Visualizing and Understanding Players' Behavior in Video Games: Discovering Patterns and Supporting Aggregation and Comparison

Dinara Moura
Simon Fraser University-SIAT
Surrey, BC, Canada V3T 0A3
dinara@sfu.ca

Magy Seif el-Nasr
Simon Fraser University-SIAT
Surrey, BC, Canada V3T 0A3
magy@sfu.ca

Christopher D. Shaw
Simon Fraser University-SIAT
Surrey, BC, Canada V3T 0A3
shaw@sfu.ca

ABSTRACT
As video games become more popular, there is an urge for procedures that can support the analysis and understanding of players' behaviors within game environments. Such data would inform game and level designers of game design issues that should be fixed or improved upon. By logging user-initiated events in video games, analysts have exhaustive information regarding players' actions within games. However, visualizing such data is a challenging task due to the amount of data one has to deal with; the necessity of a deep understanding of the game and players' possible actions within the game plus a deep understanding of questions one wants to answer; the computation that has to be done on the data; and the limitations and/or complexities of current analysis tools. In this paper, we present a new visualization system that allows analysts to build visualization and interact with telemetry data, to identify patterns and identify game design issues efficiently. Besides the system itself, we propose a new approach to visualize players' behavior that has not been explored so far. For example, instead of using heat maps to visualize a single metric (e.g. deaths), our system allows analysts to superimpose and visualize a series of actions players take in the game. This is especially important when one should understand cause and effect within the game. We present examples of the visualizations using an RPG game, Dragon Age Origins (BioWare/EA, 2009). It should be noted that the system is currently under development and testing with analysts working at BioWare.

General Terms
Visualization, Telemetry Data

Keywords
Players' behavior, Visualizations, Visualization System

1 INTRODUCTION
As games gain more popularity (Jenkins, 2006; The Entertainment Software Association, 2010; Turkle, 2003), designers start to feel the pressing need to adapt their designs to cater to a wide diverse market. Thus, understanding how different players interact and progress within video games is obligatory. For example, for role-playing games (RPGs), designers need to verify if design goals are being achieved, including if players appropriately navigate through the environment and if they find and collect key items, just to mention a few. As a result, game companies have invested time, effort and money to collect such information and test their designs through different methods (Isbister and Schaffer, 2008).

Recently, there has been a move towards collecting player activities through logging user actions. This process is called instrumentation (Kim et al., 2008). Its outcome is telemetry data – a quantitative record of player-game interactions.

Many companies use telemetry to uncover system crashes and bugs (Zoeller, 2010). Also, telemetry data allows analysts to assess virtual economy (Hopkins, 2010) and catch cheaters (Kennerly, 2003). While those applications are useful, using telemetry to analyze player behavior, uncovering design issues, has received little attention. Most current work uses visualization methods such as heatmaps and bar charts to represent actions aggregated over many players and playthroughs (e.g., Drachen and Canossa (2009). Such methods are limited in analyzing temporal progression of players within a game, which is crucial for games like RPGs. Visualizing and making sense of telemetry data is challenging due to the large amount of information involved and the limitations of current analysis tools. More generic analysis tools may be used but are too complex and expensive for design studios. Therefore, more work is needed to develop specialized visualization tools, allowing analysts and designers to make sense of telemetry data more efficiently (Romero, 2008).

In this paper we address this problem by proposing a novel interactive visualization system specialized for video games that involve navigation, such as action/adventure games and RPGs. It allows analysts to make sense of telemetry data through visualization and comparison between different player types. By interacting with our system, analysts are able to visualize player actions by cluster or aggregated over multiple clusters. Analysts can also filter the data by specifying time windows. Additionally, our system allows analysts to superimpose data categories (which are specific actions taken by the players such as collecting items or opening the menu) or filter such data out, which enables them to focus their investigations. The result is an intuitive, clean visualization that allow analysts to "tell a story" based on the data and generate new hypotheses to be tested.

This paper is organized as follows. First, we discuss related research. Second, we explain our interactive visualization system. The last section presents our conclusions and future work.

2 PREVIOUS WORK
Telemetry provides detailed, quantitative record of thousands of players' behavioral data. Understanding players' behaviors using telemetry is an open problem due to the considerable amount of information and data involved (Romero, 2008). To address this

problem, several visualization and data mining methods can be used.

