Data-Driven Web Entertainment: The Data Collection and Analysis Practices of Fantasy Sports Players

Gabriel S. Dzodom
Department of Computer Science
Texas A&M University
College Station, TX 77843-3112
gabriel.dzodom@tamu.edu

Frank M. Shipman
Department of Computer Science
Texas A&M University
College Station, TX 77843-3112
shipman@cse.tamu.edu

ABSTRACT
More and more data is available on the Internet. This data can provide the context for new forms of entertainment and edutainment. Here we study fantasy sports as an example of how this happens and how users bound their data-oriented activities. Fantasy sports are enjoyed by many tens of millions of players around the world. But how engaged are these players and how do they manage the information-rich nature of the game? To start to answer these questions, we are investigating players' fantasy sports practices. How much time do they spend playing? What information do they use to make decisions? How do they interact with other players? In this paper, we report on 160 responses to a questionnaire of US-based Amazon Mechanical Turk workers who are self-identified fantasy sports players.

Categories and Subject Descriptors
K.8.0 [**Personal Computing**]: General– *games.*

Author Keywords
Fantasy sports; gaming practices.

1. INTRODUCTION
The Web is not just a provider of content aimed at human consumption. Whether provided by organizations intending to add their content to the "Web of Data" or by institutions viewing the Web as a delivery mechanism for the data they collect, more and more data streams are being captured and made available in near real time. These datasets enable new forms of engagement and entertainment. We are interested in encouraging engagement with data through new on-line activity. To gain insight into how this can be accomplished, we explore fantasy sports as a success model for motivating people to analyze large data sets.

Fantasy sports are interactive games where players act as if they are the manager of a team of athletes. The player competes against other fantasy team managers. Similar to real team management, players decide which athletes start, substitute for injured athletes, and trade athletes with other fantasy players. Each player's score is based on statistics about the performance of their athletes, generated by a single athlete or a team during a real sport event. Up-to-date knowledge of the sport is valuable for these games [2].

WebSci'14, June 23–26, 2014, Bloomington, IN, USA.
ACM 978-1-4503-2622-3/14/06.
http://dx.doi.org/10.1145/2615569.2615649

Fantasy Sports are very successful. According to the fantasy sports trade association (FSTA), 33.5 million people played fantasy sports in the United States in 2013 and this number is expected to rise in subsequent years[1].

While there has been a variety of research into the demographics and motivations of fantasy sports players, our focus is on how players collect and use information to make decisions. To explore this topic, we asked self-identified fantasy sports players about their gameplay and experiences. The five main research questions we were trying to answer:

Q1: *How do players select athletes for their teams?* Do players collect and analyze data or rely on their existing knowledge and tools provided by the game site (e.g. autodraft)?

Q2: *What are players' data collection and analysis practices?* For players that include data collection and research as part of athlete selection,(a) What resources are they using for data gathering? (b) What tools are they using for data analysis?

Q3: *How do players scope the data they include in their decision process?* Do players find the volume and velocity of athletes' statistics overwhelming? How many data types are included in their data analysis and decision making process?

Q4: *What is the time commitment and activities of players?* What are users' habits after the constitution of their team? Is there activity during the off-season?

Q5: *What aspects of fantasy sports are problematic?* What are constraints or issues perceived by users that impede their gameplay? What are recommendations for improvement?

2. METHOD AND RESULTS
A questionnaire was administered on Amazon's Mechanical Turk to help answer our five research questions. 160 responses were received to the 57 questions about respondent demographics, fantasy sports experiences and practices, and suggestions. Due to space constraints, we limit the discussion here to a few questions of engagement with the game and game data.

Participants were limited to Mechanical Turk users from the United States that had achieved a 95% or higher acceptance rate on past Human Intelligence Tasks. The 160 participants included 136 males (85%) and 24 females (15%). In terms of age, 71 (44%) were 18 to 25, 80 (50%) were 26 to 40, and 9 (5%) were over 40. The time commitment of players is quite high with almost all respondents reporting at least two interactions a week and slightly more than half reporting at least five interactions per week. (See Figure 1.) When asked about the average length of these interactions, 87.5%

of the respondents reported at least 15 minutes. This is shown in Figure 2.

