The Chain Model for Social Tagging Game Design

Lilian Weng[1] Rossano Schifanella[2] Filippo Menczer[1]

[1]School of Informatics and Computing, Indiana University, Bloomington, IN, USA
[2]Department of Computer Science, University of Turin, Italy

ABSTRACT

We introduce the *Chain Model* for object association games, and two social tagging games based on this model. *GiveALink Slider* and *Great Minds Think Alike* harness human power to generate large streams of high-quality social tagging data. Such social annotations are utilized to help people organize Web resources and infer semantic relationship, which in turn can enhance Web applications such as search, recommendation, navigation, and categorization. The two games leverage several design features as well as external social media resources to create entertaining incentives for the players to generate reliable annotation data.

Categories and Subject Descriptors

H.3.4 [**Information Storage and Retrieval**]: Systems and Software—*Information networks*; H.5.2 [**Information Interfaces and Presentation**]: User Interfaces; K.8.0 [**Personal Computing**]: General—*Games*

Keywords

Social tagging, social media, games with a purpose, browser, iPhone

1. INTRODUCTION

Social tagging provides us with a powerful way to manage online resources collaboratively. Users can freely choose words to describe resources and therefore resources get descriptions from various users represented as sets of tags, collectively forming a *folksonomy.* Tags are widely used in many Web 2.0 and social media Web sites, by which users can organize their own collections, discover interesting resources, and find friends with similar interests. These folksonomies, generated by the power of the crowd, not only help enrich the semantic space of online resources, but also enhance the performance of many Web services, such as search, recommendation, navigation, and categorization.

In current social tagging systems, there is little motivation for the majority of users to annotate many resources with sufficient numbers of accurate tags such that they can help others, and the number of new pages that are posted per day to social annotation systems is small compared to the rate of growth of the Web [4]. This causes a significantly sparse semantic network of social annotations. Without any control on tagging behaviors, users can easily employ poor tags or even abuse the system by spamming [5]. They can use inappropriate tags, including tags that are unrelated to a resource, too general to meaningfully describe a given resource, or so specific that they are only meaningful for one individual.

Here we introduce two social tagging games, *GiveALink Slider* and *Great Minds Think Alike*, to accelerate the generation of large streams of high-quality social annotations as a side effect of enjoyable activities. Using *Games With A Purpose (GWAP)* to engage humans in the solution of computational problems is an approach that has gained popularity in recent years in the field of Human Computation [1]. GWAP, or crowdsourcing through play, are designed in a way that players can help solve hard computational problems while having fun. The best known instance of GWAP is the ESP Game [1], which takes on the challenge of image recognition by asking pairs of players to reach an agreement on image labels. Games are also used for scientific and other purposes. Foldit is a game for non-expert players to handle computationally difficult protein folding problems in biochemistry [2]. Galaxy Zoo (`galaxyzoo.org`) leverages human image recognition abilities to classify celestial objects. Microsoft uses games in Club Bing (`clubbing.com`) to promote the Bing search engine. In all the club games, players have to use Bing as a tool to collect enough hints for winning the game and therefore get a chance to try some functions that they might not have noticed before. Great Minds was partly inspired by an online word association game called *Human Brain Cloud*, which generated interactive visualizations of crowdsourced word networks. The Slider was built upon an early social tagging game prototype [11].

The development of games to help enhance both the quantity and quality of annotation data is an integral component of the `GiveALink.org` project, which broadly examines several aspects of social tagging to foster the construction and applications of socially driven semantic annotation networks. Both of the games introduced here are built upon existing work in the GiveALink project. Previous research includes the design of effective similarity relationships among pages, tags, or users [6, 10], applications to page recommendation [8], exploratory navigation interfaces [3], bookmark management [9], and social spam detection [5].

2. THE CHAIN MODEL

The two tagging games proposed here are built upon the idea of making a chain of semantically related objects. The objects are connected based on a measure of similarity among them, and the players extend the chain by making these relationships explicit. We formalize this idea in a *chain model* for object association games.