Chittaro et al. (2006) proposed VU-Flow – a visualization system showing individual or aggregated users' paths through lines. Similarly, Coulton et al. (2008) proposed a visualization tool to visualize players' navigation in Pervasive Location-Based Games, where player position is superimposed on a 2D plane with time represented as a 3rd axis. This allows one to understand players' movement through space. However, it is not scalable.

Additionally, Drachen and Canossa (2009) proposed the use of heatmaps to represent specific events, such as death and spawn points, in *Tomb Raider: Underworld*. Such variables are aggregated and shown one at a time within a map. This approach was also adopted by Romero (2008) and other analysts within the industry. While heatmaps are appropriate to draw analysts' attention to hot spots within a map, they alone cannot be used to make sense of the context behind specific actions or allow exploration of cause and effect.

To address these issues, Hoobler et al. (2004) developed Lithium – "a system that enables high-level analysis of competition taking place in multi-player games (page, 163)." It was developed based on a multiplayer game. Analysts can visualize occupancy of the map over time, players' orientation, munitions usage, and player classes.

Microsoft proposed the TRUE system, which supports the triangulation and visualization of telemetry data with other types of data (Kim et al., 2008; Romero, 2008). While such triangulation is important for analysts to understand the context of action, the visualization system used in TRUE is similar and faces many of the disadvantages of heatmaps and simple progression lines discussed above (Romero, 2008).

3 INTERACTIVE VISUALIZATION SYSTEM

In this section we present the proposed interactive visualization system. This system allows analysts to aggregate, filter, visualize, and compare players' behavior based on telemetry data. Figure 1 shows the basic interface.

Analysts can visualize data from all players within different time ranges. They can also visualize specific clusters (players with similar play styles), or a single player playthough in case of outliers, in a specific time range (top buttons in Figure 1). These clusters are important as they provide a level of abstraction over player behavior (typically formed using a clustering algorithm (Witten, Frank, & Hall, 2011), and thus provides a scalable solution.

Analysts can also specify time ranges through the slider (dark gray bar bellow the top buttons in Figure 1). Analysts can pan on that time line bar and select a specific map they want to analyze. Player actions are categorized and displayed on the left menus (top and bottom). These event categories can be turned on and off. In Figure 1, for example, the analyst selected the first hour of gameplay (see top right button). Thus, only the 5 maps related to that time range are displayed. Analysts can then evaluate if this information is in tune with the designers' intentions; e.g., did designers expect players to finish 5 maps in the first hour of gameplay?

After selecting the map analysts want to work with, they can choose the events they want displayed by using the top menu on the left (Figure 1). Event categories are color coded, making it visually easy to connect between the event category chosen and its representation in the visualization. More detailed information is provided by selecting event categories from the bottom left menu.

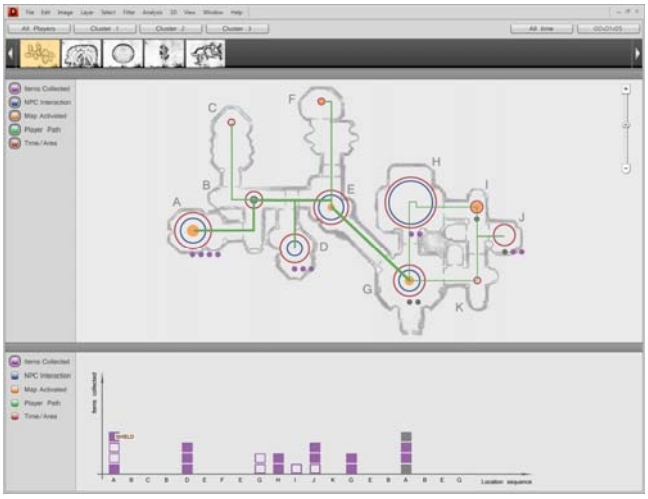

Figure 1: *Interface of the proposed visualization system. Analysts have selected the first map from the time line on the top. All event categories are selected and visualized. The bottom graph displays 'Items Collected' in the first map.*

Analysts can open new windows with new data (Figure 2); thus they can compare players' behaviors across different maps. For example, analysts can open different clusters or different maps in new windows. In addition, by right clicking on the bottom left menu, analysts can open graphs that present specific measures (see new window on the right).

