Investigation of game players' strategies using visualization techniques and data mining

Behnam Dorry[1], Seyed Abolghasem Mirroshandel[1*], and Abdollah Esmaeilpour Chirani[1]
1- University of Guilan, Rasht, Iran.
Email: mirroshandel@guilan.ac.ir (*Corresponding author)

ABSTRACT:
Nowadays, generating a profitable game is a big challenge. Thousands of commercial products are generated annually that compete with each other for getting more market share. Many tools and techniques, such as project management and user testing are used for effective game development. In recent years, game developers take notice to gamers' behavior and tend to design and modify games based on their attitude. In this paper, we present an approach called UPB, for analyzing computer games based on telemetry data, statistical information, and data mining. The goal is developing games in a way that improves user experience and increases player enjoyment, so more gamers would welcome them. For this purpose, a telemetry data gathering system and a software for analyzing telemetry data were developed. A data mining process was also conducted to discover the behavior patterns of the game players based on the telemetry data collected from them. For this purpose, we designed a new game, called "Treasure Castle", to show the effectiveness of UPB approach and data gathering. The difficulty of game levels was determined based on the behavior of the players. In addition, the playing style of players were visualized and their strategies were detected by examining how they play. Our experiments were done on the Treasure Castle game and it has been shown that the proposed approach can answer different questions about understanding players' behavior.

KEYWORDS:
Telemetry, Action Games, Gamers Behavior, Tracking Strategies, data mining, visualization.

1. INTRODUCTION

Computer games have become one of the most popular hobbies for people [1]. They are aimed at creating better and more enjoyable experiences for players, and can motivate emotional and cognitive processes [2, 3]. In addition, games are also considered as modern and effective platforms for learning and education [4, 5].

Due to development of electronic, computer, and communications technologies, computer game industry has experienced a great growth trend and expanded worldwide [1]. Data analysis is a necessity for survival of companies in current high competitive business environment that has resulted in emergence of business intelligence. In the computer games industry, a large number of video games are introduced. Producers are trying to increase their market share. This motivates researchers and game developers to examine and analyze the behavior of players in game environment and to understand the relationship between their functions and satisfaction. These analyses are not just for discovering this relationship, but help game designers to identify pros and cons of their games. As a result, their products can be improved and the company's standing in the industry will be enhanced. It shows that data gathering and analysis plays an important role in success of game business.

Games were previously analyzed to find the bugs, however, their analysis are currently conducted to obtain feedback to improve their design. Because of the efforts made in the game industry, the number and diversity of game players highly grew up. Therefore, game makers saw a great opportunity to create various games with a wide variety of genres, which are for various ages of gamers having different levels of ability, intelligence, and creativity.

In this paper, we have designed a new game, called Treasure Castle and an approach for Understanding Players' Behavior called UPB that was examined on the Treasure Castle game. Game analysis was performed by tracking players and based on telemetry data. These analyses include the visualization of events such as the failure of players at different levels of the game using tools such as heat map charts, as well as the extraction of statistical information such as the average playing time of each level and other events. However, the turning point of this article is to get the players' strategies to complete the game by means of data mining techniques. UPB approach has the following advantages:

- This approach has been tested on a real game with statistical community of 252 players that is an appropriate statistical community for test purposes.
- The statistical results have been illustrated to give a clear sight, for example, death event

تاریخ دریافت: ۱۳۹۸/۰۳/۱۵ تاریخ پذیرش: ۱۳۹۹/۰۶/۱۶

percentage, illustrated in Fig. 13, the average of playing time of each level, illustrated in Fig. 14, and other examples throughout the paper.

- Players' strategies have been examined using data mining techniques.
- UPB approach detects the difficulty of the game in preproduction phase so game developers know to what extent the implementation matches the initial plan.
- By discovering different styles of playing the game, the approach helps game developers to detect cheating or skilled players. This difference is detected by the time of finishing a level, the selected path to in each game level, acquiring helping objects, and connecting to game NPCs.

Our proposed method also has the following disadvantages:

- Generalizability and extensibility of the UPB approach need more consideration in future works

The remaining of this paper is organized as follows.

Section 2 provides a background on game design terminology and some related works on the design and evaluation of games Section 3 introduces a game called "The Treasure Castle" that is developed in our research. In Section 4, the steps of the UPB approach described. The final section is the conclusion of our paper.

2. BACKGROUND

Game designers need to understand the points of view of end users and consider their experience of dealing with games in future developments. This is a game user research (GUR) that is one of the practical aspects of game analysis [6]. Although very novel, GUR plays an important role in game development process [7]. This new science has a close relationship with Business Intelligence (BI), and it can be argued that Business Intelligence provides a platform for analyzing computer games. Although the definition does not cover all the analytical aspects of games and does not cover its relation to BI, it can be considered as a good guide in this area.

In this article, a few words are frequently used that should be clearly defined to avoid confusion. They include Game Analytics, Game Telemetry and Game Metrics. Game Analytics is the process of research and development of games. The goal of this analysis is to support the decision-making process in tactics, strategy and operations at all levels of the organization, such as design, arts, programming, marketing and research on players [8]. Telemetry is remote data gathering. The data can be quantitative and provide us with information on how a user plays. It is sent from client systems to some servers [9]. Game Metrics are interpretable measurements of the attributes of objects [10, 11]. A common source of game metrics is telemetry data that results from the behavior of players in the game or their interaction with the game. In fact, the game analytics can be extracted from the telemetry data. For example, in a shooting game, the number of gunshots is telemetry data, and the percentage of gunshots hit the target is game metrics. Another example is the start time of the game and its end time. They are considered as telemetry data, but the amount of time the player continues to play is treated as game metrics.

Game companies invest in data gathering and test their designs to optimize their products. The optimization includes improvements to the game environment, racing track [12], the levels of platform games [13], storytelling [14], improved game rules for board games [15, 16] and games such as PAC-MAN [17] and technical issues such as camera control parameters such as height, distance and so on [18, 19] The user interface is a key components of computer games. Some feedback is provided to gamers by current user interfaces such as coins, pills, scores, lives, blood splashing, and speedometer. This feedback helps to improve the game and create better experience for players. Today, feedback is considered as an important feature of computer games as it provides useful information about player performance [4, 5]. The old computer games provided very poor feedback due to their limited graphical capabilities and computational power at that time. For example, the skill games ("Pac-Man" -1980, "Space Invaders" -1978, "Pong" -1972, "Space war" - 1961, "Tennis for Two" -1958), show points (experience or level scores) and counters displaying level, remaining life, time or highest record. Recently, some aspects of feedback have been taken into account to improve the user experience of players [20]. As an example, the eagle vision is used in Assassin's Creed to highlight invaders and their patrol paths. Blood splashing in "Call of Duty" denotes that the character of the game is losing their health. In addition, modern computer games reveal players' behavior for designers and researchers using data analysis techniques. This feedback is useful for enhancing the computer games design and implementation, and comprehending the behavior of players concerning the level of the game. Data mining techniques and statistical information help to know the strategies used by players and to improve artificial intelligence approaches in games [21, 22].