FDG '11 June 29-July 1, 2011, Bordeaux, France.

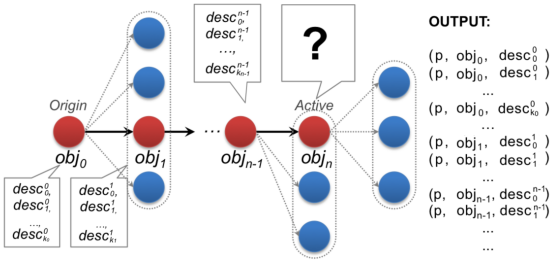

Figure 1: The chain model for object association games.

The model can be used to design games with the purpose of collecting descriptions for Web resources, tags, media, people, and geographical locations.

Let us define a chain as an ordered sequence of objects $\langle obj_0, \ldots, obj_n \rangle$, in which obj_0 is the *origin* of the chain and the last element obj_n is called the *active* object. Chain model games allow players to characterize obj_i with a set of descriptions $D^i = \{desc^i_0, \ldots, desc^i_{k_i}\}$ in some language. At each step the player p can add a new description to the active object obj_n or make a game *move* to extend the chain. A *move* can be represented as a transition $obj_n \overset{\mu}{\mapsto} obj_{n+1}$ where μ is a user-defined relation between the *source* and the *target* of the move. To support the player in the decision of the next object in the chain, the model suggests a set of *candidate* objects C^n that are computed from a system-defined measure of similarity with respect to the active object obj_n. Fig. 1 illustrates the chain model and the stream of annotation triples that are generated by playing the game.

The GiveALink Slider game uses Web pages as objects in the chain, and the players are asked to annotate each page with a set of tags as descriptions. A move is made by the selection of a Web page from the candidates that are defined through the similarity metric adopted in the GiveALink system. On the contrary, in the Great Minds Think Alike game, objects are words and the descriptions are sets of media content from public social media sites that are connected to the active word. To build the chain a player has to enter a term that she believes is related to the current one, or pick a suggested tag. In the remainder of this section we describe some common aspects that characterize the chain model as a platform to implement games with a purpose.

Output: The output of the game is a stream of posts $\langle (p, obj_i, D^i) \rangle$ where $0 \leq i \leq n$, that represent how a player with handle p describes each object obj_i with some language. The annotations can be used later to improve the game as well as a number of other applications. One significant concern here is to guarantee the relevance of the player's description with respect to the object. The relationship can be verified either internally or externally. *Internal verification* is to check the input against other player submissions. When multiple players agree on a link between an object and a description, the relationship should be trusted. *External verification* consists of providing candidate descriptions from trusted sources, so that when a player selects an option, she agrees on the corresponding object-description link. The Slider uses internal verification, but it suffers from the cold-start problem: reliable links cannot be determined when few players are available. External services, such as GiveALink and Delicious, are used in lieu of other players when this issue arises. Great Minds employs external verification, where players are provided with Web resources from various social media sites as candidates descriptions.

Scoring: The interaction with different game components involves a rewarding scheme that transforms actions into points. In our model the players gain points for describing objects, discovering interesting resources, and extending the chain. Rewards are greater for novel descriptive terms or rare objects, to encourage fresh ideas and collect unique triples. The accumulation of points drives the achievement of *levels* of expertise or the gain of special *badges*, that are common features in game frameworks.

Social Interactions: The chain model strongly emphasizes social interactions between players that derive from explicit player handles in the annotations. While playing, it is likely that a user adopts the same description for a certain object or that she builds portions of the chain in a similar way as others. The more a user plays, the greater the chances that she can connect to like-minded people. In the Slider we visualize the connections among players with a network based on their pages and tags. The player expands her own network by playing and finding similar players. In Great Minds we suggest Facebook users that have made similar associations between words, fostering the formation of a community centered on affinity criteria as well as geographical proximity.