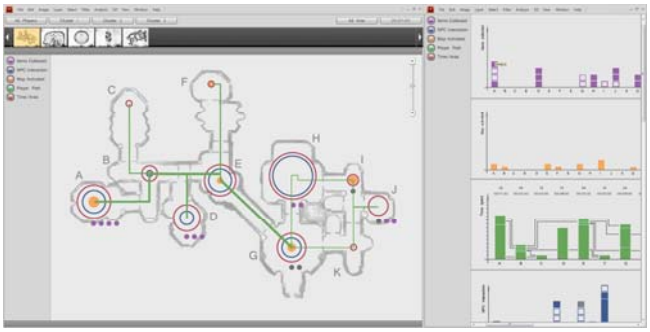

Figure 2: *Analysts can minimize the bottom part of the interface, expanding the map. They can open graphs (or maps) in new windows.*

4 VISUALIZING TELEMETRY DATA

Video games are complex systems that include a large number of actions to be tracked. The system discussed above accommodates visualization of several actions at the same time. In this section, due to limitations of space, we chose to focus on only 5 event categories; but, the visualization system allows the inclusion of other event categories.

4.1 Visualizing time spent in each area

Knowing how much time players spent playing the game is crucial. In our visualization, we detail how much time players spend in each area within a level. We named each room with a different letter. 'Area A' is the starting point. A circle represents the total amount of time players spent in one area of the level. The more time players stay in one area, the bigger the circle (Figure 3).

The size of the circle is calculated by varying the area, not the radius.

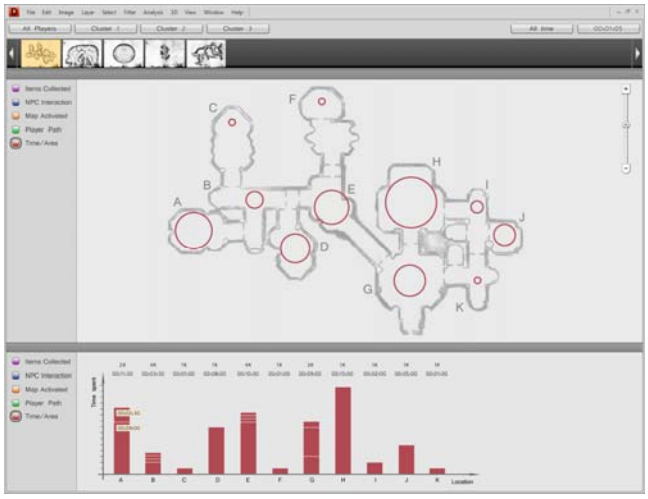

Figure 3: *An example visualization where players spent 11:30 minutes within 'area A' in total. Players visited 'area A' 2 times and spent 9:00 minutes and 2:30 minutes in the first and second visits respectively.*

For more precise details about the time players spent in each area, analysts can open up the graph related to this event category (Figure 3, bottom). In the graph, each room is represented by a different column. The graph shows time spent (Y axis) by location (X axis). Analysts can see how much time players spent in a specific room, and how many times they visited a room. Every time players visit a room a new rectangle is drawn in the respective column. The gap between the rectangles shows that a new visit occurred. For example, in Figure 3, players visited 'Area E' 4 times but spent very little time in the last 3 visits. Analysts can mouse over each rectangle to see the exact amount of time spent on average per each visit.

