Design Uncertainty in Crowd-Sourcing Systems

Anthony Tomasic, John Zimmerman, Aaron Steinfeld, Yun Huang,
Daisy Yoo, Chaya Hiruncharoenvate, Ellen Ayoob
Carnegie Mellon University
tomasic@cs.cmu.edu

ABSTRACT
Crowd-sourcing social computing systems represent a new material for HCI designers. Crowd sourcing has been successfully applied to many areas, but these systems are difficult to reliably design and prototype. One source of uncertainty during design is the behavior of the crowd at different levels of scale. A second source of uncertainty during design is the crowd's response to differing incentive functions. Using a research through design methodology we have constructed two crowd-sourcing systems where these two sources of uncertainty are central research problems to the design of both systems. This paper discusses these issues with respect to our implementations and discusses a research agenda designed to reduce the uncertainty associated with these design choices.

INTRODUCTION
The emergence of crowd-sourcing social computing systems has fundamentally changed computing. The success of Google's search engine via Page Rank [1] (participant generated cross-referencing), Wikipedia (participant generated content), and reCAPTCHA [2] (participant generated ground truth data for machine learning systems) demonstrates the many different forms this new technology can take and the range of problems it can address. Crowd-sourcing systems provide a new technique for the discovery of new solutions to many kinds of problems.

Perhaps the most striking aspect of crowd-sourcing is the broad applicability of the technique to many different problems and that the application of the technique satisfies different goals for different stakeholders in a system. Table 1 is a list of the goals of designers and users from a sample of different crowd-sourcing systems. This broad range of applications makes the design process more difficult since there are few generally applicable assumptions at the start of the process.

All crowd-source systems must face two barriers to large scale adoption. Crowd-sourcing systems behave differently at small scale than at a large scale. For several systems in Table 1, information generated by the crowd is provided to the crowd. This feedback loop generally requires a minimum amount of activity to be self-perpetuating. Thus, the designer is faced with the problem of bootstrapping the crowd until it is sufficiently large. The second design problem is the incentive structure of the system that is used to keep participants returning to the system to insure long term viability. The uncertainty in design here is more due to a lack of empirical evidence of the effectiveness of different incentive techniques.

In the next section of the paper we briefly describe two prototype research systems that use crowd-sourcing in distinct ways. The last section outlines a research agenda for attacking design uncertainty.

Table 1: Primary Value for Designers and Users of a Sample of Crowd-Sourced Systems

System	Designers	Users
Google	Find information	Find information (user) / Better, more popular, website (administrator)
Facebook	Connect people	Communicate with friends
Wikipedia	Encyclopedia	Find information (reader) / Increase knowledge (editor)
reCaptcha	Optical character recognition	Gain access (user) / Block spammers (admin)
Tiramisu	Automatic Vehicle Location	Save time, improve transit service
ESP Game	Image search	Fun
Inspiration	Concept construction	Fun

TIRAMISU
Tiramisu [3] is a transit information system where commuters share GPS traces while they ride the bus. Tiramisu processes incoming traces from the crowd of commuters and generates historical and real-time arrival time predictions for buses. In effect, the problem of automatic vehicle location (AVL) of buses is solved by crowd-sourcing participants with GPS enabled smart phones. Thus, a multi-million dollar AVL system is replaced with a crowd of smart phones and Tiramisu

CHI 2011, May 7–12, 2011, Vancouver, BC, Canada.

software. In addition to bus location, Tiramisu enables the crowd to perform participatory sensing about the fullness of buses and problem reports (e.g., broken seat, late bus, etc.)

Figure 1: Bus arrival information interface and bus location recording interface for Tiramisu.

Figure 1 is two screen shots of the bus arrival information interface and the bus location recording interface of Tiramisu. The left screen shot of an interface shows the buses expected to arrive at Forbes Ave at Morewood based on schedule information, historical information, or real-time information. The right screen shot of an interface shows the user recording the location of the 61B bus.

