

Collaborative Systems with Applications for Social Good

Rakshit Agrawal
University of California, Santa Cruz
Santa Cruz, CA 95060
ragrawa1@ucsc.edu

ABSTRACT

Crowdsourcing and collaborative systems have become an important part of Computer Science system deployments. The research discussed in this paper designs and explores the use of collaborative systems for crowdsourced user participation in different kinds of tasks. Application focus of projects discussed here is mostly towards social good. This paper provides an overview of my research objectives and approach, and identifies my work on both usability of systems as well as data specific definitions for them. Acknowledging the importance of user participation in development projects, I work on structuring systems in a way that they can extract best response from participants and help in the overall goal. The paper also describes my objectives to study user behavior based on their participation in various collaborative systems.

Keywords

Crowdsourcing, HCI, Data Science, Social Good

1. INTRODUCTION

Crowdsourcing with its own variety of definitions has become an important part of Computer Science system deployments. One significant way to define this is that crowdsourcing can be seen as inclusion of human component in the computational process in order to either accelerate, tune or improvise existing systems. This means that we can take problems and build systems around them where both machine and human serve as processing modules.

The research discussed in this paper designs and explores the use of collaborative systems for crowdsourced user participation in different kinds of tasks. Combining my previous research in the field of Information and Communication Technologies for Development (ICT4D), I work on design and development of collaborative systems for application towards social good.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author.

Copyright is held by the owner/author(s).

OpenSym '15, August 19-21, 2015, San Francisco, CA, USA

ACM 978-1-4503-3706-9/15/08.

<http://dx.doi.org/10.1145/2789853.2789858>

This paper provides an overview of my research objective and approach. I am developing collaborative systems and focus on both usability as well as data specific definitions. Acknowledging the importance of user participation in development projects, I work on structuring systems in a way that they can extract best response from participants and help in the overall goal. The paper discusses my research objectives in the next section, followed by some details of my current projects. I conclude the paper with a discussion on my expectations from dissertation and the symposium.

2. OBJECTIVES

Collaborative systems evolve as users contribute a small amount of effort to perform certain tasks which together combine to fulfill larger tasks. For eg., systems like Wikipedia, Open Street Maps require multiple users to add content or add tags, which are small individual tasks, but in all contribute to the development of such large systems. Another great example of crowdsourcing platforms is Amazon Mechanical Turk, where a requester lists down small tasks to be performed by workers that usually take less than a minute to complete, and workers get rewarded with a small amount of money for completing these tasks.

Research for social good usually tries to develop solutions for problems in developing regions such as poverty, literacy, health, etc. Solutions to a large number of these problems lie in collaborative practices. For eg, in order to moderate Government policies and their deployment, a large number of people can continuously provide data to a system when they see news on some deployment. A collaborative system with this input can help ensure correct moderation and doesn't require extraordinary efforts from any member. Similarly in situations like rural information and news distribution, crowds can act as local distribution centers and help expand a system's reach to others around them. Therefore, answers to many development problems can be provided by building dedicated collaborative systems.

My research focus is to design principles for building these systems, implement them and optimize them as they grow with data. An important aspect of my research is the management of human factors in collaborative systems. As I need the humans to represent a computational module for this system, I need to take care of both usability as well as optimization of machine to human conversion of the problems. System design should always be usable by the targeted population, and architecture must support conver-

sion of their input into a usable entity. These systems are both unmoderated and semi-moderated. Moderation can be managed in these systems by building a reputation system among the contributors and therefore assigning relevant trust to their contribution.

Through these systems, I also try to study user behavior on different parameters and make an attempt to build inference models. By creating a crowdsourced response platform for a certain task, I try to relate user biases towards problem subjectivity. This is one of the most interesting and challenging part of my research. With a general focus of work around Crowdsourcing systems, I am working on dedicated deployments and am working both on Interface design principles, as well as data based inference models.