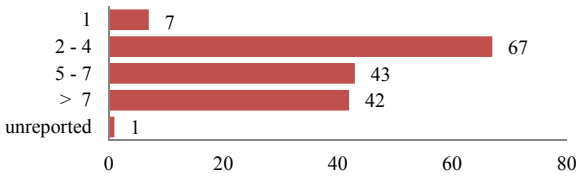

Figure 1. Interactions per week

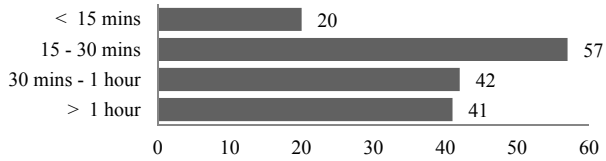

Figure 2. Length of average interaction during last season

In terms of data gathering and analysis during decision making, they reported most often using data provided by on-line sources with slightly more reporting use of external data/information sources than those that primarily used in-game data/information.

When asked about the number of variables or data streams used in making decisions, exactly half of the respondents reported using at least five data streams, as shown in Figure 3. The tools used to analyze this data were varied; most respondents reported using a combination of in-game tools for searching and sorting players and external tools, e.g.MS Excel for more complex analyses.

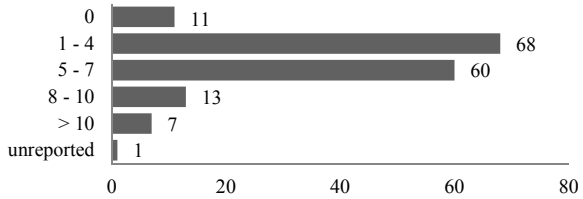

Figure 3. The number of data fields employed during the athlete selection process.

Comments from respondents indicated that the availability of data and tools in the game engine increased the likelihood of its use. This points towards the need for prediction game engines to support more than just the basic prediction and scoring functions. To encourage learners to increase their data interpretation skills, the data and the analysis tools need to be provided within the game engine.

Our survey also asked about motivations for playing fantasy sports. As found in other surveys on this topic, sports engagement, entertainment, socializing, and intellectual challenge are common to many players. Social reasons were discussed in open-ended responses more than any of the other factors. This is important for the design of prediction games in non-entertainment domains. It confirms that communication channels provided in the game (e.g. player-to-group forums and player-to-player direct messaging) are components that keep people engaged.

3. DISCUSSION AND CONCLUSION

This study provides insight into the data gathering and analysis practices of fantasy sports users, their time commitment to the game, and their motivations to their efforts.

During the athlete selection process to build their team, fantasy sports players prefer online resources like sports sites, social networks or discussion forums. They often employ strategies that involve analyzing the athletes' statistical data using a combination of tools built in to the game and external tools (excel or paper). Most of our respondents consider seven or fewer data fields during their analysis. Once their team is complete, they at least occasionally frequent the fantasy sports site to either get the latest updates on athletes, monitor their team scores, their rank and status in the league or to interact with other players.

Social interaction is a major factor for motivating our participants to play fantasy sports in the first place. Fantasy sports are considered important by many for maintaining family, friends, and workplace relationships. Other significant motivations include the engagement with sports, intellectual challenge and competition, and pure entertainment. Prizes were less commonly viewed as an important motivator.

The main complaints about fantasy sports concerned user interface issues and data limitations. The user interface issues included frustration over the inability to control data presentation and organization, e.g. the inability to sort or filter data. Participants also indicated that a simpler user interface would be welcome for novice players. Second, participants report that missing historical data and slow data updates undermine their data analysis activities. Finally, some players wanted better data analysis support, such as tools enabling data comparison/visualization and data export capabilities.

Our results have implications for improving fantasy sports systems and the design of data-driven web entertainment systems in other domains. First given the different responses regarding the volume of data and data analysis practices, future systems should consider enabling different levels of complexity to accommodate different classes of users. Systems for more advanced players should support the ability to filter data sets and compare/visualize subsets of data. They should also enable exporting data in an open format so that the user can transfer data to the tools of his/her choice and importing results that rank potential decisions.

Social interaction is an important factor for many forms of web entertainment. As such, design of data-driven entertainment should foster community building through competition or challenges to be overcome collaboratively. Game design should sustain communication at multiple levels. This includes forums for general discourse about the data domain, other forums at the contest level to discuss and set the rules or voice and resolve concerns of the members, and a chat system to facilitate user-to-user conversation.

As ever more data streams are made available via the Web, there are more opportunities for developing entertaining activities that motivate people to learn about this data. System developers can design activities that provide an intellectual challenge, generate opportunities to show one's knowledge or abilities, and foster social engagement and competition. The results here provide insight into such data-driven web entertainment systems.

4. ACKNOWLEDGEMENTS

This work was supported in part by NSF grant DUE-0938074.

5. REFERENCES

[1] FSTA. 2014. Industry Demographics. November 26, 2013. Retrieved from http://www.fsta.org/industry_demographics

[2] Shipman, F. "Blending the Real and Virtual in Games: The Model of Fantasy Sports", *Proc. ICFDG*, 2009, 169-17