A crowd-sourcing application that simply shares bus location information suffers from the bootstrapping problem. The application is only useful if a critical fraction of bus riders record their GPS location. To solve the bootstrapping problem, Tiramisu leverages existing open community transit schedules [4]. Thus, even if no user records bus locations, the application provides useful information. This solution generalizes to a design rule: *to avoid bootstrapping problems, extend an existing non-crowd-sourced application with crowd-sourced based features.*

Tiramisu has been shown to be feasible through a field trial, but the question of viability is still open. The viability of Tiramisu is governed by, in part, the incentive function used to motivate users to record bus location traces. Since Tiramisu provides historical information, one incentive of users is the self-interest to simply improve the existing historical model for a particular bus. Since the user will probably ride the same bus in the future, an accurate historical model is beneficial to a user because it is presumably more accurate than the published schedule.

INSPIRATION

Inspiration (http://musket.isri.cmu.edu) is a human-computation game that acquires concept sets from game participants. The style of the game is similar to the ESP game [5].

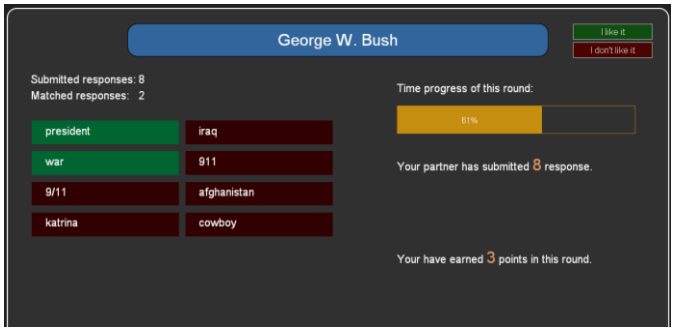

Figure 2: Screenshot of game-play of Inspiration.

Inspiration matches two anonymous participants over the internet. The participants are presented with a series of challenges. A challenge is an input concept and a set of guesses by a player. To win points, a player must guess the same related concepts (output) that the other participant is guessing. Figure 2 is a screen shot of game-play where two participants are guessing concepts for "George W. Bush". They have successfully guessed the same related concepts of "president" and "war". The player shown has six additional guesses of related concepts that have not (yet) matched.

Inspiration has two bootstrapping problems. The first bootstrapping problem comes from the anonymous pairing of participants. With few users, a participant may wish to play but have no partner. This problem is avoided by adding robot participants to the system that act like human participants by replaying previous player guess sequences. The second bootstrapping problem comes from the concepts presented to players during game play. The initial version of the game has a set of basic concepts that have been added by game developers, but this set is of limited size. To add new concepts into the system, Inspiration solves this bootstrap problem by permitting players to add new challenges, but the right to add new challenges must be earned by playing the game. Thus bootstrapping of concepts is linked to incentives to play the game.

Inspiration uses a leader board to broadcast the identity of the highest scoring players of all players. Leader boards are a useful method to motivate people to play but leader boards also have problems, since they discourage new users. As an additional incentive, Inspiration gives access to a user who had added a challenge all the guesses for any player of that challenge. In effect, the user can leverage Inspiration to gather information about concepts that they add to the game. This feedback is the only way that a user can see another player's guesses. This incentive function is another form of providing privileged information or access

as a reward. The feasibility of Inspiration has been demonstrated through a field trial, but the viability of the game is still an open question.

RESEARCH VISION

The research community has done an excellent job of understanding design methods to motivate users to participate in crowd sourced systems [6]. In our opinion there is relatively low uncertainty in designing a crowd-source system that primarily encourages participation compared to other types of crowd-sourced systems. Many design techniques are available to encourage users and increase participation. However, the design choice of systems that encourage participation in general is not straight forward.

