Player Performance Prediction in Massively Multiplayer Online Role-Playing Games (MMORPGs)

Kyong Jin Shim, Richa Sharan, and Jaideep Srivastava

Department of Computer Science and Engineering
University of Minnesota
200 Union Street Southeast
Minneapolis, MN 55455, USA
{kjshim,srivasta}@cs.umn.edu, sharan.richa@gmail.com

Abstract. In this study, we propose a comprehensive performance management tool for measuring and reporting operational activities of game players. This study uses performance data of game players in EverQuest II, a popular MMORPG developed by Sony Online Entertainment, to build performance prediction models for game players. The prediction models provide a projection of player's future performance based on his past performance, which is expected to be a useful addition to existing player performance monitoring tools. First, we show that variations of PECOTA [2] and MARCEL [3], two most popular baseball home run prediction methods, can be used for game player performance prediction. Second, we evaluate the effects of varying lengths of past performance and show that past performance can be a good predictor of future performance up to a certain degree. Third, we show that game players do not regress towards the mean and that prediction models built on buckets using discretization based on binning and histograms lead to higher prediction coverage.

1 Introduction

Massively Multiplayer Online Role-Playing Games (MMORPGs) are personal computer or console-based digital games where thousands of players can simultaneously sign on to the same online, persistent virtual world to interact and collaborate with each other through their in-game characters. This study is concerned with forecasting of player performance in the game. While many games today provide web and GUI-based reports and dashboards for monitoring player performance, we propose a more comprehensive performance management tool (i.e. player scorecards) for measuring and reporting operational activities of game players. This study uses operational and process-oriented performance data of game players in EverQuest II, a popular MMORPG developed by Sony Online Entertainment, to build performance prediction models for game players. First, we show that variations of PECOTA [2] and MARCEL [3], two most popular baseball home run prediction methods, can be used for game player performance

M.J. Zaki et al. (Eds.): PAKDD 2010, Part II, LNAI 6119, pp. 71–80, 2010.

prediction. Second, we evaluate the effects of varying lengths of past performance and show that past performance can be a good predictor of future performance up to a certain degree. Third, we show that game players do not regress towards the mean and that prediction models built on buckets using discretization based on binning and histograms lead to higher prediction accuracy.

Systematic studies of game player performance is expected to yield the following contributions. First, analysis of player performance in different dimensions (i.e. player demographics, archetypes, classes, sub-classes) can help game developers understand whether their games and game characters are being played as intended. Second, benefits for game players are two fold. a) Game players can not only have a view of their past and current performance but also they can have a view of their projected future performance. b) A recommendation engine can be built to recommend character types and tasks to players in order to meet certain objectives (i.e. move up to the next level as fast as possible, play safe by attempting easy tasks, play aggressively by attempting challenging tasks, play tasks that encourage grouping with other players). Third, players can have a view of performances of other players for the purposes of forming quest or raid teams.

2 EverQuest II Game Mechanics

2.1 Point-Scaling System in EverQuest II

In EverQuest II, there is a concept of Ding Points, which is the amount of points one needs to obtain in order to move from one level to the next higher level [4]. For instance, to move from Level 2 to Level 3, one needs to obtain 1,000 points whereas 20,000 points are required to move from Level 73 to 74. The amount of ding points increases as one advances to the next level. As players gain more experience with the game and advance to higher levels, the types of task they can perform increase and the task difficulty also increases. The higher the task difficulty, the higher the potential point gain.

2.2 Tasks in EverQuest II

EverQuest II is rich in types of task players can perform with monster kills being one of the most popular. Monster kills are discussed in details in [1]. In addition to monster kills, other sources of experience points exist in the game such as alternate achievement points (AA) which can be obtained from quests, named mobs, and discovery experience. A player can gain more experience points by having another player mentor him. The mentor levels down to the level of the mentee. The mentee receives a five percent bonus to adventuring experience points.

2.3 Archetypes, Classes, and Sub-classes in EverQuest II

In playing MMORPGs, selection of character type (i.e. archetype, class, subclass, and race) is considered an important decision as it defines the basis of opportunities and choices of roles and tasks within the game [5]. In EverQuest II, there are four archetypes where each archetype consists of three classes each of which in turn consists of two sub-classes [4]. Performance comparisons are discussed in details in [7].

3 Baseball Home Run Prediction

Prediction of future performance of humans has long been studied in various disciplines over the years. Most notably, it has been well studied in sports. Baseball has a long history of record keeping and statistical analyses that dates back to the nineteenth century. Batting average, RBIs, and home runs are some of the many statistics being kept track of today. There exists an enormous amount of public and private interest in the projection of future performance. Major league teams rely on the past statistics of a given player in deciding whether to acquire him or not and for how many seasons under the assumption that his past success is a good indicator of his future success.