3. PROJECTS

Significant projects under my current research work include:

- **WikiTrust Crowd Compute**

WikiTrust [1] is an open-source, on-line reputation system for Wikipedia authors and content. The system builds word based reputation measure for Wikipedia articles and adds a trust value to them. Based on contributor's reputation along with sophisticated algorithms, WikiTrust provides a robust trust management to the content. The process of assigning trust to a document requires a large set of computation on revisions of Wikipedia pages. Every revision associates a set of trust annotations within it which involve it's author's reputation measure. Initially, all this processing was done on central WikiTrust server.

This project builds up a system to crowdsource the computation process in building these trust documents. We have designed an architecture where a centralized server is setup in the cloud and it initiates communication with Wikipedia. For every Wiki page stored in the server database, it keeps knowledge of the latest annotated revision. The server continuously polls Wikipedia to gather recent changes in any of the stored pages. As soon as it finds a change in one of the pages, it creates a task for processing the respective page. Information about the page and its latest annotated revision is packaged into the task.

A client can pickup any such task from the server and that places a lock on the page. The server maintains this processing queue by handling locks and maintaining pointers on latest revisions per page. Once the client gets a task, it can then communicate with both Wikipedia and WikiTrust servers to collect relevant revision and author reputation information. Every client is setup with a set of algorithms which it can work on with this revision data. Once these revisions are processed by the client, it can request to push back the annotated revisions to the server. Once validated, the server updates latest annotated revision for the page and closes the task if revisions have been updated till the current revision of Wikipedia page. If only partial set of pages are processed, the server keeps the task open but releases the lock on this page so as to make this task available for processing.

- **Rank It! Crowdsourced Rankings**

Rank It! is a mobile interface for crowdsourced ranking of digital media, including books, images, movies and music, designed as a vehicle to investigate the size of items to rank, especially if the number of items is large enough to require partial ranking list, and the interaction modality and metaphor. With user-centric development methodology, we finally developed differently styled interfaces for ranking different media on Android platform and identified list lengths of 4 and 5 to be ideal lengths for ranking on mobile interface.

- **Crowdsourced Disaster Response Hosts**

This project builds up on the open source project Sahana Eden [2] and helps create easy hosts for crowdsourced data management operations on a disaster response information management system. Motivated by the fact that there is a large number of volunteers after a disaster who might not be technical experts but can provide some assistance in information management operations, we have created easier interfaces for Sahana Eden which automatically synchronize data with a base deployment and run individually on user machines. The entire process of managing data on the system can be crowdsourced to these small hosts.

4. EXPECTATIONS

As explained in the objectives, my research focus is on developing collaborative systems that can be directed towards applications for social good. I expect my dissertation to lay down well evaluated principles of design and development for such systems. Also, I expect to gain better user understanding with my deployments and their continued usage. Over a long term of implementation, I also expect to develop partial behavioral models for user response under different systems and situations.

The Open Symposium is an ideal venue for my research domain. My learning and my projects are closely related to Wikipedia, which is a primary contributor for this event. The amount of co-relation between my research direction and the Symposium is immense and I want to contribute to this area of this research by actively being a part of the community. I have developed some collaborative systems and am in process of working on more. The OpenSym will serve as an excellent venue to share my experiences with experts in the area. I look forward to getting my research reviewed by peers and experts in the area and collaborating on more projects.

5. REFERENCES

- [1] B. T. Adler, K. Chatterjee, L. de Alfaro, M. Faella, I. Pye, and V. Raman. Assigning trust to wikipedia content. In *Proceedings of the 4th International Symposium on Wikis, WikiSym '08*, pages 26:1–26:12, New York, NY, USA, 2008. ACM.
- [2] M. Careem, C. De Silva, R. De Silva, L. Raschid, and S. Weerawarana. Demonstration of sahana: Free and open source disaster management. In *Proceedings of the 8th Annual International Conference on Digital Government Research: Bridging Disciplines & Domains*, dg.o '07, pages 266–267. Digital Government Society of North America, 2007.