Research on user behavior involves several approaches, techniques, and methodologies, such as psychological research, machine learning, and human–computer interaction. The goal of this research is to know how games are played by players and how players interact with games [23, 24]. Also, the user-oriented game testing [25] has been introduced over the years and has been an important part of game production [26-29]

In recent years, a series of new structured methodologies of human-computer interaction field were adapted for research in the field of user-oriented game testing, playtesting and usability-testing [30].
In 2017, researchers used Kolb's theory of experiential learning to detect different playing styles. They utilized multiple linear regression to show how players' personality traits are [31]. In 2018, modeling of individual differences in the behavior of playing was performed using the Hidden Markov Model (HMM). Using the HMM output, the behavioral attributes are generated that are used to classify the players' characteristics [32]. The idea of creating clusters from tracking previous games to obtain different playing styles has been implemented in game Tetris. These clusters were used to decide how much the system should help new players to have good pieces [33]. Another approach is adjusting the difficulty of the game adaptively regarding serious games using the player performance model. [34]

The main theme of the telemetry data used in this paper is the user telemetry data. User telemetry data can include data that is obtained from gamers' behavior. This behavior includes player interaction with games, the examination of physical movements by means of joy sticks, the interaction of players with each other in game environment (in online games and usually in the MMO style) or interaction of players with a non-player character (NPC) [35-38]. The user metrics or player metrics are related to the human user or the person who plays, and are considered from two points of view: the customer and the player. The first aspect, which is the aspect of being customer, addresses issues that are related to financial interests, such as the average revenue per user (ARPU), the number of active users per day, customer support efficiency analysis, and so forth. They are the metrics which are examined in game companies. The second aspect relates to the way in which people deal with or interact with a game. In other words, it focuses on examining the behavior of the players in the game, and in this paper the latter point of view is considered and the former one is out of its scope.

In this paper, the statistical data of players' gameplay and data mining were used to understand the difficulty of game levels. Moreover, it illustrates different strategies of gameplay of the levels by visualizing the results of data mining. Discovering the different gameplays helps game developers to investigate issues such as cheating or identifying skilled players.

3. EXPERIMENTAL STUDY

In order to test the efficiency of the proposed approach, we have developed a new game named Treasure Castle. To perform our experiments, 252 people were asked to play the Treasure Castle game. It is a two-dimensional game which has ten levels. Important metrics and data for the analysis of games that contain character actions, generally focus on fast reactions, accuracy, and time. Such games are usually single-character. The game is free and automatically sends behavioral data of players to a server for which was intended. Fig. 1 shows a screenshot of the game. It should be noted that this game and the images related to data mining are available on cloud[1]:

In this game, players lead a boy that is the main character of the game. Each level of the game ends when the player reaches the exit door of that level. The player can find and use weapons such as swords, guns and dynamite. Things like boxes, lifts, and diving hats are also available. There are several types of non-player characters (NPCs), such as Angry Cloud and Angle. The first two levels of the game make players familiar with the game environment. In level 3 there are more than one path to complete the level. The game runs on the Windows operating system.

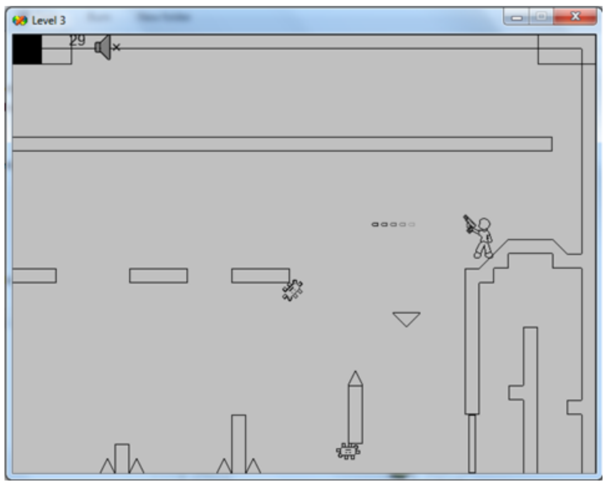

Fig. 1. A screenshot of the third level of the Treasure Castle game

4. THE UBP APPROACH

Understanding the behavior of gamers plays an important role in game development, because it shows that whether games have been well designed and implemented. Also, it shows that if the interaction between games and players are done correctly.

Therefore, some principles need to be defined before entering into the discussion, because understanding a concept needs to satisfy some levels of knowledge. In other words, if after the research, we can answer the questions raised by game designers, it

[1] https://drive.google.com/open?id=19_xnoNLJwyrkt1nPsIqoE2VNg2VJB9bf

shows that we understand the behavior of gamers. The questions are as follow:

- In which region of the game, do players play more?
- Which game configuration i.e. tools & objects is most used by players?
- What route does the player choose to finish the game?
- Which objects are found in the game? For example, was the sword found?
- What percentage of each event occurred in each level?

Fig. 2 illustrates the UPB approach for understanding players' behavior. It shows that the gameplay of each player is stored in a telemetry file to be processed. For next steps, data are integrated into a database. Now, there are three ways to achieve the goal, which is understanding the behavior of users: first, visualizing the data and understanding the visualization intuitively, the second, the extraction of statistical data from telemetry data and the third, by means of data mining techniques.

It should be kept in mind that the output of each method itself provides an understanding of how the players behave in the game, and all three methods are accomplished in this paper, but the best method is that all the three processes are applied on data so that a very clear view is reached for game designers.

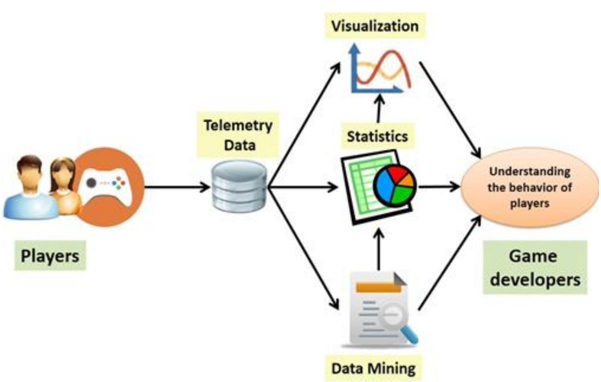

Fig. 2. An approach for understanding players' behavior (UPB)

4.1. Collecting and processing telemetry data

The telemetry data of this game is gathered concerning events such as starting the game, ending the game, squeezing the arrow keys, and pressing and releasing the other keys. The data stores on the client side and recorded in a file. If the character's locations in specific time intervals were considered, the volume of telemetry data would be enormous and it would be difficult to transfer it to the server. There is no need to add an event to know the current location of the character and his path, because each activity in this 2D game has specific coordinates (in axes X and Y). A number of these events are given in Table 1.