PECOTA [2] and MARCEL [3] are widely known methods in baseball home run prediction. PECOTA [2] is considered a very sophisticated method for home run prediction in baseball. For a given ball player at the age of X, the method uses a nearest neighbor analysis of both minor and major league players from the past that exhibited similar performance at age X. It uses the historical performance of these past players to predict the given player's future performance. MARCEL [3] uses data from the three immediate past seasons of a given ball player, and it assigns more weight to more recent seasons. One drawback of this approach is that prediction models solely based on individual players cannot be generalized to the global population. A variation of the MARCEL approach attempts to regress predictions to the global population mean. One drawback of this approach is that prediction models built on the global population can become too coarse.

We consider game player levels in EverQuest II similar to seasons in baseball. Players perform tasks, gain points, and move up to the next level as ball players would attain different types of achievement (i.e. home runs, single, double, triple hits, run batted in, etc.) in each season and proceed to the next season. Unlike in baseball where there is not necessarily a fixed number of home runs, triples, doubles, etc. required to move to the next season, EverQuest II employs a point scaling system where there exists a fixed number of experience points at each level in order to move up to the next level. Because the ding point is a fixed constant, we measure a game player's total play time at each level and uses it as a performance measure in this study.

4 Player Performance Prediction in EverQuest II

In this study, we develop performance prediction models for game players in EverQuest II. The objective is to predict a given player's play time at level i, a future state, based on his past performance at levels $i-1$, $i-2$, and so forth, where performance at any level is measured as the total play time spent at that level. Play time in EverQuest II excludes any idle periods where being idle is defined as any contiguous time blocks of 30 minutes or beyond.

4.1 Methods

MARCEL [3] method uses data from the three immediate past seasons of a given ball player, and it assigns more weight to more recent seasons. One drawback of this approach is that prediction models solely based on individual players cannot be generalized to the global population. A variation of the MARCEL approach attempts to regress predictions to the global population mean. One drawback of this approach is that prediction models built on the global population can become too coarse. Algorithm 1 [7] delineates the steps taken to generate MARCEL-like prediction models for game player performance prediction.

Our preliminary data analysis of the game data reports that play times at each player level exhibit a skewed distribution [7]. EverQuest II game players do not regress towards the mean, and therefore prediction models built under the assumption that they do regress towards the mean will become too coarse and will perform poorly for players whose performances deviate from the mean. To overcome this problem, for a given player, PECOTA [2] uses past performance of those players whose performance patterns are similar to that of the given player.

In this study, we perform data discretization based on two unsupervised techniques, binning and histogram analysis, in order to create buckets of players where all players in a given bucket are termed neighbors. Neighbors share similar performance patterns, and a prediction model is built for each bucket. This is similar to the way PECOTA [2] uses a nearest neighbor analysis to group players into buckets and builds a prediction model for each bucket. Algorithms 2 and 3 [7] delineate the steps taken to create buckets based on binning and histogram analysis, respectively. Algorithms 4 and 5 [7] delineate the steps taken to create MARCEL-like prediction models and Regression-based prediction models, respectively, both using discretization.

4.2 Dataset

The study uses one month worth of performance data from March 1, 2006 to March 31, 2006. The dataset contains over 36 million player-to-task records where over four million of them are monster kills related tasks. The dataset contains 24,571 distinct players across player levels 1 through 70. Since then, Sony Online Entertainment has added an additional ten levels to the game, making 80 the maximum level one can reach.

All of the players and their performance data has been extracted from XP table in the EverQuest II database housed at National Center for Supercomputing Applications (NCSA) at the University of Illinois. The dataset contains at the minimum the following information about game players: character id, character sub-class, race, task, timestamp of task completion, group size (whether a given character grouped with one or more other characters), average group level (if a given character played with one or more other characters, this value represents the average of player levels of all players involved in that group), experience points, location (location in which the task was completed).

4.3 Evaluation

In prediction (i.e. regression, time series analysis, etc.), a common practice has been to specify coverage probabilities by convention, 90%, 95%, and 99% being typical choices. A previous study [6] reports that academic writers concentrate on 95% intervals while practical forecasters prefer 50% intervals. In this study, we compute prediction coverage at varying confidence intervals at 80% and 90%. Algorithm 6 [7] delineates the steps taken to compute prediction coverage.

