Gaming as a gateway: Ensuring quality control for crowdsourced data

Shaban Shabani, Maria Sokhn

HES-SO Valais Wallis, Institute of Information Systems, Sierre, Switzerland
shaban.shabani@hevs.ch, maria.sokhn@hevs.ch

Abstract. Crowdsourcing is a growing topic that has proved to be capable and cost effective solution for various tasks. Over the last decade it has been applied to numerous domains, both in research and enterprise contexts. Though there are several issues that remain open and challenging in crowdsourcing, here we address the issues of quality control and motivation. In this paper we present an ongoing work which explores the use of gamification in crowdsourcing settings, as means to: improve the task assignment and performance, incentivize people to participate and control the quality of their work. The developed crowdsourcing hybrid mobile application is applied to data within cultural heritage domain.

Keywords: crowdsourcing, gamification, quality control

1 Introduction

As an emerging service platform on the Internet, crowdsourcing has shown to be an effective solution for problems which for computers are difficult to solve and that require human intelligence [1]. The rise of the Web, the popularity of online platforms and mobile crowdsourcing [2], has made it easy to reach crowds of workers that are available at any time to solve micro-tasks called HITs (Human Intelligent Tasks). However, there are significant issues that appear in crowdsourcing. In these online platforms, money is the main encouragement for crowds to participate. For some workers, this is a motivation plus to increase their profits and become lazy by providing random answers and not considering seriously the HITs [3] or malicious workers that try to sabotage the system by providing intentionally wrong answers [4]. In these scenarios, crowdsourcing may yield to relatively low-quality results, hence control mechanisms should be applied to maintain the quality of the work. On the other side, in a crowdsourcing environment with no monetary reward, a significant challenge it how to incentivize people to participate. As a result, main incentives used are services, entertainment or learning [5].

This paper demonstrates a crowdsourcing application within cultural heritage domain that addresses these two issues through the application of game based elements to increase the engagement of participants and to improve the quality of their work.

2   Related work

A widely applied preventive quality control mechanism is gold questions [6] which consists of a set of questions for which answers are known in advance. Its goal is to remove unethical workers from a task and educate incompetent workers to improve the accuracy of their answers. However, choosing the test questions without further considering worker's behavior and profile is challenging as that effects the worker's output. Reputation based systems [7] additionally rely on reputation score that is computed based on worker's feedback history to detect fraudulent workers.

Unpaid crowdsourcing systems additionally face the issue of participation as workers need extra incentives. A good example are Games with a Purpose (GWAPs) where users solve tasks while playing the game [8] or by learning [9]. Another method is stimulating contributors [10] by social achievements in form of scores, ranks and badges.

3   Crowdsourcing application design

In this work, we use data from two archival institutions for digital culture: Mediatheque[1] and Digital Valais[2]. Datasets contain historical documents (audio, images, video and text) about the canton of Wallis in Switzerland. The crowdsourcing application enables users to participate through:

- Data sharing - allows users to insert new data to the multimedia repository and provide metadata by following a 5-step process in an interactive and self-descriptive user interface (illustrated in Figure 2a). The application makes it relatively easy for the users to transfer their valuable documents such as photo and video albums, and share them through the application with other users.
- Data annotation - while integrating existing data and inserting new data into the repository, two different issues occur: *missing information* and *conflictual information*. For instance, some data miss particular information about the location, description, and sometimes two different dates appear. In this section participants are asked to annotate missing parts and/or solve conflicting data. These issues are designed and generated as small tasks/questions where users are asked to choose or provide an answer (Figure 2c). As the data is specific and related to historical cultural heritage of the country, maintaining a high quality of users' annotations is crucial, therefore we apply a quality control method through gamification described as follows.
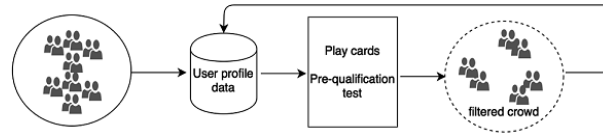
---