Event	Telemetry data	Sample
The first run of the game	assign a character number to the player	ID=7854239279
Start of the game	record the start time of the game	StartGame(10:10:41)
End of the game	record the end time of the game	EndGame(10:19:50)
Start of each level of the game	record the start time of the level	StartRoom> Level1(10:11:6)
End of each level of the game	record the end time of the level	EndRoom> Level1(10:11:10)
The player starts to move to the left	record the time and location the player the left arrow key	PressLeft(640|480)(10:11:8)
The player jumps	record the time and location where the player presses the up arrow key	PressUP(588|480)(10:11:9)
The player starts to move to the right	record the time and location where the player presses the right arrow key	PressRight(640|480)(6:31:1)
The player presses the button down	record the time and location where the player presses the bottom arrow key	PressDown(406|480) (6:36:47)
Player Death Event	Save the character's death location	die(839|480)
The player starts running	record the time and location the character starts running	PressR(816|256)(7:14:44)
Player stops running	record the time and location the character stops running	ReleaseR(816|256)(7:14:44)
Event of picking up and dropping objects in the game	record the time and location of the event	PressE(1095|944)(7:17:43)
Use of weapons	record the time and location of the event	PressZ(293|944)(7:18:29)
Start quick mode	record the time and location that the player starts using this ability	PressSpace(1232|695) (7:23:13)
End quick mode	record the time and location that the player stops using this ability	ReleaseSpace(1232|467) (7:23:14)
Calling an NPC called Angry Cloud	record the time and location the player starts using this NPC	PressX(844|816)(7:27:45)
Releasing an NPC called Angry Cloud	record the time and location the player stops using this NPC	ReleaseX(844|816)(7:27:47)
Calling an NPC called Angel	record the time and location the player starts using this NPC	PressC(888|816)(7:27:53)
Release an NPC called Angel	record the time and location the player stops using this NPC	releaseC(900|816)(7:27:53)

Table 1. Some events and telemetry data.

There are a lot of objects in the game. In fact, everything in the game is an object. Some of them are more important in this study. The first very important object is called Tracker. It isis responsible for recording events such as the character's moves and its interaction with other objects. The object is not visible to the player and does all its activities while the player takes no notice of. Another object in the game that is very important is the telemetry sender. The player has no control on it either and is not also visible. The task of this object is to send the player's behavior data to the server in a compressed format. Fig. 3 shows the workflow of the sender object. The role of this object is so important and this data collection process can be applied to every game.

The process can also be described in a way that different events from telemetry data are sequentially stored on the client side and sent to the server at the end of each level.

During two months, the game was played by 252 different players and the information of 2,173 plays were received, in which 1,196 failures and 977 victories were happened. At the end of the data collection phase, 239,801 records were received and saved, which is a great source of information for game analysis.

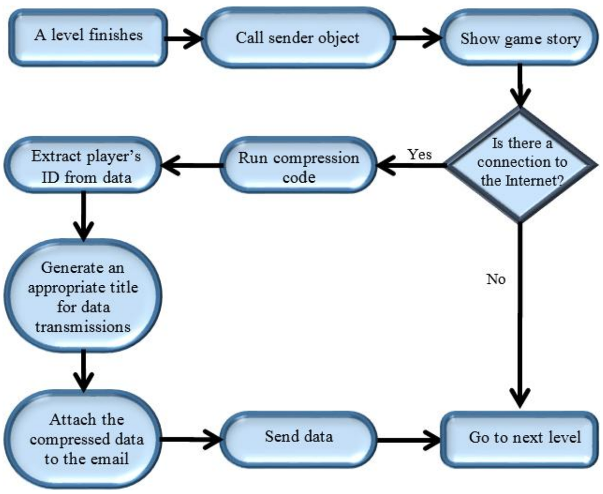

Fig. 3. The workflow of the sender object.

4.2. Visualizer

The analysis tool should help game designers to find the answers of their questions without involving them in the complexity of computer programming. Game designers want to be able to analyze data as a whole as well as case by case. In other words, they would like to have the ability to analyze all data or a group of data, as well as the ability to analyze just a single item.

Regarding the fact that the follow-up policy in this game is "event-based", in order to obtain the playing character track, the occurrence of each event has been taken into account. So the game designers have three distinct options that also can be combined. In the first option, the location in the game where an event occurred, is marked with an image associated with that event. This is shown in Fig. 4. In the second option, the locations of the game where the event happened is marked. The color of marks denotes some meaning. Fig. 5 illustrates this method. In this figure, the arrow keys are in green and the death event is shown in red. This method is called Color Marking.

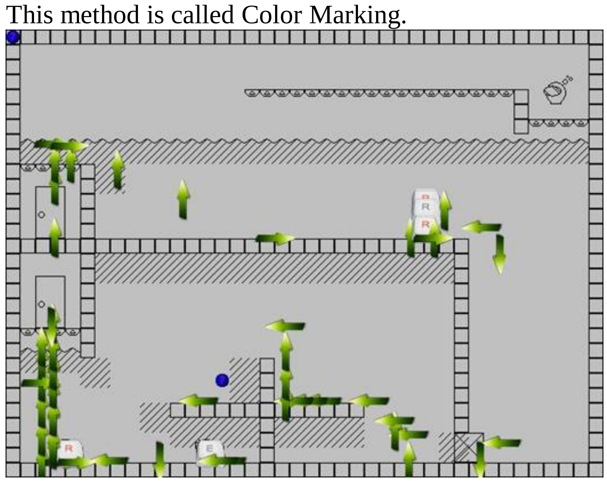

Fig. 4. Tracking the player character by marking events with images such as arrows and the icon of the buttons used.

The sign marking method provides better sight rather than color marking for game designers, because they understand the concept without having to remember the color separation. It is also a good way to examine a single data item. But when the number of telemetry data items increases, the direction of motion cannot be depicted because of the accumulation of marks and their overlap. In the third option, both lines and color separation were utilized to connect the points to each other, so a clear view was achieved. Game events were presented. It is illustrated in Fig. 5.

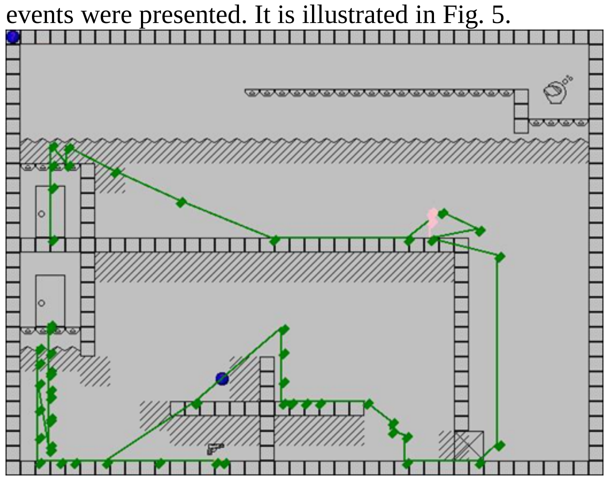

Fig. 5. Tracking a character's motion by marking events and connecting events by drawing a line between the marks.

The heat map is another solution by which we can monitor events. Once the heat map and the map of the levels of the game is compared, if a particular event – such as the death of a playing character– happens more than expected in a region of a level, that region is identified easily. In other words, with this method, a simple but functional view is obtained from the data both as a whole and item by item. Fig. 6 shows 441 times the death event in the 3rd round of the game. The highest concentration of player deaths can be seen at the bottom of the figure, and the bright color of that area also shows that this area is more difficult than the rest of the game.