5 Experiments and Results

5.1 Past Performance as Indicator of Future Performance

A series of experiments have consistently shown that the three immediate past levels contribute the most to the prediction of a player's future performance. Extending beyond the three immediate past levels does not positively contribute to prediction coverage. One possible explanation might be that game players, in playing tasks such as monster kills in EverQuest II, do not tend to degrade in their performance suddenly, and therefore, a given player's performance at the most recent level $(i - 1)$ should be most informative about his performance at the current level (i). However, this may not necessarily be true in all cases such as when a player all of a sudden decides to attempt monsters whose levels are far beyond average, in which case, the player's performance at the current level may degrade due to the fact that his skill level is suddenly not matching the task difficulty. Additionally, we try a variety of weighting schemes for use with MARCEL [3] approach. Broadly, weighting functions are categorized into 1) even weight distribution and 2) decaying weight distribution. The former assigns an equal amount of confidence to each of the past levels whereas the latter assigns more weight to more immediate past levels. Our findings suggest that with the three immediate past levels, both even weight distribution and decaying weight distribution produce comparative results.

5.2 Discretization Improves Prediction Coverage

Given the dataset used in our analysis, our findings suggest that the bucket number of six leads to high prediction coverage. In some player levels though we observe that a bucket number slightly lower or higher than six leads to even higher prediction coverage.

Our results show that discretization using binning and histogram analysis leads to higher prediction coverage overall across all 70 player levels where the number of buckets is six. Figure 1 shows that MARCEL [3] approach produces

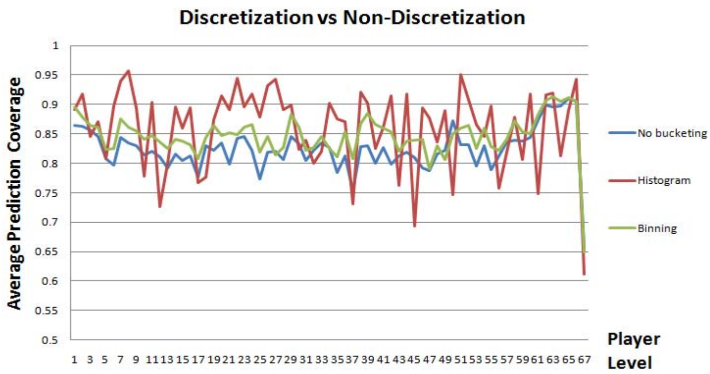

Fig. 1. Discretization Improves Prediction Coverage (MARCEL approach)

an average prediction coverage of 82.4% whereas the same approach employing binning produces 84.7% and that employing histogram analysis produces 86% prediction coverage (confidence interval of 80%). Figure 2 shows results consistent with MARCEL approach where the base linear regression model produces an average prediction coverage of 83.2% whereas the model employing binning produces 85% and that employing histogram analysis produces 85.7% prediction coverage (confidence interval of 80%).

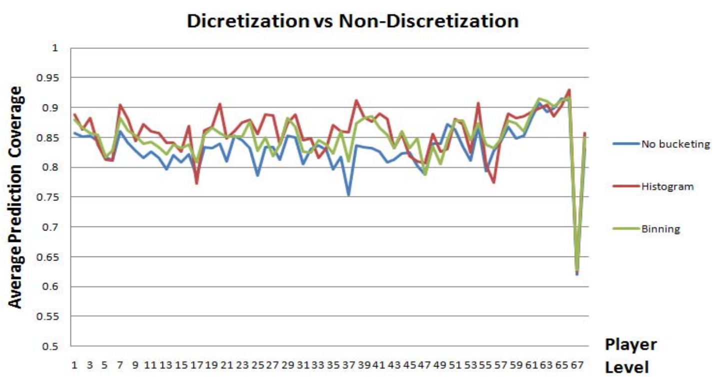

Fig. 2. Discretization Improves Prediction Coverage (Linear Regression)

5.3 Comparison of Prediction Models

Figure 3 shows prediction coverage computed at confidence interval of 80%. MARCEL [3] approach in combination with histogram-based discretization performs the best while all other schemes produce results that are comparative to that of MARCEL [3] approach.

