) and the high committed players (dark gray points in Figure 3) are better than average ones in both aspects.

The supervised stage may start. The method has been configured to use the C4.5 algorithm (Quinlan 1993), a Decision Tree approach. By inspecting the generated tree, one can figure out what is happening in each timestamp. A classifier is inducted for each timestamp (each month).

It is important to highlight that players' profile may change from month to month. This happens because in the MMORPG scenario, external events (e.g., a vacation month, a month with several holidays, or a new game version release) may affect the interest of players. This means that values for high commitment in a month x may be different in a month y . To take this variation into account, we have adopted an ensemble classifier (Kittler et al. 1998) with a majority voting policy. In this way, each classifier gives its "opinion" on the player commitment class. The most voted class is the label of that player. The number of classifiers that forms the ensemble depends on the number of months a player plays the game (the ensemble uses only the classifiers inducted based on clustered data, it does not create new classifiers). To compute the commitment to a timestamp, a simple sum can be made for each commitment class predicted by the ensemble (equations 1, 2 and 3):

$$Low = \sum_{i=1}^n P_{low} \quad (1)$$

$$Average = \sum_{i=1}^n P_{avg} \quad (2)$$

$$High = \sum_{i=1}^n P_{high} \quad (3)$$

where P_{low} , P_{avg} and P_{high} has value 1 if the player corresponds to its degree of commitment, otherwise a 0 value is assumed, n is the number of active players in the timestamp (e.g., MAU).

The next step computes the following commitment metrics: the number of players with *low*, *average*, and *high* commitment, the number of players who changed their commitment (e.g., *low* to *average*, *average* to *low*, and so on) from one timestamp to another, and the number of players who did not change their commitment. In total, we have six commitment metrics as presented in Table 1. Those metrics are used to compute the actual risk situation.

In Table 1, Influence means the actual value of a specific commitment metric. For instance: the metric *Low to Average* indicates the number of players that moved from low commitment to average commitment. From the game producer point of view, this is positive, i.e., players are spending more time and improving their abilities.

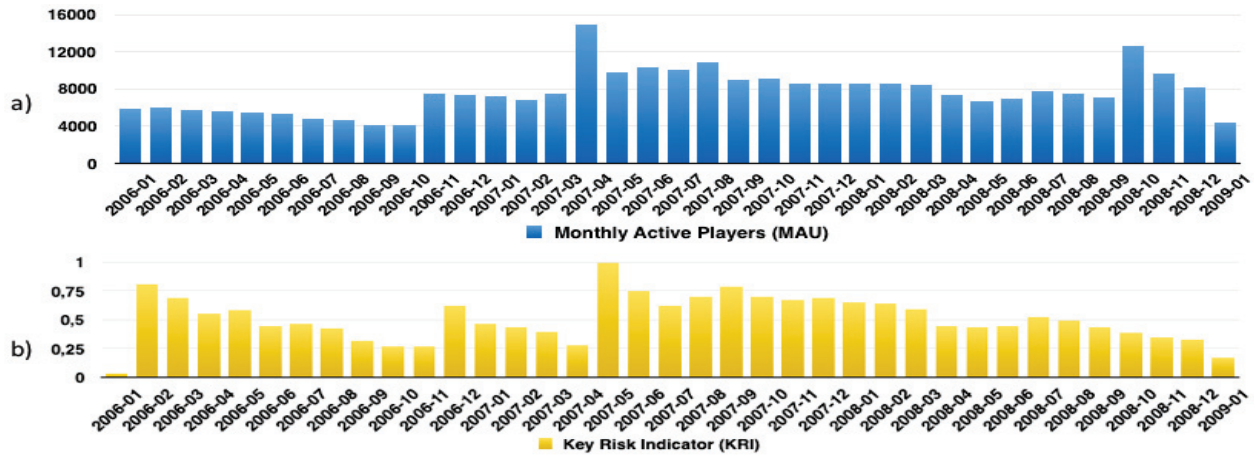


Figure 4. a) MAU (WoW dataset) b) KRI (WoW dataset).

Table 1. Commitment Metrics

Metric	Label	Influence
Low to Average	LA	Positive
Average to High	AH	Positive
Low to High	LH	Positive
Average to Low	AL	Negative
High to Average	HA	Negative
High to Low	HL	Negative

The KRI for each timestamp j is computed by equation 4:

$$KRI_j = \frac{(LA + AH + LH) - (AL + HA + HL)}{\max(KRI_n)} \quad (4)$$

where $\max(KRI_n)$ is the largest KRI already computed. For the first timestamp, KRI assumes 1.

Experimental Results and Analysis

In this section, we detail experiments to test our proposed method for risk prediction and report the results. The dataset used was obtained from (Lee et al. 2011). The WoWAH dataset (“World of Warcraft Avatar History”) contains a list of online players gathered every ten minutes. The player attributes contain the method assumptions (player identification, instant of time and score). In that case, the score is the player’s level. The dataset has 37 months of observation, 91,065 unique players and a total of 36,513,647 examples. The authors collected the usage data

through a Lua script applied to the game internal console. We pre-processed the data to change the granularity from hour/minute to monthly. The new data have a total of 282,780 instances, following the vector v_i format.

First, we evaluated which classifier should we use in the commitment’s prediction supervised stage. We selected three classifiers: C4.5 (Quinlan 1993), MLP (Rosenblatt 1961) and SVM (Cortes and Vapnik 1995). To estimate the best classifier, we compared them using a t test with $p < 0.05$. We ran experiments for each month (#37) and calculated the accuracy using 10-fold cross validation. No differences were found between SVM and MLP, however, both are better than C4.5. Besides this, the C4.5 was chosen because in a future work we want to explore the generated rules to explain the method results to users (game producers). The C4.5 achieved an accuracy rate of 98.5%.