4.2 Visualizing interaction with characters

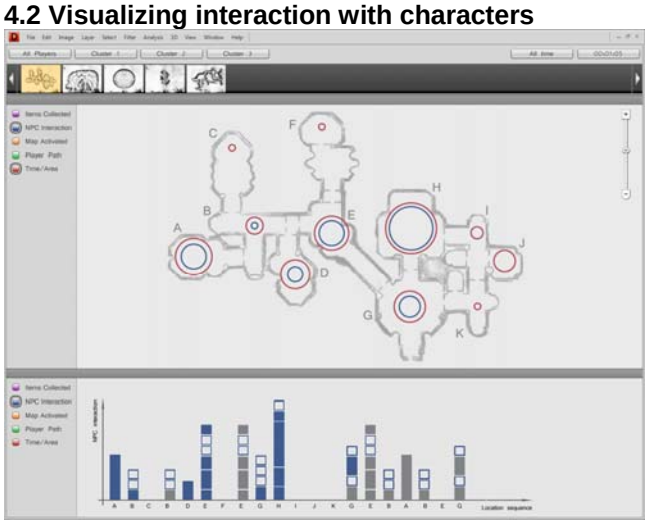

Figure 4: *Both, time players spent per area and time players spent talking to NPCs are visualized. On the bottom, analysts can visualize how many NPCs are within each room and how many on average were talked to (blue rectangles).*

Knowing what players do in each area is also important. For example, for an RPG, interacting with non-players characters (NPCs) is important because they are a source of information and a key aspect in the narrative. Game designers need to know whether players are interacting with NPCs and for how long. Figure 4 shows such relation. The amount of time players spent on average talking to NPCs per area is represented by blue circles; again the amount of time is represented by the area of the circle.

The graph on the bottom (Figure 4) shows the locations players visited (X axis) by representations of NPC-player interactions (Y axis). NPCs that players interact with are represented by blue rectangles. NPCs that players do not interact with are represented by white rectangles. The graph also shows the sequence of areas visited. The player first went from area A to B, then C, then B to D, and so on. Thus, analysts can track when players talked with any specific character. Gray rectangles represent NPCs that players interacted with in a previous visit to a specific area. The size of the rectangle is related to the amount of time players spent talking to an NPC. If analysts mouse over a rectangle in the graph, they will see how much time players spent talking to that NPC on average, and if that NPC was a key NPC or not.

4.3 Visualizing where maps are activated

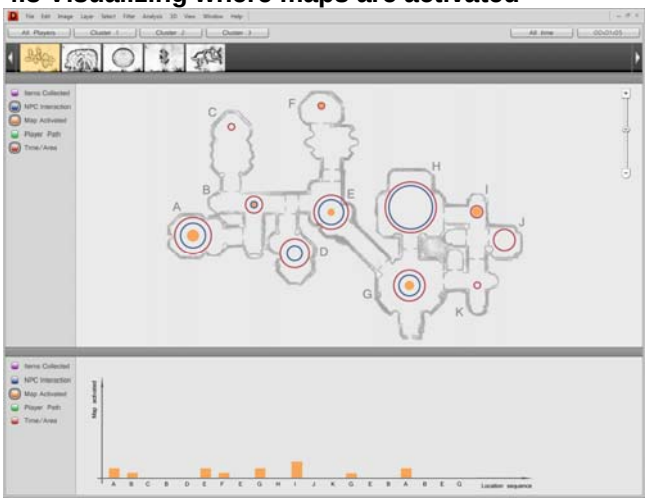

Figure 5: *Analyzing when and where players interacted with the map*

It is important to know where players open the map and how much time they spend with the map active. Based on this information designers can hypothesize (and then test) if players are getting lost in a specific level or not. The yellow dots on top of the map represent locations and relative amount of time players spent checking the map (Figure 5). Based on this visualization, analysts can see the most problematic areas in the level. Also, by analyzing the graph on the bottom of the screen, analysts can verify when and where players started checking the map and for how long.

4.4 Visualizing items collected

There are key items in the game that should not be missed by players. Figure 6 shows how the system visualizes such data. Each item collected in the game is represented by a purple dot on the map. Dots are placed according to the area the items are located in the level. Gray dots represent items that were not collected by the player. The graph on the bottom of Figure 6 works similarly to the graph of Figure 4.