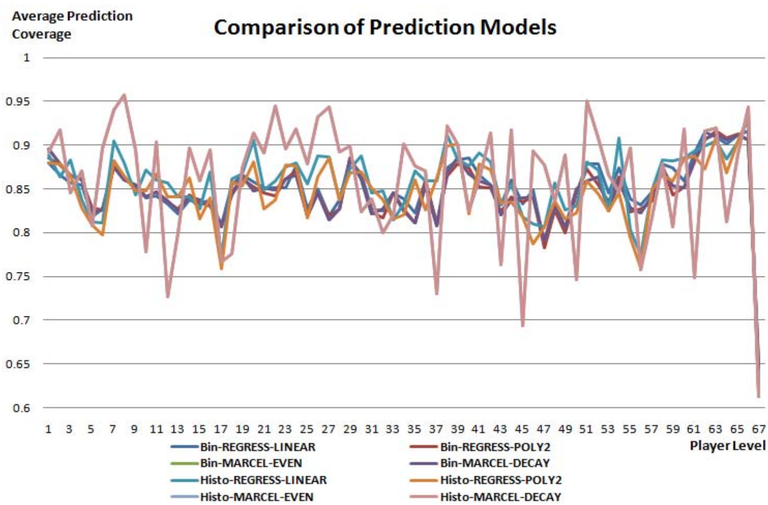

Fig. 3. Comparison of Prediction Models (80% Interval)

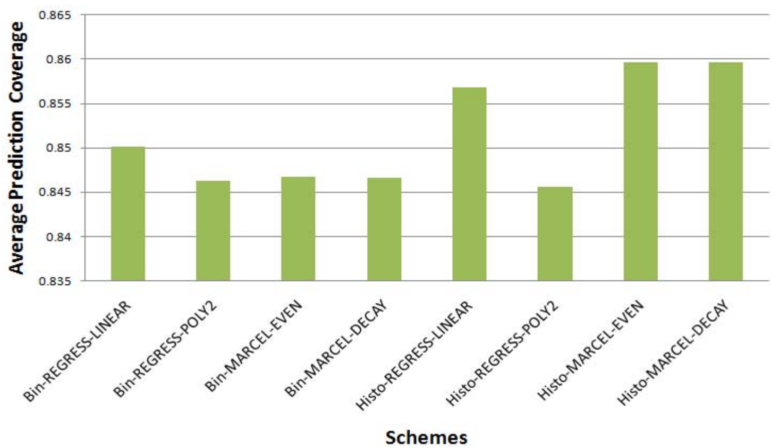

Fig. 4. Comparison of Prediction Models (80% Interval)

Figure 4 charts the average prediction coverage computed at confidence interval of 80% across 70 player levels. MARCEL [3] approach in combination with histogram-based discretization performs the best while all other schemes produce results that are comparative to that of MARCEL [3] approach.

Figure 5 shows prediction coverage computed at confidence interval of 90%. Linear regression model in combination with binning-based discretization performs the best while all other schemes produce results that are comparative to that of linear regression model.

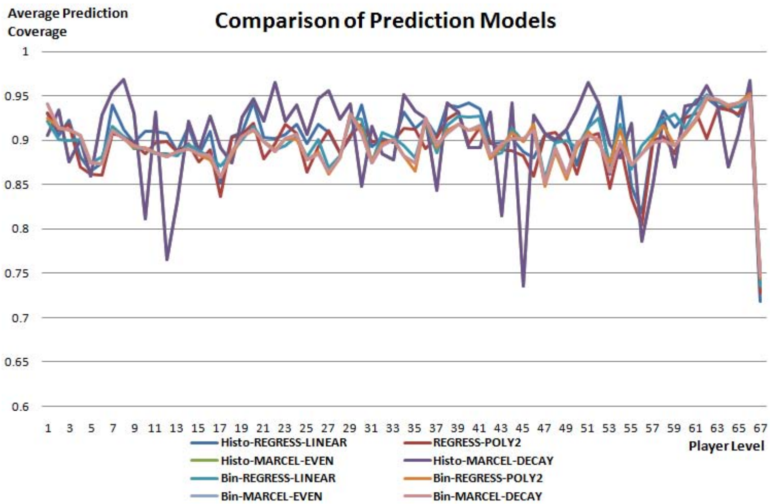

Fig. 5. Comparison of Prediction Models (90% Interval)

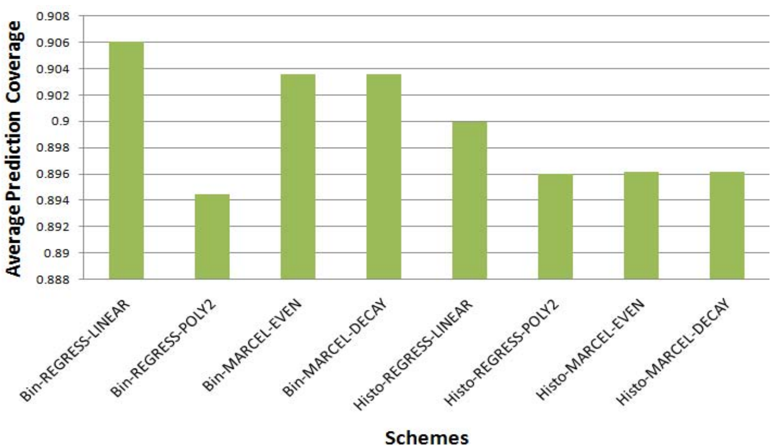

Fig. 6. Comparison of Prediction Models (90% Interval)

Figure 6 charts the average prediction coverage computed at confidence interval of 90% across 70 player levels. Linear regression model in combination with binning-based discretization performs the best while all other schemes produce results that are comparative to that of linear regression model.

