A Game Design Analytic System Based on Data Provenance

Lidson B. Jacob¹, Troy C. Kohwalter¹, Alex Machado², and Esteban W.G. Clua¹

¹ Instituto de Computação, Universidade Federal Fluminense - Niterói - RJ, Brazil
{ljacob, tkohwalter, esteban}@ic.uff.br

² Departamento de Computação - Instituto Federal do Sudeste de Minas Geral
Rio Pomba, MG – Brazil
alexcataguases@hotmail.com

Abstract. This paper presents a game system approach to assist game designers to make decisions and find critical points in their game through data provenance collected from a game. The proposed approach is based on generating graphs from collected data to quickly visualize the game flow. We test and validate our approach with an infinity run genre for mobile devices.

Keywords: Provenance, Game Analytics, Infinity Run.

1 Introduction

Developing a successful game that holds the user's attention for a long time in today's market is a challenging endeavor [3]. Due to this, it is essential to generate income for both games that contains advertisers and those that offer items for sale. [3] also says that thousands titles are published each year, generating a competition for players' time. Making game analysis one of the main resources for ensuring game quality, rate of success, understanding player behavior, and enhancing the player's experience. It also have led to a shift in development design strategies for digital games, bringing data-driven intelligence practices into the fray for help decision making.

A system capable of displaying played and captured game data would support game designers by finding expected patterns as confirmation of such functioning. On the other hand, it could also find unexpected patterns, which indicates critical points where the game designer should analyze for potential problems, like bad difficulty balancing or items values. It is very important that these data are displayed in a simple and clear way for a quicker understanding.

As a proof of concept, we use an infinity run prototype due to the large numbers achieved by this genre games. According to [10] these are in important demand from AppStore. Raising examples as Temple Run by Imangi Studios has more than 170 million downloads and Temple Run 2 has achieved the Top 10 most sold games on iOS. Another example is the Subway Sufers by Kiloo launched on June 2012 and it established itself as the top action title. This genre is characterized by making the player run an endless path. In it, the longer the player stays alive, or the distance achieved, greater are the chances of collecting more items and points.

After collecting the relevant data provenance while running the game, it is possible to generate graph for an easier viewing and understanding of the game, from its beginning to its end. According to [2] we can observe that few systems address the issue of query game data, requiring the user to have knowledge of how to write queries in complex languages. However, the designer doesn't always have such knowledge. These queries are importance to game designer because they may be used to find critical issues. To solve this problem, this paper proposes to develop a system for assisting the game designer in making decisions and finding critical points in their games without necessarily knowing complex languages to query the data.

This paper is organized as follows: Section 2 presents the concept of data provenance, treating about the capture, storage and query of provenance data. Section 3 presents an approach for querying previously collected provenance data. Finally, section 4 presents the conclusions of this work and highlights the contributions and future work.

2 Data Provenance

According to [8][9], data provenance refers to the documentation of an art object, or the documentation process life cycle of digital objects allowing you to see all your history. Due to the historical documentation of an object, it gets some property, allowing the comprehension and the possibility to assess more accurately the importance and application context of the object. Computational methods must be transformed in order to build a qualified provenance, so that it can be retrieved and analyzed [8].

In [1], it is possible to see that the central foundation of provenance is to collect evidence about the time and place of object. Techniques of comparative analysis and expert opinions can also be used for this purpose, but what establishing the provenance is documentation of object. The provenance is the reference to the origin of data and contains identification, responsible, creation date and the processes applied to it.

According to [7], a provenance graph may be an object execution representation in past or present. Thus, provenance can only have information about up to where the player was able to reach. However, this information can be used to help understand why the player could not advance as far as was expected by game designer, showing possible fails at the game design process.

2.1 Provenance Components

According to [4], when referring to computational tasks, may be used a variety of mechanisms for modeling process. Tasks related to games are complex requiring the usage of other tools, data sources, specific libraries, or Web services. We use Web services for the collection and storage of data in order to simulate a real environment where users from different places can play the game.

There are two ways for approaching computational tasks in provenance: prospective and retrospective. Prospective provenance captures the steps (or processes) adopted for generation a product, allowing a record of specifications for the computational tasks. On the other hand, retrospect provenance is used for capturing the steps performed, in addition to information about the environment used for the derivation of a product specific data, such as detailed execution log of a given computational task. It is preferably captured information about the input and output of data, execution user and initial time and final time of execution. However, the retrospective provenance is not dependent on the presence of the prospective, because it allows capturing process execution without the need to know the sequence of activities involved [4]. Assuming in a first case where the user take an item and not deviate from the obstacle. In a second case, the user doesn't deviate from an obstacle and get the item. These are different circumstances because, in the first case, the item may be influencing the user by not deviating he obstacle, while it was different in the second case. Thus the sequence of activities involved is relevant to our work.

