Games with a Purpose for Social Networking Platforms

Walter Rafelsberger, Arno Scharl
MODUL University Vienna, Department of New Media Technology
Am Kahlenberg 1, 1190 Vienna, Austria
(t) +43-1-3203555-500; (f) +43-1-3203555-903
www.modul.ac.at/nmt

{rafelsberger, scharl}@modul.ac.at

ABSTRACT

The online games market has matured in recent years. It is now a multi-billion dollar business with hundreds of millions players worldwide. At the same time, social networking platforms have witnessed unprecedented growth rates and increasingly offer developer interfaces to leverage and extend their built-in core functionality. Benefiting from these trends, games with a purpose are a proven way of leveraging the wisdom of the crowds to address tasks that are trivial for humans but still not solvable by computer algorithms in a satisfying manner. This paper presents an application framework to develop interactive games with a purpose on top of social networking platforms, suitable for deployment in both mobile and Web-based environments. A set of analytic tools helps to evaluate the results and to pre-process the gathered data for use in external applications.

Categories and Subject Descriptors

H.5.3 [**Group and Organization Interfaces**]: Web-based Interaction; K.8 [**Personal Computing**]: Games

General Terms

Measurement, Design, Human Factors, Languages, Verification.

Keywords

Games with a purpose, social networking platforms, application framework, sentiment detection.

1. INTRODUCTION

Games with a purpose leverage collective intelligence, which is described as combining "behavior, preferences, or ideas of a group of people to create novel insights" [1]. Collective intelligence from groups of people often produces better results than individual domain experts [2]. Games with a purpose have been used successfully to solve problems that computers cannot yet solve, such as tagging images [3] and annotating content [4]. The main challenges when creating such games are motivating users to play the game while generating useful data, and ensuring that the process yields unbiased results.

The social application framework presented in this paper builds upon the lessons learnt from the *Sentiment Quiz*, a Web-based social verification game for sentiment detection released in June 2008 and developed as part of the *US Election 2008 Web Monitor* (www.ecoresearch.net/election2008). The project aims to gain new insights into information diffusion via interactive online media, and the interdependence of news media coverage and public opinion [5].

To capture the editorial slant of news media coverage, the system automatically measures media attention (= frequency of candidate references) as well as media sentiment (positive versus negative). Sentiment detection is a particular challenge in natural language processing, and the focus of the related research project RAVEN (Relation Analysis and Visualization for Evolving Networks; www.modul.ac.at/nmt/raven).

Following a crowd-sourcing approach in the form of games with a purpose, *Sentiment Quiz* serves three main purposes in the context of this research: (i) it helps create tagged dictionaries in multiple languages, which are a prerequisite for most sentiment detection methods; (ii) it yields a large set of manually tagged sentences to evaluate and fine-tune the sentiment detection algorithms; (iii) in conjunction with extensive user polling between January and November 2008, it allows analyzing *hostile media effects* (= the different perception and interpretation of content depending on the reader's political orientation).

There are various methods to detect sentiment automatically, usually based on a combination of natural language processing and machine learning. Yet most of these methods require human input in the form of dictionaries and evaluation sets. Leveraging the wisdom of the crowds by engaging users in online games addresses the scarcity of human resources to tackle such tasks. Crucial to the success of this approach are effective strategies to attract Internet users and achieve a critical mass of players. Deploying games with a purpose with social networking platforms like Facebook (www.facebook.com) enables developers not only to harness the functionality of the platform's API, but also to tap into a large online community with a proven affinity towards Web 2.0 applications. Other benefits include the possibility to identify players via their Facebook ID, and the availability of built-in viral notification systems to increases the visibility of the application among the users of the network.

Preliminary results from the *Sentiment Quiz* reflect the potential of games with a purpose for research projects. Between July and October 2008, more than 1,000 Facebook users installed the application and evaluated more than 30,000 quotes on the US presidential candidates according to a five-point sentiment scale (very