Our prediction models capture information essential about the relationship between progression of player level and progression of player performance (as a function of play time) over a range of three player levels. Our results consistently show that the relationship is linear to a certain extent. This trend is observed across all 70 player levels.

6 Conclusion

In this paper, we show that variations of PECOTA [2] and MARCEL [3], two most popular baseball home run prediction methods, can be used for game player performance prediction. MARCEL approach in combination with bucketing inspired from PECOTA approach leads to high prediction coverage. The method uses data from the three immediate past levels and assigns more weight to more recent levels. In game player performance prediction, our findings suggest that the results from even weight distribution and decay weight distribution are comparative. To account for an observation that game players in EverQuest II do not regress towards the mean in terms of their play times, prediction models are built on buckets using discretization based on binning and histograms. This approach leads to higher prediction coverage. Further, we build regression-based models and show that the relationship between progression of player level and progression of player performance (as a function of play time) over a range of time is linear to a certain extent. The regression-based models produce prediction coverage comparative to that of existing methods.

Prediction models we propose in this study are expected to be a useful addition to many existing player performance monitoring tools by providing a projection of a given player's future performance given his past performance. Game player performance data such as that of EverQuest II is rich of not only outcome data (i.e. number of monsters killed, number of experience points gained, number of deaths occurred, number of quests completed in a given time duration) but also process data, from which we can construct a progression of a given player's performance at any given time point. Existing player performance monitoring tools can be greatly enhanced to dynamically capture player performance progression, provide instant feedback on player's progress, and recommend tasks tailored towards a given player's objectives of playing the game (performance-oriented tasks vs. social activity-oriented).

7 Future Directions

An extension to the current work involves investigating model dynamics by examining the balancing of past consistency with advancing player level. An issue arises when a player performs way below the average for a couple of levels and springs back up to a very good performance. All of the prediction models discussed in this study so far lack the ability to integrate such dynamics into prediction. Another extension to the present study seeks to define performance in many dimensions of different granularity levels (i.e. task types, archetypes, classes, sub-classes, races, roles, etc.). For instance, the present study defines performance as a function of play time or active time. Another measure of performance is the level of consistency and commitment. Results from such analyses can reveal player behavioral patterns indicative of player churning. Yet another addition to this study is to leverage a variety of social networks in EverQuest II (i.e. housing network, trust network, raid group network, and guild network) to measure the impact of social interactions on player performance.

Acknowledgments

The research reported herein was supported by the National Science Foundation via award number IIS-0729421, and the Army Research Institute via award number W91WAW-08-C-0106. The data used for this research was provided by the SONY corporation. We gratefully acknowledge all our sponsors. The findings presented do not in any way represent, either directly or through implication, the policies of these organizations. Acknowledgements are also due to the members of the University of Minnesota, Northwestern University, University of Southern California, University of Illinois, University of Michigan, and University of Chicago.

References

1. Shim, K.J., Ahmad, M., Pathak, N., Srivastava, J.: Inferring Player Rating from Performance Data in Massively Multiplayer Online Role-Playing Games (MMORPGs), cse. In: International Conference on Computational Science and Engineering, vol. 4, pp. 1199–1204 (2009)
2. Silver, N.: Introducing PECOTA, pp. 507–514. Baseball Prospectus, Hyattsville (2003)
3. Tango, T.: Marcel The Monkey Forecasting System. Tangotiger .net, `http://www.tangotiger.net/archives/stud0346.shtml` (March 10, 2004)
4. Prima Development. Everquest II: Official Strategy Guide (Prima Official Game Guides)
5. `http://pc.gamespy.com/pc/everquest-ii/564411p1.html`
6. Granger, C.W.J.: Can We Improve the Perceived Quality of Economic Forecasts? Journal of Applied Econometrics 11, 455–473 (1996)
7. Shim, K.J., Sharan, R., Srivastava, J.: Player Performance Prediction in Massively Multiplayer Online Role-Playing Games (MMORPGs), Technical Report 10-003, Department of Computer Science and Engineering, University of Minnesota