For example, the principle value difference between Tiramisu and Inspiration is between time efficiency and entertainment. But we can use entertainment style incentives to improve Tiramisu – we could design Tiramisu as an iPhone application that "shoots" a bus to record it – and then provide a leader board of the best shooter. We can also use utilitarian incentives to improve Inspiration by providing monetary compensation to users of Inspiration. But both of these design choices do not appear to be good ones, but an effective design rule that illuminates why these choices are bad is difficult to determine. Such a design rule would reduce the uncertainty in building crowd-sourcing systems. More research on this issue is needed.

Some sites require an *exchange* of information as a condition of continued access. On-line dating sites require users to submit a profile description before the user can access the repository of profiles. This exchange incentive function insures that the number of profiles on the site grows to a steady state of profiles. Similarly, job salary sites (e.g., http://www.glassdoor.com) require a user to register and provide information about the user's current job after a brief trial period. This incentive function effectively eliminates the free-rider problem. But the circumstances that an exchange model works better or worse than a pure participation model are not clear. For example, an application of the exchange model to Facebook could require users to periodically write to their wall (or add a new friend). This rule might reduce the total number of Facebook users, but it also might increase the loyalty of the users that remain.

Amazon Mechanical Turk has been a very successful system for crowd-sourcing the acquisition of labor in exchange for money. E-Bay, stock exchanges, and gambling lotteries are all systems that effectively crowd-source the exchange of value between users. These systems extract rent (a fee) at the point of exchange. But other crowd-source systems that use a financial incentives function have failed [7]. Again further research is needed to determine the design decision rules for financial incentives.

The above discussion has focused on *primary effects* of an incentive function. However, *secondary effects* also exist, particularly in the impact of brand and perceived quality of service. For example, consider an expert service (such as a lawyer or physician). How is the brand of an expert service impacted if the decisions provided by those experts are simply crowd-sourced to a third party? A patient might pay *more* for a crowd-sourced analysis of a chest x-ray if the crowd consisted of three independent, board certified, radiologists since the patient is effectively gaining the information of a second (and third) medical opinion without the hassle or risk of a second physical x-ray. Additional research is needed in this area.

ACKNOWLEDGMENTS

We thank Rafae Aziz, Jon Chu, Mikhil Dhruv, Andrew Ko, Nitin Seemakurthy, Nikhil Ravi Thiruvengadam, Luis von Ahn, and Jing (Eric) Yang for extended discussions about crowd sourcing

REFERENCES

1. Sergey Brin, Larry Page, "The Anatomy of a Large-Scale Hypertextual Web Search Engine". *Proceedings of the 7th international conference on World Wide Web (WWW)*. Brisbane, Australia, 1998. pp. 107–117. http://dbpubs.stanford.edu:8090/pub/1998-8.
2. Luis von Ahn, Ben Maurer, Colin McMillen, David Abraham and Manuel Blum. "reCAPTCHA: Human-Based Character Recognition via Web Security Measures". Science, September 12, 2008.
3. John Zimmerman, Anthony Tomasic, Daisy Yoo, Chaya Hiruncharoenvate, Charlie Garrod, Rafae Aziz, Nikhil Ravi Thiruvengadam,Yun Huang , Aaron Steinfeld, "Field Trial of Tiramisu: Crowd-Sourcing Bus Arrival Times to Spur Co-Design," Conference on Human Factors in Computing Systems (CHI 2011), Vancouver, B.C., 2011.
4. General Transit Feed Specification. http://code.google.com/transit/spec/transit_feed_specification.html Accessed January 13, 2011
5. Luis von Ahn and Laura Dabbish. General Techniques for Designing Games with a Purpose. Communications of the ACM, August 2008. pp 58-67.
6. Kraut, R. E. and Resnick, R. Encouraging Contributuon to Online Communities. In Kraut, R. E., Resnick, P., Kiesler, S., Riedl, J., Konstan, J. and Chen, Y. (Eds): Designing From Theory: Using the Social Sciences as the Basis for Building Online Communities.
7. A Farewell to Bing Cashback. http://www.bing.com/community/site_blogs/b/search/archive/2010/06/04/a-farewell-to-bing-cashback.aspx Accessed January 13, 2011.