Enjoyability: To successfully design a framework for games with a purpose, we have to focus on mechanisms to both assure the achievement of the goal and increase the participation of the players. We thus need to make the game enjoyable. An example in our games is the ability for the players to explore the semantic space around their own actions. They may discover rare or interesting resources during the play. Because of the social aspects of the games, the players are able to know players with similar interests, which increases the fun of playing the games. Another way to make the game entertaining is to foster competition, therefore we display player rankings.

Platform: The deployment of a game is bound by the particular characteristics of the platform on which it is implemented. Although the Slider and Great Minds games are based on the same underlying model, they are implemented for different platforms. The former is a browser game in a desktop environment while the latter is a native iOS application. As a consequence, the Slider occupies a big screen with large images, such as thumbnails of Web pages and rich instructions. The limited space of mobile device displays, on the other hand, dictates a simpler user interface layout for Great Minds, with minimal interactions and information visualized.

3. GAME DESIGN

Both of our tagging games, built upon the proposed chain model, are aimed at collecting annotations of Web resources with trusted descriptive tags. To connect related Web pages or tags in the chain, both games need an effective measure of similarity among objects. To this end we employ a scalable, collaborative measure of similarity for social annotation systems, named Maximum Information Path (MIP) [7]. MIP similarities among pages or tags are available via the public GiveALink API (`GiveALink.org/api_doc`).

3.1 Browser Game: GiveALink Slider

In the GiveALink Slider,[1] the chain objects are Web pages. Players build chains of pages and generate descriptions by tagging these pages (Fig. 2). The origin page obj_0 in the chain is provided randomly by the game, then the player extends the chain by tagging the first page with one or more relevant tags. Each time the player enters tags for a page, the game displays a small set of pages based on the player's tags as candidates for the next object in the chain. The player is free to choose any candidate as the next page to tag.

The game mechanisms incorporate two important design principles from the chain model: (i) Players gain points for relevant tags, and lose points when the chain gets disconnected because of unre-

[1]Available at `http://slider.givealink.org/`

Figure 2: Slider interface for tagging and selecting pages.

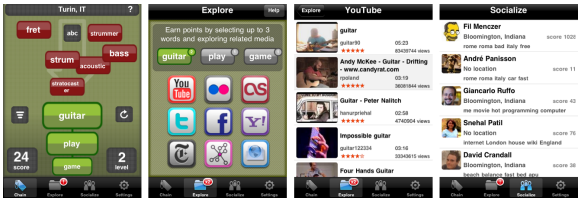

Figure 3: Great Minds main operational modes.

lated or unverified tags. (ii) Links between tags and resources that exist in the GiveALink triple store are trusted, so they can be used as benchmarks. New links are recorded until verified by multiple players internally.

The score of each tagging step is obtained by adding scores associated with each tag. Each tag is worth a number of points that depends on its relevance to the page and specificity. Specific, trusted, and novel tags are most valued. To make the game more enjoyable, there are six types of badges corresponding to different game tasks, and each has multiple levels of difficulty. These tasks deal with pages in the chain, connections with other players, and score milestones. The harder the task, the greater the bonus. Another incentive comes from visualizations of social ego-networks based on social connections among players.

3.2 iPhone Game: Great Minds Think Alike

Great Minds Think Alike (available on the App Store via `greatminds.givealink.org`) is a word association game that lets the player build semantic concept networks and explore similarity relations between people, tags, and media content (Fig. 3). Starting from a word, the player builds a chain by entering or selecting related terms among suggested options from the GiveALink knowledge base. The player gains points typing new words or exploring content from social media sites like Flickr, YouTube, Last.fm, Twitter, and others. Tags in the chain are geo-tagged and players are linked to their Facebook profiles so they can find like-minded people nearby.