negative, negative, neutral, positive, very positive). The application has been re-launched after the election in November 2008. Players are not asked to evaluate sentences anymore, but single words in a specific context. This yields tagged dictionaries in multiple languages, which will guide further research into sentiment detection and improve the algorithm's accuracy.
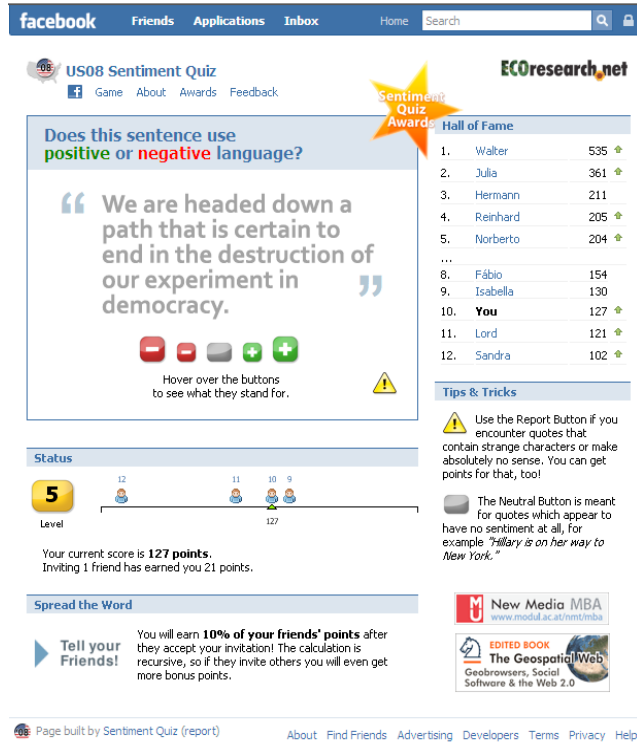

Figure 1. Screenshot of the US08 Sentiment Quiz
(www.modul.ac.at/nmt/sentiment-quiz)

2. GAME DESIGN

Games have to include incentive schemes that engage users while delivering valuable information to solve the underlying problems. Implementing games via social networking platforms addresses a vast number of potential players. On such platforms, users are not aware whether their input is matched against real-time input from another anonymous player, stored records from previous users, or system-generated values. Players score if the results match – independent of the origin of the other values. If a certain number of players agree on a solution, it will be assumed correct and taken out of the game. There are two types of games:

- Closed-end games provide a set of responses, either in binary format (are A and B synonymous?), scaled (how relevant is this concept for the domain?), or multiple choice (choose correct relation type between A and B). Users score if their input matches the answer of other players.

- Open-end games invite users to make arbitrary suggestions (e.g., image tagging games). Users have a limited amount of time to type in their suggestions. They get points for every suggestion made by the other player as well. The most popular suggestions might then be fed into a 'closed-end' game to further refine the confidence values.

2.1 Incentive Scheme

Successful games can motivate users to spend significant amounts of time on productive tasks [3]. When games deployed within a specific community (for example, environmental activist platforms and networking sites), users will have a high level of intrinsic motivation to play games that help build a shared knowledge repository. Unless group features are supported and accepted, this is not the case on social networking sites. In such a generic setting, the main motivation to play the game will be the level of entertainment it provides, as well as additional incentives such as awards and prizes that might be built into the application. Developers of applications for social networking sites, however, have the advantage of embedded viral mechanisms that can be leveraged. Users are encouraged to invite friends to install the application and receive a certain percentage of the points made by the players whom they invite. In the high score ranking, users not only see players with the highest scores, but also their friends and other players who are currently playing.

2.2 Task Complexity

Applying the game concept to more complex tasks poses new challenges. With the exception of image labeling games, previous projects usually featured simple multiple-choice questions. The findings in terms of usability and user motivation learned from the *Sentiment Quiz* will be applied for conceptualizing and implementing more complex gaming environments. Merging several types of games in a meta-gaming platform (e.g. combining sentiment detection, image tagging and ontology learning) should increase the game's attractiveness, reduce the risk of cheating, allocate collective intelligence more efficiently by prioritizing tasks across game types, and help avoid the situation that dedicated players run out of new challenges.

2.3 Validation of Results

A crucial task when applying games with a purpose is to make sure that the games yield unbiased results. On a social networking site, users can identify other players and might collaborate to manipulate the game; e.g. by agreeing in advance on the answers to a limited set of questions.

A number of simple measures can be taken to ensure output of high quality: (i) hide the identity of the other player; (ii) analyze the temporal distribution of answers; (iii) assign trust values to each player, which in turn determine the impact of their answers – e.g. insert questions with known answers into the exercise queue and identify users who tend to score low on these questions; (iv) avoid exploitable patterns in the sequence of answers, since users who identify the pattern could quickly earn credits without actually solving the puzzle.