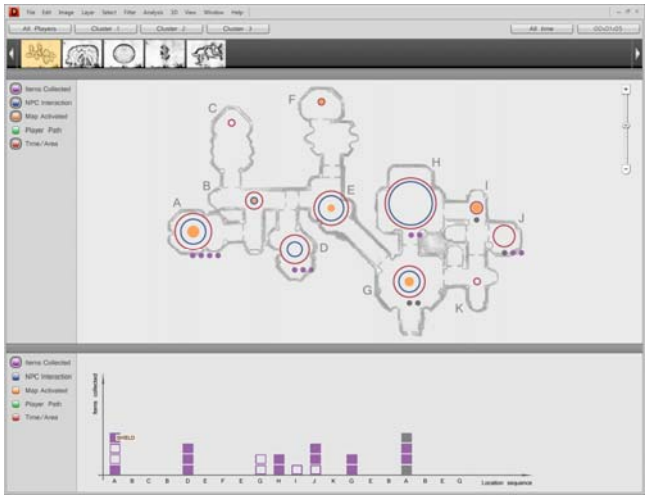

Figure 6: *Analysis of items collected*

4.5 Visualizing players' paths

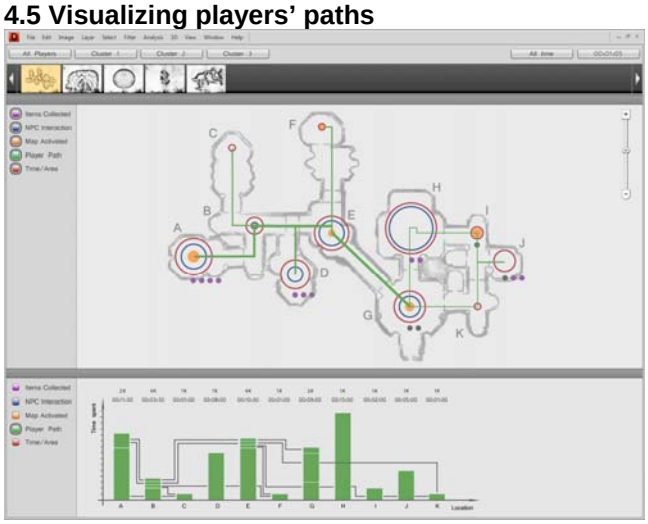

Figure 7: *Analysis of player paths and movement*

It is important to visualize and understand how players navigate through the level. Paths followed by players are represented by green lines on top of the map (Figure 7). The more players navigate through an area, the thicker the line representing that path will become.

The bottom graph shows locations (X axis) and the time spent in each location (Y axis), similar to the graph seen in Figure 3. Analysts can see how much time players spent in each area and how many times they visited each area. This bottom graph represents players' paths in an abstract way. Each area is represented by a column. Each visit to an area is represented by a rectangle that varies in size based on the amount of time players spent in each visit. Every time a player returns to an area, a new rectangle is placed on top of the first one. Finally, arrows connect the rectangles based on the sequence of locations the player visited.

5 COMPARING VISUALIZATIONS

To exemplify the utility of our system, in this section we explain the visualizations generated by 5 different play sessions producing 3 different clusters where the average times for these clusters were 2:21 (left, n=1), 7:34 (center, n=1), and 6:00 (right, n=3) minutes respectively (Figure 8).

In this first part of the game, areas C and F were locked. There were 11 items to be collected and 17 NPCs in the unlocked areas. Players are able to explore the environment, talk to NPCs, open the map, and open the menu/inventory. There was no combat.

In the first visualization (left), the player navigated fast. He did not collect any item and talked to only 2 NPCs. He opened the map every time he entered in a different room until he saw the exit in 'Area G'. In the second visualization (center), the player took time exploring the game. He collected 2 items, visited 'Area D' (missed by the first player), opened the map, and opened the inventory (light blue). Finally, in the third visualization (right), players were more interested in talking to NPCs. Note that they talked to NPCs in every area they visited.

Comparing the first 2 visualizations we notice that, as players explore, the visualization becomes busier. The first player left area A as soon as he finished the first dialogue. Note that the red circle (total time in the area) and the dark blue circle (time spent in dialogues) have almost the same area. In contrast, player 2 spent more time in 'Area A' because after finishing the dialogue he opened the inventory and explored the environment collecting items. In the third visualization, players also did not leave as soon as the dialogue was over, which gave them time to find one item.

The data showed here comes from only 5 participants so we cannot draw conclusions at this point. However, using visualizations like these, games designers can verify whether players are performing as expected and then test and/or improve the game accordingly. For example, if players are going from one room to the next too fast, the designer can lock certain areas until specific items are collected. If players are taking too long exploring the inventory, designers can investigate the user interface of the inventory. That is, based on telemetry data, designers can construct hypotheses, test and improve their games.