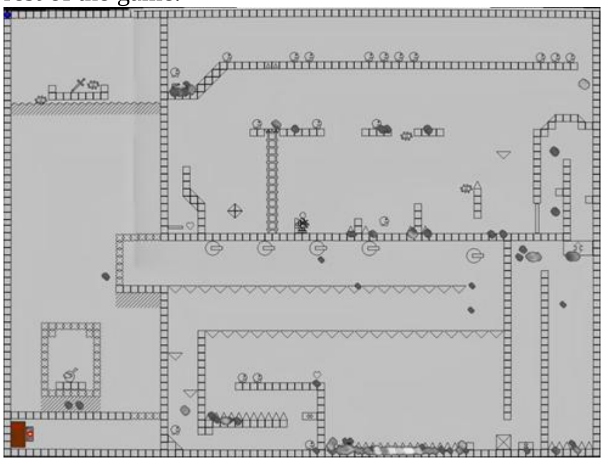

Fig. 6. Heat map from level 3 of the game to show the location of the character death.

The software also allows several events participate in forming a heat map. The heat map of Fig. 7 comprises events including the death, picking up objects and events other than motional ones (directional keys). The areas with a brighter color shows places where the most activity has occurred. This sample shows that how designers can gain a clear understanding of players' behavior by comparing heat maps and game maps.

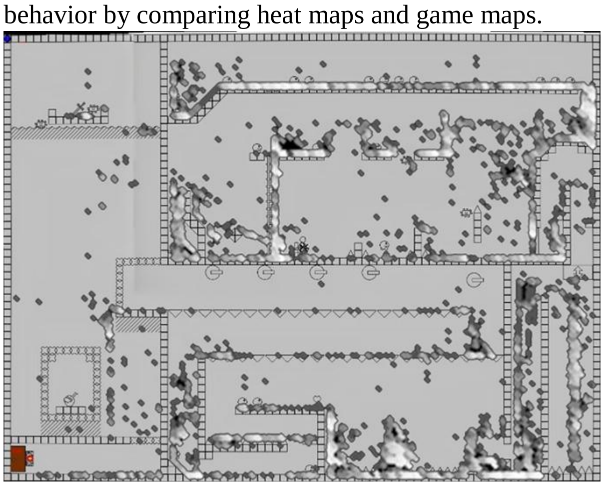

Fig 7. Heat map of level 3 to show the location of non-motional events

Getting quantitative data is crucial because they can be compared. For instance, the percentage of each event in the game data help to understand the behavior of the players in each level. In the visualizer program, the percentage of game events is visible. The percentage is illustrated in a chart that displays the perception of players’ activities at each level of the game, regardless of the place where the event has happened. Fig. 8 shows a chart associated to the level 10 of the game.

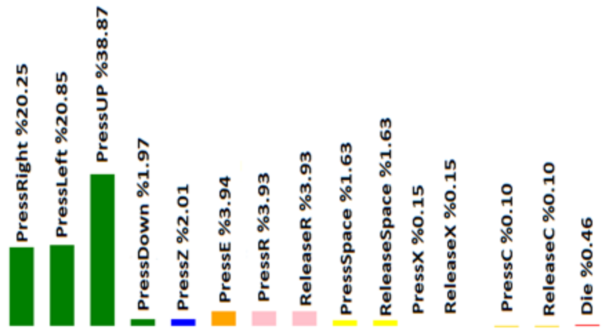

Fig 8. A chart that shows the Percentage of events in the 10th level of the game

One of the advantages offered by the designed software is that it works completely based on the UPB approach to analyze the game. As a result, researchers reach a good perception of players. These analyses are based on collecting telemetry data from players and extracting their behavioral patterns using processes depicted in Fig. 2. They provide complete understanding of the behavior of players in the game environment, the difficulty of each level, and different strategies of players.

The software answered the first question of game designers that wanted to know "In which region of the game, players play more" with three methods: color marking, sign marking, and by means of a heat map. All the three methods visualize the telemetry data to display players’ behavior in each level of the game intuitively. About the second question “Which game configuration is most used by users (players)?” and the fifth one “What percentage of each event occurred in each level?” Considering the chart of the percentage of the events, the two questions are answered both in a visually and quantitatively. To answer the third and fourth questions, “What route does the player choose to finish the game?” and “Which objects are found in the game? For example, was the sword found?”, one can track the character by means of marking the events and connecting the location of events to each other, using a line (Fig. 5). A heat map or a combination of both methods can also be used in this regard because the use of the game objects needs to move the character player on that object. Another case is hitting a particular key that fires an event as well as the former one, and the event can be stored and analyzed utilizing the services described before. For example, in the levels that a sword was provided as a weapon, only 21 percentage of players attempted to pick it up.

4.3. Extracting statistical data from the telemetry data

After preprocessing the telemetry data, extraction of statistical information commenced. This information includes the player's ID, the levels that players started, the time a level took to finish, and most importantly the number of occurrence of each type of event. Then the total and average of playing time of each level were calculated. The total and the average of occurrence of each event were calculated, too. Figs. 9 and 10 show the average and the percentage of occurrence of events, respectively.