3. APPLICATION FRAMEWORK

Most social networking and Web 2.0 platforms are commercial in nature and hosted by companies participating in highly competitive markets. To meet the requirements of scientific projects, social applications should provide analytical functions that are often missing in commercial implementations. The framework presented in this paper addresses this shortcoming by supporting crowd-sourcing applications in the form of Web-based games with a purpose offered through social networking platforms. Fol-

lowing a generic approach, the framework is compatible with a range of developer platforms such as *Facebook*, *iGoogle*[1] and *Netvibes*.[2] It acts as a wrapper that gives developers more flexibility and enables them to implement applications on multiple social platforms. The framework has been evaluated via two social media applications for the Facebook platform: (i) *Election Monitor*, a user poll focusing on the US Presidential Election 2008 that invites users to cast their votes, and (ii) *Sentiment Quiz,* an interactive game with a purpose, which asks users to assess sentiment of quotes related to the presidential candidates.

- *Election Monitor* allows multiple votes after a given time offset, and also captures the users' opinion of how likely it is that their preferred candidate will win the election. Using privacy options, users can specify if they want to take part anonymously or make their votes public, and whether they will provide their demographic data for statistical purposes.

- *Sentiment Quiz* asks players to evaluate the sentiment of sentences like *"Candidate X is going to win the election"*. The goal is to evaluate the language of the sentence, and not whether the player agrees with the statement. After clicking on the button that best represents the players' opinions, the game will match their choices with answers from other players. Points are being awarded depending on whether the paired answers match. Players gain three points for exact matches, one point for the same direction (positive vs. negative), but loose one point for conflicting answers. The application uses Facebook's viral notification features to attract users, offers high score rankings and a level-based incentive system with network-marketing inspired elements (e.g., the possibility of inviting contacts from the social network to the game and thereby earning 10% of their points).

The applications serve multiple purposes: (a) By playing *Sentiment Quiz*, users help optimize methods for automatically detecting the sentiment of sentences from an archive of election-related news media articles. Based on a user's progress in the game, it is possible to select sentences with a certain length or sentiment. The game automatically offers short sentences with obvious sentiment in the beginning, making the game easier for first-time users and continuously increasing difficulty as the user progresses. (b) Individually, both applications track user behavior and investigate information dissemination within social networks – visual means for this investigation will be introduced in the section on "Visualization Services". (c) Analyzed in combination, the gathered information reveals different perceptions and interpretations of content depending on the readers' political orientation (= hostile media effect). Explanations for this effect range from selective recall (preferentially remembering hostile content), selective categorization (perceiving the same content differently) and conflicting standards (considering hostile content as invalid or irrelevant). Recent research suggests that selective categorization best explains hostile media effects [6].

3.1 Supported Platforms

Facebook has been chosen as a platform for both applications, as it offers access to one of the largest online communities. Whereas Facebook can be seen as a walled garden where most applications

bring information into its ecosystem, Google creates services which act upon existing information on the Web and multiple online communities – a combination of their Gadget Platform,[3] Open Social[4] and Social Graph APIs.[5] Similar to Google's iGoogle, Netvibes first concentrated on a widget platform and introduced social networking in a later stage. Facebook first focused on community building, but is working on widget functionality in an attempt to open its services for regular users (in addition to their popular development platform).

The framework acts as a wrapper for the APIs of social networking services and allows developing the same application for multiple platforms with one generic data repository. It offers a plug-in architecture and caching on the database, file and memory levels. At its core, the framework offers user management to match multiple accounts of a user on various services and follows a combination of the *model-view-controller pattern* [7] and *object-relational mapping* [8] as design principle. Once a user starts playing the game on a given platform, the system creates a meta ID – an internal representation mapped to the user's platform ID. Users who continue to play the game on a different platform are again mapped to the meta ID. This tracking of users across multiple sessions and platforms avoids assigning the same task twice and helps maintain individual scores in a consistent manner.

Users who play games with a purpose complete simple tasks, which in a different context might be considered repetitive work. What motivates users to do this? The tasks are presented in a way that entertains the users and motivates them by using game design patterns [9]. From the developer's perspective, the primary goal is to maximize the number of successfully completed tasks. In addition to user management via meta IDs, the framework offers interfaces to access data sources and map them to such tasks. In the case of *Sentiment Quiz*, sentences are sourced from the repository of Web documents gathered through the *US Election 2008 Web Monitor*[6]. A rule set filters the data to comply with a given task.