According to [6][7], actions can be represented by a series of attributes in context which are related, allowing the creation of a provenance graph containing some information as reason, triggered, responsible, and time. A good reason of using provenance is to produce a graph containing details that can be determined why something occurred. Similar to this idea we propose to collect data and generate graphs leaving the analysis for the game designer.

3 Graphs Representation

A graph representation of the provenance is fundamental for fast and efficient view of the elements and events. From a game design point of view, it is necessary that this representation is associated with game design elements.

The graph representation of data is defined as a “distance by round” as illustrated in Fig. 1 representing in this case the distance that the user advanced during each round. By observing case 1, the users in the first rounds reached short distances. This may be a consequence of the first contact with the game, where obstacles were a novelty, thus quickly losing the game. Another factor that could have influenced is the unfamiliarity with the character controller. After getting used with the game, users reached greater distances. However, it stabilizes at 2000. This may be due to the increased difficulty as the distance increases.

These analyses can help the game designer in the following circumstances: First, the users spend many rounds to known the game. This can cause abandonment because it's might be very difficult for the player and frustrating. Developing a tutorial at beginning of the game would be a good solution to the problem. Second, the maximum distance that users were able to reach might not be the distance expected by game designer. For example, the coins distribution and special items are only available after 2000 distance, but users hardly reach this distance. The designer will have to rearrange the items or change the game difficulty.

Fig. 1. Distance x Rounds

A large permanency in a distance as the Case 2 from Fig. 1 can demonstrate a difficult obstacle where users couldn't surpass. It may be necessary to remove or simplify the obstacle so users don't get frustrated and stop playing the game. This type of graph can be made to analyze the experience of a single or several users.

Another example is Case 1 from Fig. 2, which represents the number of acquired coins over the total advanced distance. This information can be used to reach conclusions about a uniform distribution for the coins. Perhaps the game designer wanted to benefit users who achieved greater distances, while also motivating to go farther. This is one of the main goals of the Infinity Run games. To represent a distribution in this context would be the case 2 of Fig. 2. A graph for special items would also be similar to the graph for coins.

Fig. 2. Coins x Distance

Graphs can also be combined for analysis, as shown by Fig. 3, to find features that influenced other features. For example, the user loses when acquires a special item and pass only when don't acquire it.

Fig. 3. Combined Analysis

4 Conclusion

This paper presents an approach for a system that helps the game designer by using visual representations of the data to facilitate understanding of the game flow and to make decisions faster and more punctual. This is very important as (Gaspar et al., 2012), where this type of system should provide the designer with facilities as features for viewing, querying and analyzing the data. It is also important for (Costa, et al., 2010), which highlights as an important issue to query these data to be one of the most important things for scientists that analyze the collected data from experiments. Experiments for game designer are your game and, therefore, it's necessary the existence of a system query such data.

Currently this work doesn't have gone through the implementation phase. The next phase consists of the implementation of this approach applied to a game and verify by experiments if it helps game designers during analysis. And to know if the proposed analyses approach is easier and quicker than existing ones.

References

1. Almeida, F.N.: Descrição da Proveniência de Dados para Extração de Conhecimento em Sistemas de Informação de Hemoterapia. Interunidades em Bioinformática (2012)
2. Costa, P.V.C., Braganholo, V.: Uma nova abordagem para consulta a dados de proveniência (2010)

3. Drachen, A.: GameAnalytics. [Online] GA, 02 de 11 de 2012. [Citado em: 22 de 04 de 2013.], <http://blog.gameanalytics.com/blog/announcing-game-analytics-maximizing-the-value-of-player-dat.html>
4. Freire, J., et al.: Provenance for Computational Tasks: A Survey. *Computing in Science and Engineering*, pp. 11–21 (2008)
5. Gaspar, W.A., Braga, R.M., Campos, F.A.: SWfPS: Um Sistema de Proveniência de Dados e Processos no Domínio de Workflows Científicos (2012)
6. Kohwalter, T.C., Clua, E.W.G., Murta, L.G.P.: Provenance in Games. *SBGames* (2012)
7. Moreau, L.: The Open Provenance Model core specification(v1.1). *Future Generation Computer Systems*, 743–756 (2011)
8. Moreau, L., et al.: The open provenance model(v1.00). Technical report, University of Southampton (2007a)
9. Moreau, L., et al.: The Provenance of electronic data. *Communications of the ACM*, 4:52-4:58 (2007a)
10. Todor, A.: Monetizing an Infinite Runner. *GameAnalytics*. [Online] GameAnalytics, 22 de 02 de 2013. [Citado em: 22 de 04 de 2013.], <http://blog.gameanalytics.com/blog/monetizing-an-infinite-runner-guest-post.html>