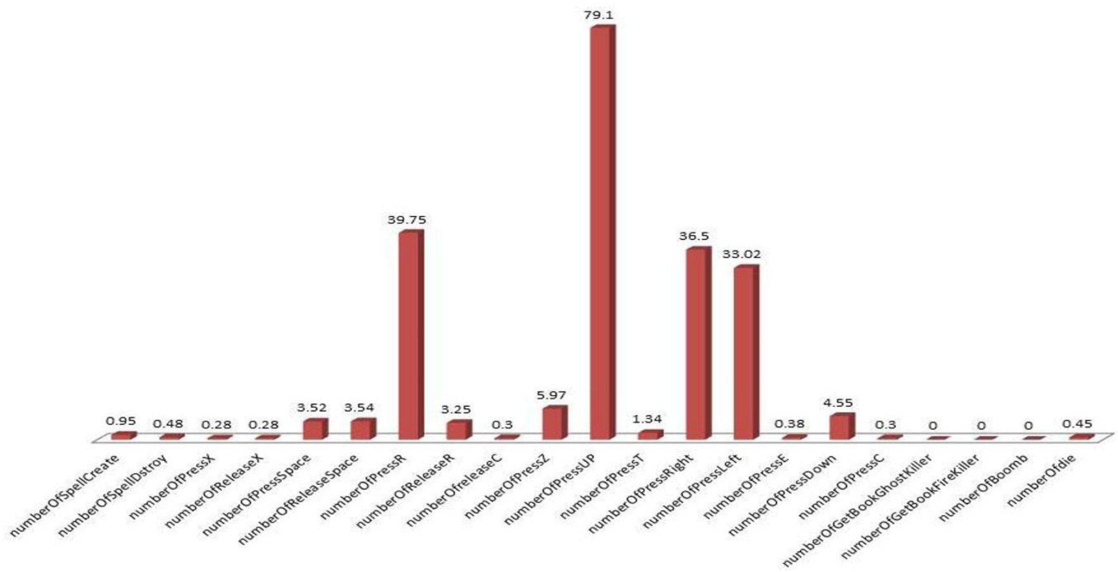

Fig. 9. The average of occurrence of events in the first level

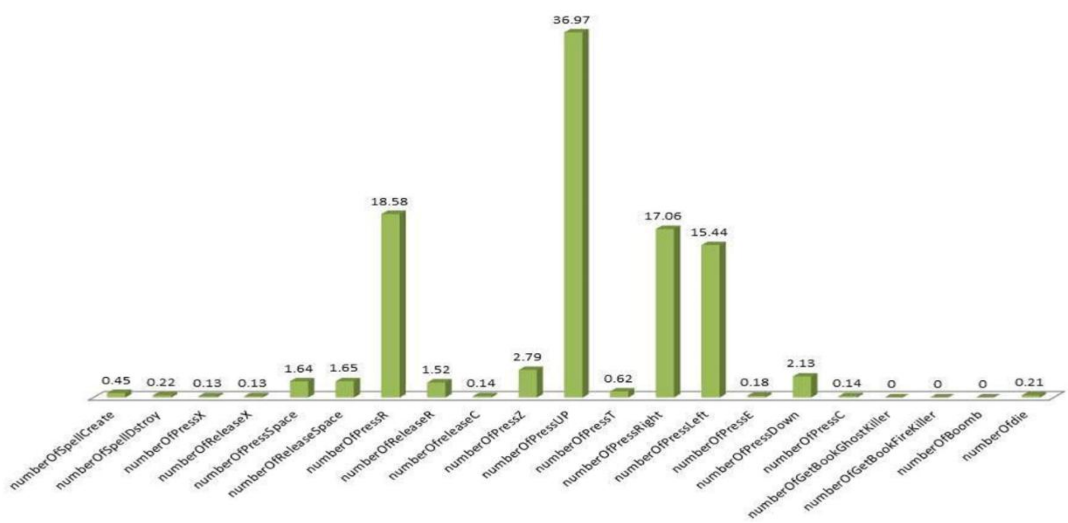

Fig. 10. The percentage of occurrence of events in the first level

Fig. 11, shows the percentage of motional activities of players in all levels. Fig. 12 also shows the non-motional activities. The comparison between motional and non-motional activities is important because of the fact that the higher the non-motional activity rate is, the more players use the game equipment to reach the end of the game. It gives the designer the ability to compare the statistics with their predictions about the amount of motional and non-motional activities required to complete the levels. In the same way, it can be understood that the game has many motional features, and it is more obvious in the training levels, especially levels 1 and 1.5. This is due to two reasons: first, in the training levels, all features and tools (such as weapons and equipment) are not available to the players, and second, the guidelines displayed to the players emphasize on the characteristics of the motional features.

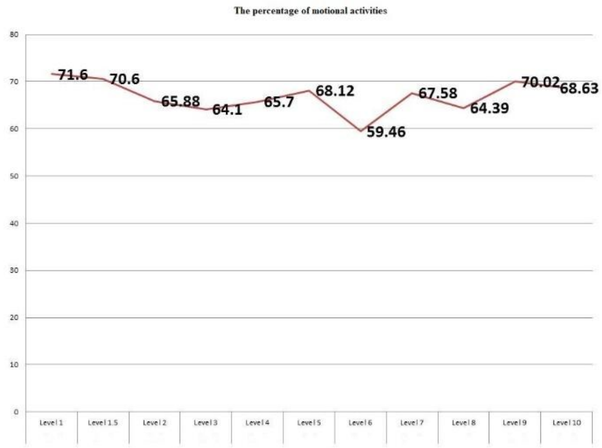

Fig. 11. The percentage of motional activities of players in all levels

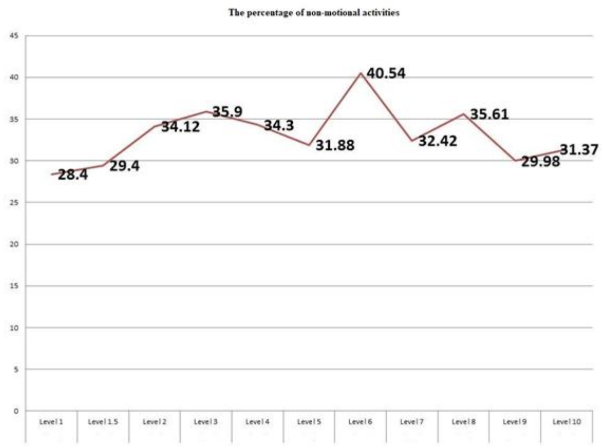

Fig. 12. The percentage of non-motional activities of players in all levels

The percentage death event in Fig. 13 indicates the players' failure rate. It can also be considered as the game difficulty. In the training levels (levels 1 and 1.5), the total loss rate was 52.64%, denoting that the game is very difficult for players to start and train. A lot of players fail at the beginning, so they lose their incentive to continue playing. There are sinusoidal ups and downs in the chart in which the third level of the game with 69.74% failure ratio seems to be the hardest part of the game.

Of course, this can be understood by looking at Fig. 14, because apart from level 1, which is the training level and takes the maximum time, level 3 with the highest time denotes its difficulty rather than the other levels. Considering the player behavior, the difficulty of a game can be found out from the two factors of time and failure rate. Examining the two Figs. 13 and 14, it is concluded that the difficulty of the levels of the game, contrary to the prediction of the game designers, has no upward trend and has ups and downs. This makes players dissatisfied. Therefore, either the levels should be moved or some changes ought to be done on the game regarding the analyses.

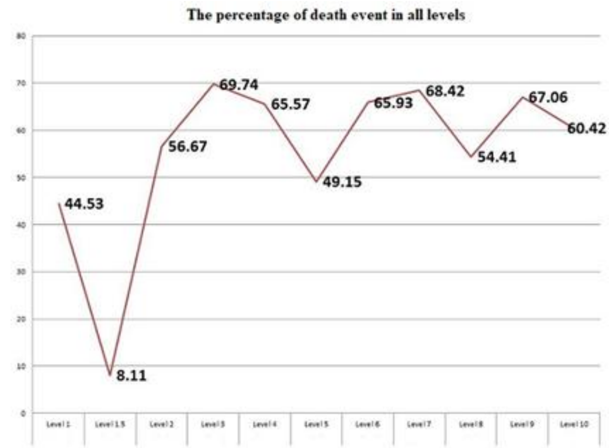

Fig. 13. The percentage of death event

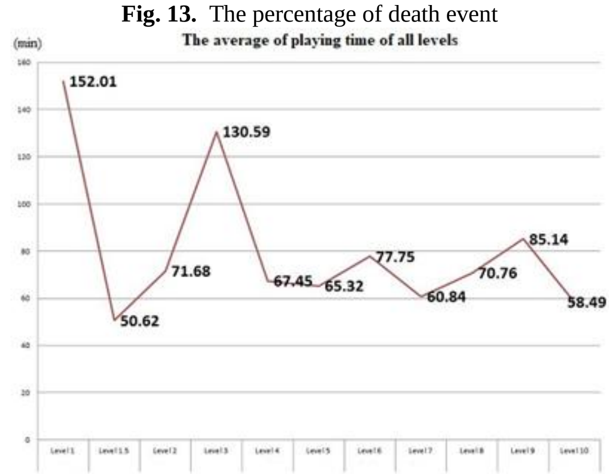

Fig. 14. The average of playing time of all levels

In Fig. 15, interactions with game NPCs are illustrated. The behavior of players coincided with the prediction of the initial design, and the gamers did like the expected pattern. The trend started from level 3 ascendingly and proceeds to level 8. In the last two levels, there are a few NPCs that are too hard to play. According to this design, the interaction with NPCs has been reduced in that levels.