The main purpose of the game is to collect annotations from a set of social media sites that are defined as reliable sources. This collection process is carried out by the navigation of media content exposed through public Web services and it is based on the assumption that a player agrees on the resource-tag links when he consumes the social content resource by tapping on it. A *move* is the selection of a *target* term that the player feels is related in some way to the *active* word. The target can be selected by entering a word from the keypad or choosing one of the candidate terms. When a move is performed, the game stores on the server a set of metadata that enables geo-social features.

Great Minds allows a player to gains points by (i) extending the chain with a move or (ii) exploring social media in the descriptive phase. In the former case, the game rewards the player when the target is specified by typing a new word to foster the introduction of serendipitous associations and fresh thoughts. In the latter case, the systems rewards the generation of trusted annotations: the more triples, the more points. To foster competition and make the game enjoyable we associate expertise levels with cumulative scores. Access to the player's Facebook profile provides a way to discover other players and socialize with them.

4. CONCLUSION AND FUTURE WORK

In current tagging systems users can easily employ poor tags or even abuse the system by tagging spam. Here, we leverage games with the purpose of improving both the quantity and quality of social annotations data. We designed and developed two social tagging games, the GiveALink Slider on browsers and Great Minds Think Alike on mobile devices. The basic design principles are derived from the Chain Model for games that generate object descriptions by object association moves.

The games are publicly available and are currently collecting annotations data generated by real players. We plan to extensively analyze the semantic relationship networks uncovered by the games and compare them with other social tagging systems. Since the proposed games are designed to generate high-quality social annotations, we plan to evaluate how efficiently and effectively this purpose is achieved.

Acknowledgments. Thanks to A. Panisson and S. Patil for assistance in the game implementation, and K. Gabler for sharing *Human Brain Cloud* data. This work is funded by NSF award IIS-0811994 and the project "Information Dynamics in Complex Data Structures" (PRIN).

5. REFERENCES

[1] L. V. Ahn and L. Dabbish. Designing games with a purpose. *Communication of the ACM*, 51(8):58–67, 2008.

[2] S. Cooper, A. Treuille, J. Barbero, A. Leaver-Fay, K. Tuite, F. Khatib, A. C. Snyder, M. Beenen, D. Salesin, D. Baker, and Z. Popović. The challenge of designing scientific discovery games. In *Proc. of the 5th FDG*, 2010.

[3] J. Donaldson, M. Conover, B. Markines, H. Roinestad, and F. Menczer. Visualizing social links in exploratory search. In *Proc. of 19th ACM Hypertext and Hypermedia*, 2008.

[4] P. Heymann, G. Koutrika, and H. Garcia-Molina. Can social bookmarking improve web search? In *Proc. of 1st ACM WSDM*, Palo Alto, CA, 2008.

[5] B. Markines, C. Cattuto, and F. Menczer. Social spam detection. In *Proc. of the 5th International Workshop on Adversarial Information Retrieval on the Web*, 2009.

[6] B. Markines, C. Cattuto, F. Menczer, D. Benz, A. Hotho, and G. Stumme. Evaluating similarity measures for emergent semantics of social tagging. In *Proc. of 18th WWW*, 2009.

[7] B. Markines and F. Menczer. A scalable, collaborative similarity measure for social annotation systems. In *Proc. of 20th ACM Conference on Hypertext and Hypermedia*, 2009.

[8] B. Markines, L. Stoilova, and F. Menczer. Social bookmarks for collaborative search and recommendation. In *Proc. 21st National Conference on Artificial Intelligence (AAAI)*, pages 1375–1380. AAAI Press, 2006.

[9] H. Roinestad, J. Burgoon, B. Markines, and F. Menczer. Incentives for social annotation. In *Proc. of 32nd ACM SIGIR*, page 838, 2009.

[10] R. Schifanella, A. Barrat, C. Cattuto, B. Markines, and F. Menczer. Folks in folksonomies: Social link prediction from shared metadata. In *Proc. of 3rd ACM WSDM*, 2010.

[11] L. Weng and F. Menczer. Givealink tagging game: An incentive for social annotation. In *Proc. of ACM SIGKDD HComp*, Washington, DC, 2010.