During the first phase of the *Sentiment Quiz,* the research interest focused on sentences that express sentiment towards a US presidential candidate. The corresponding rule set used a set of regular expressions to make sure that each sentence presented to the user contains a reference to a single candidate (excluding general election coverage as well as sentences talking about more than one candidate). After filtering, the pre-processed data defines a set of tasks for the users to act upon.

3.2 Types of User Activity

The framework distinguishes between general and task-specific user activities. The general category includes standard functionality such as viewing help pages or initial requests to view the gaming interface. The task-specific category includes all user activities which act upon data related to a task. The core functionality of the framework is to process these requests and trigger processes on various levels. It frees developers from dealing with general Web request processing and lets them focus on the definition, presentation and evaluation of tasks.

User activities trigger different types of processes. The first step is usually the validation of the submitted data through the *task rule set*, assessing whether the task has been completed. Within the *Sentiment Quiz,* for example, a sentence has to be evaluated by ten different players who were not identified as "rogues". Once this number is reached, the task cannot be assigned to other players anymore and is made available for further analysis. The *gaming rule set* defines the feedback given to the player – this can but does not have to depend on the task rule set. Examples for feedback are inline pop-ups when reaching a new level or updating the user's score and the built-in monthly, weekly and daily rankings.

3.3 Lifestream and Viral Notification Services

Depending on the capabilities of the underlying platform, both rule sets can trigger lifestream updates or send notifications to the users' social network. Mini-feeds and profile pages document user activities – e.g., users completing a task, overtaking another player or reaching a new level. While the gaming rule set is essential to motivate a single user, the layer for accessing lifestreaming and viral notification services of social networking platforms not only entices existing users to return to the game, but also attracts new players. This layer has to be used in moderation. If users install the application, they are giving the application permission to access information about confirmed contacts and address them through notifications (with certain restrictions). This gives the application the privilege to enter a circle of trust and helps spread information among contacts more easily, given the users' reputation in their social sub-network.

Applications that abuse this functionality by sending misleading, unrelated or simply too many notifications to lifestreams or other users reverse the effect and will only gain limited acceptance. Keeping these principles in mind, the framework's messaging system was designed to send notifications only in the context of a user's current activity and avoid so-called "mini-feed spam" (the messaging system uses Facebook's JavaScript implementation in conjunction with the Google Charts API to display voting results and user activity).

3.4 Visualization Services

Rhizome visualization based on *Processing* (www.processing.org) is used for tracking activities within the social network. It offers an OpenGL-based graph layout engine and interface component to access information for social network analysis. *Rhizome* uses a special form of radial layout that combines a circular layout similar to *Circos* [10] with force-directed, partially fixed graphs. The interface provides a *Selection Lens*, a spatially bound visual filter also referred to as *Magic Lens* in the literature [11]. The dynamic, physics-based display allows tracking content elements and social network characteristics simultaneously.

4. CONCLUSION AND OUTLOOK

The *Sentiment Quiz* outlined in this paper invites Facebook users and their network of online friends to evaluate whether sentences and dictionary terms express positive or negative sentiment. The built-in notification system and real-time statistics on the users' progress help engage users and leverage the wisdom of the crowds for scientific purposes.