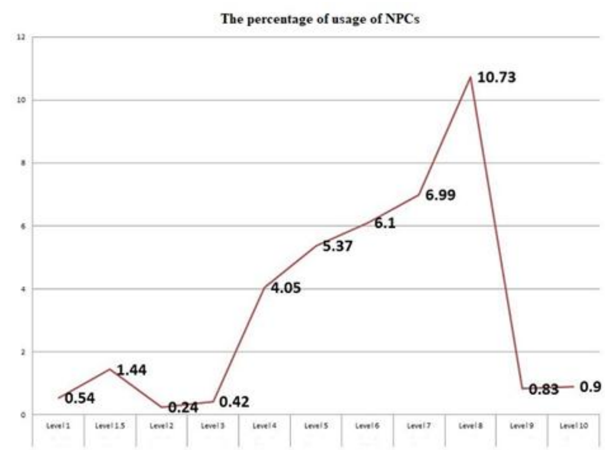

Fig. 15. The interaction with NPCs

4.4. Data Mining

As stated above, using visualization of the telemetry data and statistics of events, one can have a view of the game and that how it is played. However, the decision to apply clustering was taken to go through the various styles of playing the levels of the game. For this purpose, two consecutive clustering algorithms were used. In the first application of clustering, location of events was ignored, but in the second, clustering was fulfilled based on the location of events.

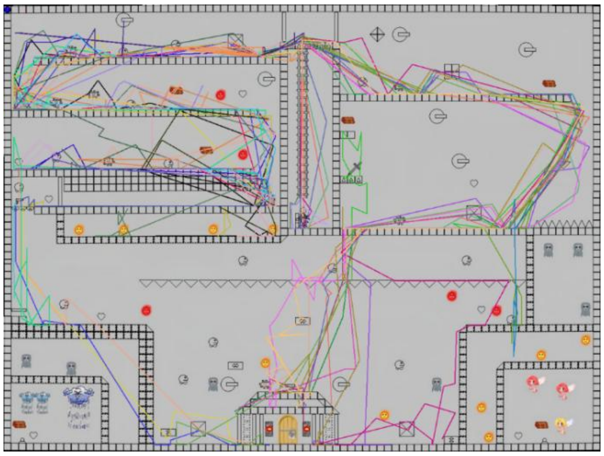

Fig. 16. Visualization of cluster 4 from level 10

1) Non-location-based clustering

The goal of this clustering is to obtain the number of clusters in each level concerning the players' performance. It is extracted from the game statistical data. The number of clusters should be selected in such a way that special samples are not faded due to merging into large clusters. Moreover, the generality of clusters must be preserved. So, two rules were followed as stated below:

- First rule: A cluster with a small number of members must exist.
- Second rule: There must be at least one cluster with an understandable visual pattern.

First rule attempts to reveal unconventional ways of playing the game by players, and the second rule attempts to prevent this cluster shrinkage from continuing. To apply non-location-based clustering, K-means algorithm was selected whose MATLAB implementation was used in this work. For all levels of the game, the clustering of 5 to 15 clusters has been performed and the result is taken according to the rules previously defined. The clusters were visualized by means of a program that was written for this purpose during this study. The program takes the game map and the generated clusters to display the behavior of the players on the game map. Fig. 16 illustrates the program output. It shows the 4th cluster of non-location-based clustering out of 9 clusters in the 10th level of the game. Each color line represents one time playing the level of the game. Regarding the rules considered when clustering, each level was assigned to one of the clusters 5 to 15. Its pseudocode is illustrated in Fig. 21. Then they were visualized, so the appropriate number of clusters for each level was found out. However, because the dataset that was analyzed by clustering, did not contain the location information of events, their visualization did not provide a clear pattern of players' behavior. The players have reached the final levels of the game but it is not known that how they have done.

It should be noted that the most obvious goal of this clustering was to discover the number of appropriate clusters for the data of each level, not to obtain the player's style of play.

2) Location-based clustering

Since the visualization of non-location-based clustering does not provide a clear pattern of player behavior, there is a need for clustering based on the location of events. This is done by dividing the game map into separate areas and counting the events occurred in that area. The question is, any level should be partitioned into how many areas? To answer this question, two rules were followed.

- First rule: There must be at least one cluster with a maximum of two members.
- Second rule: There must be at least one cluster with at least 20% of the total data (for each level of the game).

These two rules extract game specific samples and prevent the production of too many clusters. To determine the optimal number by which each area of any level should be partitioned, all levels were partitioned using seven ways, which can be seen in Table 2.

The number of areas	Column	Row
4	2	2
9	3	3
16	4	4
25	5	5
36	6	6
169	13	13
576	24	24

Table 2. Seven ways of partitioning the game map to determine the optimal number of partitions

Then the events happened at each level in each partition were counted. Moreover, one map was created for each level, so that the events were marked on the associated map with a red mark. The number of each event was counted regarding each partition. This counting was repeated for each level seven times (according to Table 2). Fig. 17 displays 5 different ways of partitioning for level 2 including 2x2 denoted by black line, 3x3 denoted by blue lines, 4x4 denoted by yellow lines and 5x5 denoted by green lines. It is worth mentioning that the analysis for 6x6, 13x13 and 24x24 partitioning were done but because of the fact that showing all of these partitioning ways would reduce the visual perception, they were not displayed in the Fig. 17. The result of this step is shown in Table 3.

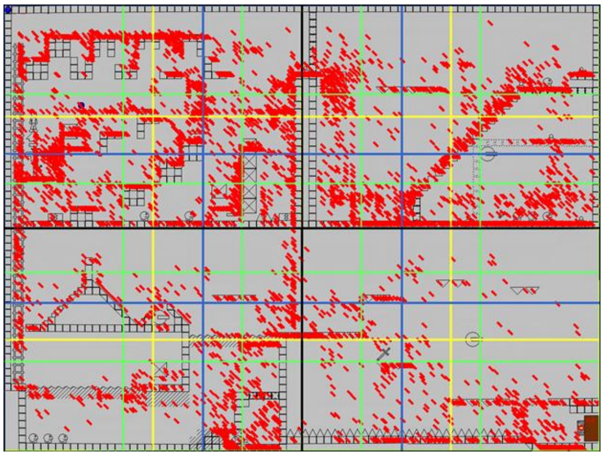

Fig. 17. Color map of level 2 with different ways of partitioning including 2x2 (black lines), 3x3 (blue lines), 4x4 (yellow lines) and 5x5 (green lines)

To perform location-based clustering, like non-location-based clustering, MATLAB implementation of K-means algorithm was used. The number of clusters obtained through the non-location-based clustering (shown in Table 3) was considered for each level, as well. The statistical data containing the number of events happened in the areas obtained from partitioning of each level were fed into the clustering algorithm as the data set. The goal of location-based clustering is determining how many areas are appropriate for each level.