[1] http://www.mediatheque.ch/
[2] http://www.valais-digital.ch/

Fig. 1: Quality control process

(a) Inserting new data

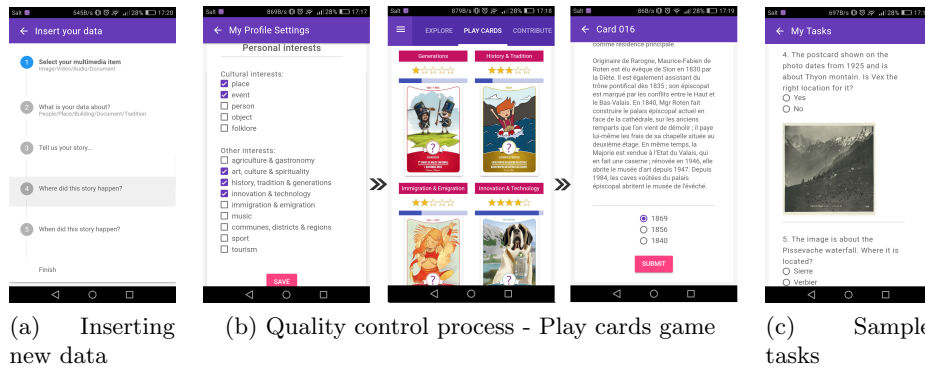(b) Quality control process - Play cards game

(c) Sample tasks

Fig. 2: Crowdsourcing application screenshots

To enable the requester, assess the quality of posted contributions, and to make the crowdsourcing tasks more attractive and engaging, we apply a game based quality control mechanism that considers users' profiles and their interests. This process is illustrated in Figure 1. The *play-cards* game adapts game-design elements in a non-gaming context. It acts as a pre-qualification test for users that are motivated to annotate data, similar to *gold questions* [6]. This test consists of 195 cards grouped in 13 categories and each one of these cards has a story behind. Initially, users provide information about their profile, especially information about their interests which fall within one or more of these 13 categories. Depending on this information, they are forwarded to read and answer questions/cards related to their predefined interests, hence avoiding unfair exclusion of workers due to non-relevant questions. For instance, a user that has chosen *sport* as his area of interest, he is asked to answer questions within that topic. He chooses to read the story of the cards and tries to guess the year which is related to the topic. To boost the motivation of users, two joker cards can be used. If the user successfully answers 70% of the cards, he is considered later as a potential worker for solving micro tasks related to that category. User's performance is visually displayed: on each category, the accuracy is shown in form of stars and a progress bar shows the completeness of the questions of that category (depicted in Figure 2b).

To further motivate qualified users to participate in data annotation, we apply a reputation mechanism. Depending on their level of contribution, users

gain reputation points and titles. Top contributors will have the chance to receive public recognition by the archival institutions.

4 Conclusion and future work

Game design elements are important to affect the human motivation and participation in crowdsourcing. In general, gamified approaches have reported an increase in engagement and output quality of workers [11]. In the near future, we plan to run an experiment by inviting users to use our application and collect data. To ease the challenge of participation, we have developed a hybrid mobile, targeting stationary and mobile users. Further analysis will be employed to evaluate the quality of users work as well as their interaction and perception about the crowdsourcing application.

Acknowledgement

References

1. Yuen, M.C., King, I., Leung, K.S.: A survey of crowdsourcing systems. In: IEEE Third International Conference on Social Computing. (2011)
2. Ren, J., Zhang, Y., Zhang, K., Shen, X.: Exploiting mobile crowdsourcing for pervasive cloud services: challenges and solutions. IEEE Comm. Magazine (2015)
3. Ipeirotis, P.G., Provost, F., Wang, J.: Quality management on amazon mechanical turk, ACM (2010)
4. Li, H., Yu, B., Zhou, D.: Error Rate Bounds in Crowdsourcing Models. ArXiv e-prints (2013)
5. Garcia-Molina, H., Joglekar, M., Marcus, A., Parameswaran, A., Verroios, V.: Challenges in data crowdsourcing. IEEE Trans. on Knowledge and Data Engineering (2016)
6. Oleson, D., Sorokin, A., Laughlin, G., Hester, V., Le, J., Biewald, L.: Programmatic gold: Targeted and scalable quality assurance in crowdsourcing. In: Proceedings of the 11th AAAI Conference on Human Computation, AAAI Press (2011)
7. Daltayanni, M., de Alfaro, L., Papadimitriou, P.: Workerrank: Using employer implicit judgements to infer worker reputation. In: Proceedings of the Eighth ACM International Conference on Web Search and Data Mining. (2015)
8. von Ahn, L., Dabbish, L.: Designing games with a purpose. (2008)
9. von Ahn, L.: Duolingo: Learn a language for free while helping to translate the web. In: Proc. of the International Conference on Intelligent User Interfaces, ACM (2013)
10. Lee, T.Y., Dugan, C., Geyer, W., Ratchford, T., Rasmussen, J., Shami, N.S., Lupushor, S.: Experiments on motivational feedback for crowdsourced workers (2013)
11. Morschheuser, B., Hamari, J., Koivisto, J.: Gamification in crowdsourcing: A review. In: 49th Hawaii International Conference on System Sciences (HICSS). (2016)