Besides validating and optimizing algorithms for automated sentiment detection, social evidence sources can be used in several ways to support knowledge acquisition. Such sources include content created by users, resource annotations and other types of metadata as well as relations between users. All these sources can be used to build a new generation of Web-based applications. Creating shared meaning through collaborative ontology building via Web-based interfaces is another excellent example. User suggestions and feedback regarding ontological entities can be complemented by multi-player online games to evaluate suggested changes and solve potential conflicts on the concept, property, relation and instance level. Knowledge derived from folksonomies can be incorporated into the knowledge repository directly, while social bookmarking tools help identify valuable resources for focused crawling.
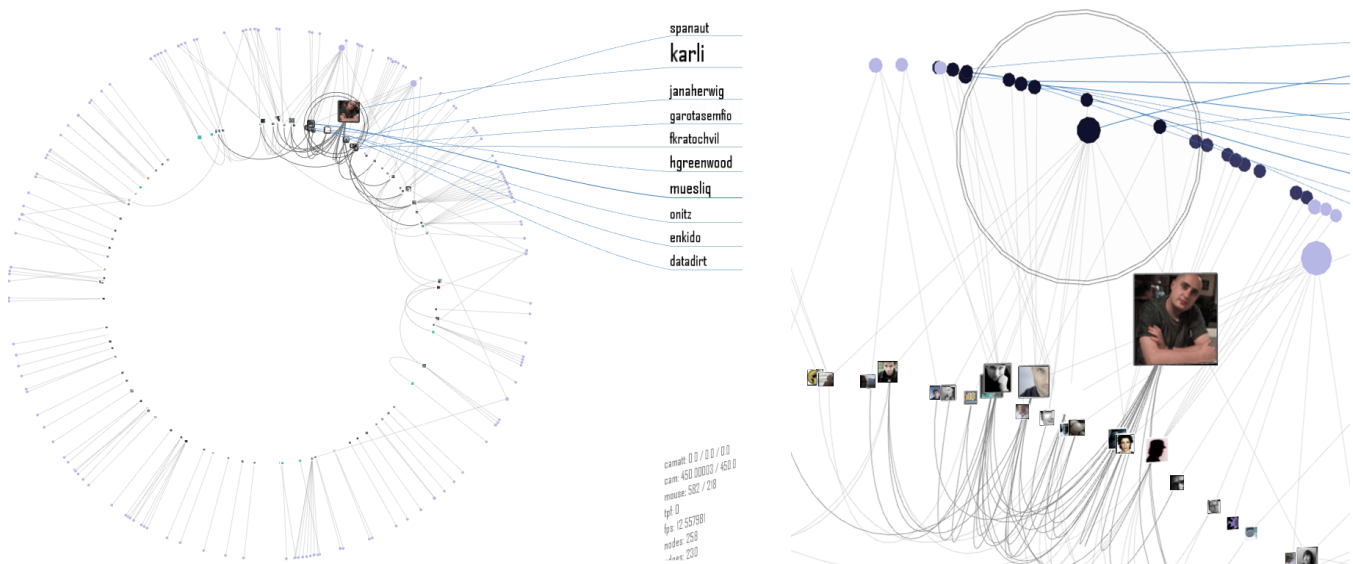

Figure 1. Real-time Rhizome visualization of Twitter conversations (www.rhizomenavigation.net)

Future research will go beyond single applications and investigate social application frameworks flexible enough to accommodate a range of tasks. Allocating human resources effectively in such a framework requires continuously updated lists of pending tasks. Each task has to be assigned a priority that takes into account both the urgency of the problem (i.e., whether the information is required for a real-time application) as well as the expected information gain.

5. ACKNOWLEDGEMENTS

6. REFERENCES

[1] Segaran, T. Programming Collective Intelligence: Building Smart Web 2.0 Applications. O'Reilly Media, Inc., Beijing, 2007.

[2] Surowiecki, J. The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations, London, 2004.

[3] Ahn, L. v. Games with a Purpose. Computer, 39, 6 (2006), 92-94.

[4] Siorpaes, K. and Hepp, M. Games with a Purpose for the Semantic Web. IEEE Intelligent Systems, 23, 3 (2008), 50-60.

[5] Scharl, A. and Weichselbraun, A. An Automated Approach to Investigating the Online Media Coverage of US Presidential Elections. Journal of Information Technology & Politics, 5, 1 (2008), 121-132.

[6] Schmitt, K. M., Gunther, A. C. and Liebhart, J. L. Why Partisans See Mass Media as Biased. Communication Research, 31, 6 (2004), 623-641.

[7] Curry, E. and Grace, P. Flexible Self-Management Using the Model-View-Controller Pattern. IEEE Software, 25, 3 (2008), 84-90.

[8] Ambler, S. W. Agile Database Techniques. Wiley & Sons, 2003.

[9] Bjork, S. and Holopainen, J. Patterns in Game Design. Charles River Media, 2004.

[10] Krzywinski, M., Schein, J., Birol, I., Jones, S. and Marra, M. Circos - An Information Aesthetic for Comparative Genomics, http://mkweb.bcgsc.ca/circos. Canada's Michael Smith Genome Sciences Centre, Vancouver, BC, 2009.

[11] Bier, E., Stone, M. and Pier, K. Enhanced Illustration Using Magic Lens Filters. Computer Graphics and Applications, 7, 6 (1997), 62-70.