Level	Non-location-based clustering (The number of appropriate clusters)	Location-based clustering (The number of appropriate partitions for each level)	The number of extracted playing styles ~~General playing styles~~
1	8	4*4 = 16	6
1.5	4	4*4 = 16	2
2	8	4*4 = 16	6
3	14	3*3 = 9	10
4	13	3*3 = 9	11
5	10	5*5 = 25	6
6	9	4*4 = 16	9
7	11	5*5 = 25	7
8	12	3*3 = 9	7
9	14	3*3 = 9	11
10	9	6*6 = 36	8

Table 3. Results of Location-based and Non-Location-based Clustering.

Comparing Fig. 16, 18, and 19 shows players who have reached the end of the level while Fig. 18 and 19 reveal two styles of game playing (existing patterns or player strategies) at the final level of the game (Level 10). They were resulted from visualizing the clusters. Each color line represents one time playing of the level. They obviously depict two distinct strategies of players as described below. In Fig. 18, players choose the right path and pick up the dynamite object in order to blast the wall and proceed into the final area of the game (close to the exit door). This group of players enter into the door from the right and the left. In some cases, they explode the ceiling and proceed to the exit door through the hole. In contrast, the second group, as illustrated in Fig. 19, move from the left to the exit door and meet challenges to complete the level.

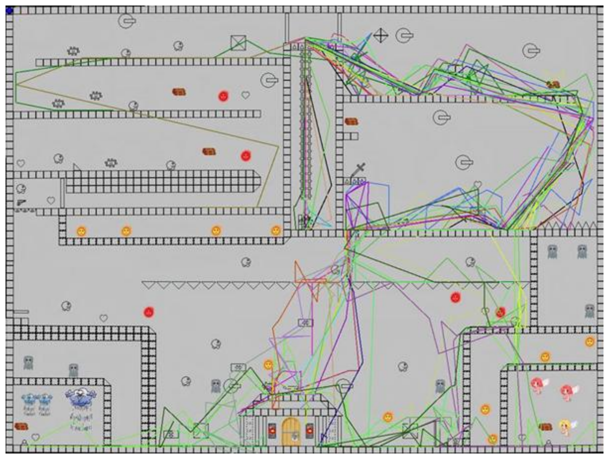

Fig. 18. A view of the location-based clustering of the 10th level of the game

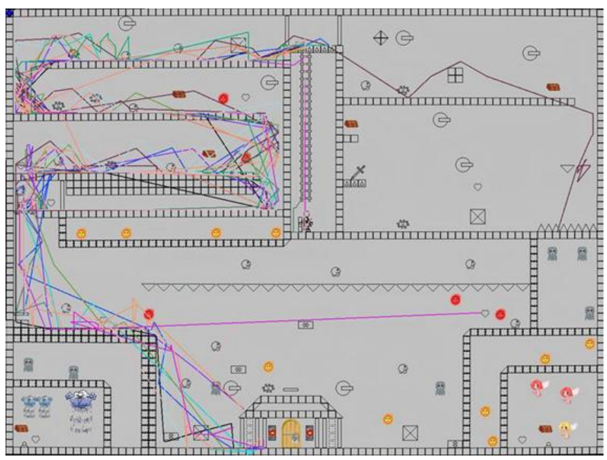

Fig. 19. Another view of the location-based clustering of the 10th level

The results of location-based and non-location-based clustering are shown in Table 3. The number of clusters in non-location-based clustering includes the number of game playing styles in general and specific modes. In fact, the general game playing styles are extracted after the game is depicted which is displayed in the corresponding column. By a specific game playing style, we mean the ways the game is played by some players that differ from all other ones. An example of a specific game playing style is illustrated in Fig. 20. A player in level 8, unlike other players, selected the best path to complete the game, without any error. He/she found some helpful objects and interacted with the NPCs to complete the level quickly. Checking the player ID showed that he/she had completed the game 3 times before and was aware of the paths and available objects based on the previous experience.

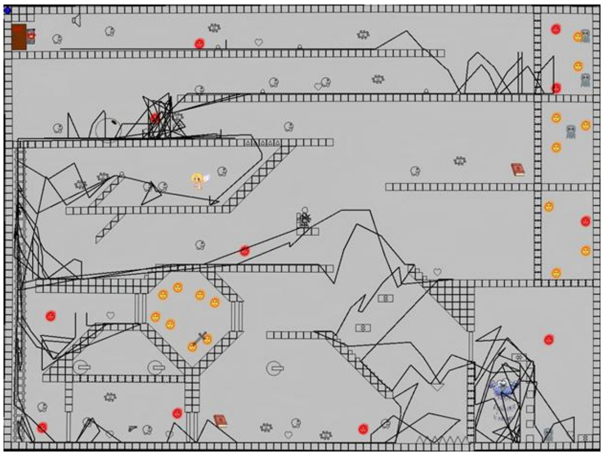

Fig. 20. An example of a specific game playing style in level 8

```
//Non-location-based clustering
MinNCL=5; //Minimum Number of Clusters in Level
MaxNCL=15; //Maximum Number of Clusters in Level
For level= 1 to 10 {
    //Number of Clusters in Level
    For NCL=MinNCL to MaxNCL {
        //Result of k-means
        Result[level,NCL]=k-means(DataOf(Level),NCL)
    }
    SendReportToDeveloper(Level)=…
    VisualizationOfCluster(GameMap(Level),Result[level,All])
    //Appropriate Number Of Clusters For Level
    ANC(Level)=get feedback from developer
}


//Location-based clustering
For level= 1 to 10 {
    For g=1 to ANC(Level) {
        NPL=[4,9,16,25,36,169,576] //number of partition for level
        For i=1 to 7 {
            DataOf(Level)=PartitionOfLevel(DataOf(Level),NPL[i])
            //Result of k-means
            Result[level,i]=k-means(DataOf(Level),i)
        }
    }
    SendReportToDeveloper(Level=…
    VisualizationOfCluster(GameMap(Level),Result[level,All])
    //number of appropriate partitions for each level
    NAP(Level)=get feedback from developer
}
```

Fig. 21. Pseudocode of non-location-based clustering and location-based clustering

5. CONCLUSION

In this paper, UPB approach was proposed for analyzing computer games based on the perception of players' behavior. This is an important issue for the effective development of the games. This approach was practically used to analyze our designed game, called the Treasure Castle. The purpose of this research is to understand how the game is played by the players. These analyses are based on collecting telemetry data of players. In addition to detecting the behavior patterns of players using data mining techniques, it also involves determining different playing styles and strategies. Extracting statistical information of each level, revealed the difficulty of each level to game designers, compared to other levels of the game. To determine the difficulty of any area, a heat map was presented for each level. Heat maps visualize events of one player or a group of players or all collected data at a level so they intuitively provide a clear view for game designers. Generally speaking, the proposed approach provides a complete comprehension of the behavior of players, the game environment, the degree of difficulty of each level of the game and various strategies of players.

Although we have examined the UPB approach on the Treasure Castle game, it can be used for similar games with some tweaks. In other words, we have designed a general framework for strategy visualization and mining of the players' behavior on different games. The different styles of games are because of their different features. To understand how different players act , following a specific approach is useful for game developers. As illustrated in the paper, UPB starts with collecting telemetry data. It includes at least one of the three processes: data mining, statistical analysis, and visualization. Each process, increases game designer's awareness, so they improve their design and produce better games.