The visualizations presented here show sequences of actions, enabling analysts to understand the context in which certain events happened. For example, analysts will be able to verify if players got stuck in an area because they missed an NPC and if such event is related to players quitting the game. Understanding cause-effect is crucial to improve games.

Figure 8: *Comparing actions within Dragon Age Origins*

6 CONCLUSIONS

We have presented an interactive visualization system for visualizing players' behaviors within RPG and Action/Adventure games where navigation, collection, and talking to NPCs are important. Our system has several novel contributions: (a) it allows analysts to interact with large data and see temporal progression and cause and effect of player actions, specifically in terms of navigation and actions. This is not possible with regular heatmaps or tableau type visualizations. (b) It visualizes and enables comparisons of behaviors across different play styles, allowing analysts to see how their game fairs to different tastes and styles. (c) We believe that it presents a clean and analyst friendly interface to large data, making a daunting task simpler than would have otherwise been given the current tools and systems available. This claim needs to be evaluated and tested with designers and analysts which is a goal we are currently working towards. Specifically, we are currently collaborating with Bioware to evaluate this system with DAO telemetry. Future publications will be concerned with the value of the system and its use in a game development pipeline.

ACKNOWLEDGMENT

The authors would like to extend their sincere thanks to GRAND-NCE for making this work possible. Also, thank you for the reviewers for their valuable comments.

REFERENCES

Chittaro, L., Ranon, R., & Ieronutti, L. (2006). VU-Flow: A Visualization Tool for Analyzing Navigation in Virtual Environments. *IEEE Transactions on Visualization and Computer Graphics*, *12*(6), 1475 - 1485.

Coulton, P., Bamford, W., Cheverst, Pk., & Rashid, O. (2008). 3D Space-Time Visualization of Player Behaviour in Pervasive Location-Based Games. *International Journal of Computer Games Technology*.

Drachen, A., & Canossa, A. (2009). Analyzing spatial user behavior in computer games using geographic information systems. *Proceedings of the 13th International MindTrek Conference: Everyday Life in the Ubiquitous Era on - MindTrek '09* (p. 182). Presented at the the 13th International MindTrek Conference: Everyday Life in the Ubiquitous Era, Tampere, Finland.

Hoobler, N., Humphreys, G., & Agrawala, M. (2004). Visualizing competitive behaviors in multi-user virtual environments. *IEEE Visualization 2004* (pp. 163-170). Presented at the IEEE Visualization 2004, Austin, TX, USA.

Hopkins, C. (2010, April 28). Second Life Economy At Record High. *Read Write Web*. Retrieved from http://www.readwriteweb.com/archives/second_life_econom_at_record_high.php

Isbister, K., & Schaffer, N. (2008). *Game Usability: Advancing the Player Experience*. Morgan Kaufmann.

Jenkins, H. (2006). *Convergence Culture: Where Old and New Media Collide* (illustrated edition.). NYU Press.

Kennerly, D. (2003). Better Game Design Through Data Mining. *Gamasutra*. Retrieved December 14, 2010, from http://www.gamasutra.com/view/feature/2816/better_game_design_through_data_.php

Kim, J. H., Gunn, D. V., Schuh, E., Phillips, B., Pagulayan, R. J., & Wixon, D. (2008). Tracking real-time user experience (TRUE): a comprehensive instrumentation solution for complex systems. *Proceeding of the twenty-sixth annual CHI conference on Human factors in computing systems - CHI '08* (p. 443). Presented at the Proceeding of the twenty-sixth annual CHI conference, Florence, Italy.

Romero, R. (2008). *Successful Instrumentation*. Presented at the Game Developers Conference.

The Entertainment Software Association. (2010). The Transformation of the Video Game Industry. Retrieved September 22, 2010, from http://www.theesa.com/gamesindailylife/transformation.asp

Turkle, S. (2003). Video games and computer holding power. *The New Media Reader*. The MIT Press.

Witten, I. H., Frank, E., & Hall, M. A. (2011). *Data Mining: Practical Machine Learning Tools and Techniques, Third Edition* (3rd ed.). Morgan Kaufmann.

Zoeller, G. (2010). *Development Telemetry in Video Games Projects*. Presented at the Game Developers Conference 2010.