Figure 4a plots the MAU for the WoWAH dataset. It is interesting to observe in Figure 4a how game upgrades impact on MAU. Four improvements were done over the 37 months: November, 2006, April, 2007, August, 2007 and October, 2008. After those improvements one can note an increase in MAU and few months later a decrease. That drop shows that the upgrade was not successful in maintaining players playing.

Figure 4b plots the KRI for each timestamp. Now, one can follow the motivation of players. The KRI behavior differs from the MAU as one can see on different timestamps. In November, 2006 an upgrade was done and that reflected on MAU and KRI (both increased), but a few months later the number of active users remained stable and motivation dropped drastically. This means that users still playing, but less motivated.

In October, 2008 another upgrade was released. In that occasion the MAU increased, but the KRI still dropping. In this context, an analysis of the risk indicator could show to the game producer when a game upgrade was not successful. That analysis would be impossible using only the

MAU because in terms of new players the upgrade was initially successful.

Conclusions and Future Works

The main contribution of this paper is a new key risk indicator, especially designed for MMORPGs. This risk indicator was designed to provide a high-level overview of the performance of the game. The risk indication is influenced by commitment instead of only using MAU (as usual in industry). We support that motivation is a key factor to follow and game producers should be aware of it.

The application of the proposed method showed that the commitment perspective can identify risky situations where the MAU analysis cannot. Observing the way players play is possible to identify moments of lack of motivation, even when the number of active players is satisfactory (e.g., an increase in usage after an upgrade).

In the near future, we intend to use the risk indicator to model and predict stages of the game usage lifecycle. We are also interested in studying the changes in C4.5 rules generated over time aiming in identifying patterns. We also intend to better study the unsupervised stage, analyzing the clusters formation. We will study the impact of different window size when analyzing player's commitment. In this case, instead of using all data available (37 months), use a variable number of months (the n newest months). Finally, we intend to validate the results with specialists.

References

- Chiang, C. Y., Cichocki, A., Erramilli, S., McNerney, K., Shur, D. and Loeb, S. 2015. Prediction of online game performance degradation under network impairments. In Consumer Communications and Networking Conference (CCNC), 2015 12th Annual IEEE, 720-725.
- Cook, D. 2007. *The Circle of Life: An Analysis of the Game Product Lifecycle*. Gamasutra (2007). Visited in 2016: http://www.gamasutra.com/view/feature/129880/the_circle_of_life_an_analysis_of_.php.
- Cortes, C., Vapnik, V. 1995. Support-vector networks. *Machine Learning*, 20 (3): 273–297
- Drachen, A., Sifa, R., Bauckhage, C., and Thurau, C. 2012. Guns, swords and data: Clustering of player behavior in computer games in the wild. In 2012 IEEE conference on Computational Intelligence and Games (CIG), 163-170.
- Galway, L., Charles, D., and Black, M., 2008. Machine learning in digital games: a survey. *Artificial Intelligence Review*, Volume 29, Issue 2, pp 123-161.
- Ghali, R., Frasson, C., and Ouellet, S. 2016. Towards Real Time Detection of Learners' Need of Help in Serious Games. *FLAIRS Conference*: 154-157.
- Kang, A. R., Woo, J., Park, J., and Kim, H. K. 2013. Online game bot detection based on party-play log analysis. *Computers & Mathematics with Applications*, 65(9), 1384-1395.
- Kittler, J., Hatef, M., Duin, R.P.W., Matas J. 1998. On combining classifiers. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 20(3):226-239.
- Lee, Y. and Chen, K. 2010. Is Server Consolidation Beneficial to MMORPG? A Case Study of World of Warcraft, *Proceedings of IEEE CLOUD*, 435-442.
- MacQueen, J. B. 1967. Some Methods for classification and Analysis of Multivariate Observations. *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability*. University of California Press. pp. 281–297
- Quinlan, J. R. 1993. C4.5: *Programs for Machine Learning*. Morgan Kaufmann Publishers.
- Rosenblatt, Frank. 1961. *Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms*. Spartan Books, Washington DC.
- Sheffield, B., and Alexander, L. 2008. Focus On Korea: T3's Kim Talks Hitmaking With Audition. Gamasutra (2007). Visited in 2016: http://www.gamasutra.com/view/news/107795/Focus_On_Korea_T3s_Kim_Talks_Hitmaking_With_Audition.php.
- Silver, D., Huang, A., Maddison, C.J., Guez, A., Sifre, L., Driessche, G.V.D., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T., and Hassabis, D. 2016. *Mastering the game of Go with deep neural networks and tree search*. *Nature* 529, 484– 489.
- Tarng, P.Y., Chen, K.T., and Huang, P. 2009. On Prophesying Online Gamer Departure. *Proceedings of ACM NetGames*, 16.
- Lee, Y.T., Chen, K.T., Cheng, Y.M., and Lei, C.L. 2011. World of Warcraft Avatar History Dataset. *Proceedings of ACM Multimedia Systems 2011*, 123-128.
- Pingen and Geert L.J. 2014. An implementation of a data-mining approach to strategy selection in Starcraft: Brood War, *Dissertation presented at Radboud University Nijmegen*.
- Rusbult, C.E., Martz, J.M., and Agnew, C.A. 1998. The investment model scale: measuring commitment level, satisfaction level, quality of alternatives, and investment size. *Personal Relationships* 5(4):357 – 387.
- Speller, T.H. 2012. The business and dynamics of free-to-play social-casual game apps. Doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA.
- Zhu, L., Li, Y., and Zhao, G. 2010. Exploring the Online-Game Life Cycle Stages. *IEEE International Conference on E-Business and E-Government (ICEE)*, 2436-2438.