In the future, we will work on dynamic leveling. By dynamic leveling it is meant that games adjust themselves to skills of players. For skillful players, games must be harder so it remains attractive and challenging for them. On the other hand, they must become easier for players that has low skills so they do not get disappointed and do not abandon the game. Since the skill level of players can be detected using UPB approach, it can be utilized in dynamic leveling.

REFERENCES

[1] G. Graham, "FUN INC: Why games are the 21st Century's most serious business," Logos, vol. 20, no. 1, pp. 171-173, 2009.

[2] R. L. Mandryk and M. S. Atkins, "A fuzzy physiological approach for continuously modeling emotion during interaction with play technologies," International journal of human-computer studies, vol. 65, no. 4, pp. 329-347, 2007.

[3] L. Nacke and A. Drachen, "Towards a framework of

player experience research," in Proceedings of the second international workshop on evaluating player experience in games at FDG, 2011.
[4] D. Burgos, C. Van Nimwegen, H. Van Oostendorp, and R. Koper, "Game-based learning and the role of feedback. A case study" 2007.
[5] I. R. Hamilton, "Automating formative and summative feedback for individualised assignments," Campus-Wide Information Systems, vol. 26, no. 5, pp. 355-364, 2009.
[6] M. A. G. Maureira, D. P. Janssen, S. Gualeni, M. Westerlaken, and L. Calvi, "Comparing Game User Research Methodologies for the Improvement of Level Design in a 2-D Platformer," in Advances in Computer Entertainment: Springer, 2013, pp. 77-92.
[7] M. S. El-Nasr, H. Desurvire, B. Aghabeigi, and A. Drachen, "Game analytics for game user research, Part 1: A workshop review and case study," IEEE computer graphics and applications, vol. 33, no. 2, pp. 6-11, 2013.
[8] A. Drachen, M. S. El-Nasr, and A. Canossa, "Game Analytics–The Basics: Maximizing the Value of Player Data," 2013.
[9] A. Gagne, M. El-Nasr, and C. Shaw, "A Deeper Look at the use of Telemetry for Analysis of Player Behavior in RTS Games," Entertainment Computing–ICEC 2011, pp. 247-257, 2011.
[10] A. Canossa, A. Drachen, and J. R. M. Sørensen, "Arrrgghh!!!: blending quantitative and qualitative methods to detect player frustration," in Proceedings of the 6th international conference on foundations of digital games, 2011, pp. 61-68.
[11] A. Drachen, A. Canossa, and J. R. M. Sørensen, "Gameplay metrics in game user research: Examples from the trenches," in Game analytics: Springer, 2013, pp. 285-319.
[12] J. Togelius, R. De Nardi, and S. M. Lucas, "Making racing fun through player modeling and track evolution," 2006.
[13] G. Smith, M. Treanor, J. Whitehead, and M. Mateas, "Rhythm-based level generation for 2D platformers," in Proceedings of the 4th International Conference on Foundations of Digital Games, 2009, pp. 175-182.
[14] M. Riedl and N. Sugandh, "Story planning with vignettes: Toward overcoming the content production bottleneck," Interactive Storytelling, pp. 168-179, 2008.
[15] V. Hom and J. Marks, "Automatic design of balanced board games," in Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE), 2007, pp. 25-30.
[16] C. B. Browne, "Automatic generation and evaluation of recombination games," Queensland University of Technology, 2008.
[17] J. Togelius and J. Schmidhuber, "An experiment in automatic game design," in Computational Intelligence and Games, 2008. CIG'08. IEEE Symposium On, 2008, pp. 111-118: IEEE.
[18] H. P. Martínez, A. Jhala, and G. N. Yannakakis, "Analyzing the impact of camera viewpoint on player psychophysiology," in Affective Computing and Intelligent Interaction and Workshops, 2009. ACII 2009. 3rd International Conference on, 2009, pp. 1-6: IEEE.
[19] A. J. M. S. H. Perez and M. G. Yannakakis, "Investigating the Interplay between Camera Viewpoints, Game Information, and Challenge," 2009.
[20] E. Fagerholt and M. Lorentzon, "Beyond the HUD-user interfaces for increased player immersion in FPS games," 2009.
[21] A. Canossa, Play-Persona: Modeling Player Behavior in Computer Games. Danmarks Designskole, 2009.
[22] B. G. Weber and M. Mateas, "A data mining approach to strategy prediction," in Computational Intelligence and Games, 2009. CIG 2009. IEEE Symposium on, 2009, pp. 140-147: IEEE.
[23] K. Isbister and N. Schaffer, Game usability: Advancing the player experience. CRC Press, 2015.
[24] R. J. Pagulayan, K. Keeker, D. Wixon, R. L. Romero, and T. Fuller, User-centered design in games. CRC Press Boca Raton, FL, 2002.
[25] J. S. Dumas, "User-based evaluations," in The human-computer interaction handbook, 2002, pp. 1093-1117: L. Erlbaum Associates Inc.
[26] A. H. Jørgensen, "Marrying HCI/Usability and computer games: a preliminary look," in Proceedings of the third Nordic conference on Human-computer interaction, 2004, pp. 393-396.
[27] C. Swain, "Master metrics: the science behind the art of game design" presented at the NLGD, Utrecht, Holland, 2008.
[28] M. Nørgaard and J. Rau, "User testing in the combat zone," in Proceedings of ACE, Toronto, ON., 2007.
[29] M. C. Medlock, D. Wixon, M. Terrano, R. Romero, and B. Fulton, "Using the RITE method to improve products: A definition and a case study," Usability Professionals Association, vol. 51, 2002.
[30] R. Pagulayan and K. Keeker, "Measuring pleasure and fun: Playtesting," ed: Morgan Kaufmann, 2007.
[31] Bontchev, Boyan, and Olga Georgieva. "Playing style recognition through an adaptive video game." Computers in Human Behavior 82 (2018): 136-147.
[32] Bunian, Sara, et al. "Modeling Individual Differences in Game Behavior using HMM." arXiv preprint arXiv:1804.00245 (2018).
[33] Lora, Diana, et al. "Dynamic Difficulty Adjustment in Tetris." FLAIRS Conference. 2016.
[34] Hendrix, Maurice, et al. "Implementing adaptive game difficulty balancing in serious games." IEEE Transactions on Games (2018).
[35] K. J. Shim et al., "Analyzing human behavior from multiplayer online game logs: A knowledge discovery approach," IEEE Intelligent Systems, vol. 26, no. 1, pp. 85-89, 2011.
[36] L. Mellon, "Applying metrics driven development to MMO costs and risks," Versant Corporation, 2009.
[37] A. Drachen and A. Canossa, "Evaluating motion: Spatial user behaviour in virtual environments," International Journal of Arts and Technology, vol. 4, no. 3, pp. 294-314, 2011.
[38] C. Thompson, "Halo 3: How Microsoft labs invented a new science of play," Wired Magazine, vol. 15, no. 9, pp. 15-09, 2007.