MobileWorks: A Mobile Crowdsourcing Platform for Workers at the Bottom of the Pyramid

Prayag Narula¹, Philipp Gutheim¹, David Rolnitzky¹, Anand Kulkarni², Bjoern Hartmann³

¹School of Information, ²Industrial Engineering and Operational Research, ³Computer Science Division
University of California, Berkeley
{prayag,gutheim,dr}@ischool.berkeley.edu, anandk@berkeley.edu, bjoern@cs.berkeley.edu

Abstract - We present MobileWorks, a mobile phone-based crowdsourcing platform. MobileWorks targets workers in developing countries who live at the bottom of the economic pyramid. This population does not have access to desktop computers, so existing microtask labor markets are inaccessible to them. MobileWorks offers human OCR tasks that can be accomplished on low-end mobile phones; workers access it through their mobile web browser. To address the limited screen resolution available on low-end phones, MobileWorks segments documents into many small pieces, and sends each piece to a different worker. A first pilot study with 10 users over a period of 2 months revealed that it is feasible to do simple OCR tasks using simple Mobile Web based application. We found that on an average the workers do the tasks at 120 tasks per hour. Using single entry the accuracy of workers across the different documents is 89%. We propose a multiple entry solution which increases the theoretical accuracy of the OCR to more than 99%.

I. INTRODUCTION

Amazon’s Mechanical Turk has about 200,000 workers who predominantly reside in the US and India. About 56% of these workers come from the United States. With about 36% of the worker population, Indians represents the second biggest nationality on Mechanical Turk. Similarly, smaller crowdsourcing platforms like Crowdflower give out about 19% of its tasks to Indian workers and about 12% tasks to Pakistani workers[1] Mechanical Turk Workers (“Turkers”) in India are often more educated, earn higher wages, and have a higher standard of living than the average Indian[2] The average Indian Turker is a young, well-educated, male[3]. This might be attributed to the lack of computer access in India. The computer penetration in India in 2004 was only about 0.9%[4]. Moreover, most Turk tasks currently available currently require a high level of English literacy[5]. These barriers currently make most crowdsourcing work inaccessible to the typical Indian Turker.

In contrast, mobile phone penetration in India is very high: about 50%. This is largely due to the plummeting cost of mobile Internet for cellphones in India, currently available for about Rs.50 ($1.10 US) a month in India. In many locations, three-day plans are available from Rs 15 (About $0.33 US). In addition, a large percentage of the phones that are capable of surfing the Web are simple, “candy-bar” style phones. Because of the ubiquity of inexpensive cell phones, even people who earn less than $2.00 US a day have access to the Internet. As a result, the mobile Internet is a good way to send microtasks to people at the bottom of the economic pyramid[1].

In response to this opportunity, we developed MobileWorks, a mobile Web-based microtask platform for developing countries (Figure 1). The goal of MobileWorks is to extend participation in microtask markets to marginalized workers, those that either don’t have access to desktop computers, or don’t know how to use them.

Simple candy-bar phones present a number of challenges to widespread adoption of micro-task work. First, they have very small screen and limited interface capabilities, making more sophisticated tasks difficult. Moreover, the bandwidth that these phones support is low; most use General Packet Radio Service 2nd Generation (GPRS 2G), with a maximum data download rate of 60 kbits/second. In practice, the data rates are often even lower than those operating on 2G due to infrastructure issues. Finally, because these simple phones typically have small screen
sizes and tiny keyboards, ergonomic issues can discourage workers from extended micro-task work sessions on their mobile device. The key research question is whether mobile interfaces can be productive enough to provide livable wages to workers, while being cost-competitive to the task requester.

To investigate these issues, we deployed a mobile phone-based micro work platform for doing human Optical Character Recognition (OCR) tasks. We ran a pilot study with 10 workers in two different cities in India over a period of 30 days. These workers were asked to work from their own cellphones using a standard data plan. We collected data for individual tasks by logging user activities, and conducted interviews with selected members of these groups to understand the user experience of doing microtasks on a candy-bar mobile phone.

II RELATED WORK

A. TxtEagle

Using cellphones to distribute microtask work has been explored before. TxtEagle[6], deployed in Kenya, used SMS text messages to provide micowork to people using candy-bar style phones. However, although using text messaging limits the range of tasks that can be performed, TxtEagle explored a variety of tasks: audio transcription, software localization, citizen journalism, search relevancy, local language translation, new sentiment analysis and market research. Nevertheless, text messaging is limiting, since tasks are limited to those that only require plain-text responses of limited length. As a consequence, this length restriction also limits the efficiency of the workers. Because of these limitations, TxtEagle discontinued many of these tasks and now focuses primarily on market analysis and brand engagement work using text messaging.

By using the mobile Web, we can avoid the limitations of text messaging. Though in the short-term this may prevent access to our service by a small number of users, in the long-term we expect that the decreasing costs of Internet access and growing availability of data plans make the mobile Web a much more promising way of distributing microtasks to a much larger number of people.

B. Samasource

Samasource is a grass-roots organization that seeks to empower marginalized workers by providing work that can be done online. Samasource collaborates with partners in developing countries to establish rural outsourcing centers and provide local residents with microtask work. During the relief efforts after the earthquake in Haiti, Samasource partnered with Inveneo and 1000Jobs/Haiti to establish a computer center in Mirebalais to translate, tag and map messages containing vital information which was collected from citizens via SMS using a mapping system like Ushahidi (Nesbit, 2010). Samasource does not use mobile crowdsourcing; instead they partner with established, local organizations to create computer centers and provide employment opportunities. The workers in these centers are recruited, trained, and managed directly by these partners. Samasource’s model works well for creating outsourcing centers and providing permanent employment in rural areas of Africa and South East Asia. However, there are scalability issues with this strategy, particularly worker training (Thies, 2011). Samasource employs about 900 people in various developing countries. But since this model requires the establishment of physical infrastructure coupled with a large commitment of resources to train workers, costs can be quite high, especially to establish new centers. Since Samasource provides full-time employment, it is more closely related to many rural BPOs, and therefore doesn’t benefit from the scalability advantages that a crowdsourcing solution provides.

C. Improved Mechanical Turk Interface

[5] proposed a new interface for Mechanical Turk that is much more accessible to low-income workers in India. The authors found that using the existing interface and instructions of Mechanical Turk, none of the participants were able to conduct tasks like drawing bounding boxes in images. They recommended a number of design alterations to make these tasks accessible to these workers. The recommendations included localization of language, providing video tutorials, and improving the task interface. These changes made the tasks much more accessible to the workers and they found that using the new interface and instructions the workers were able to do 90% of the tasks correctly.

Such a solution would make the Amazon Mechanical Turk marketplace much more accessible to the users. This allows the workers to use an already mature platform to earn money. However, such a solution requires that workers have access to desktops and are fairly comfortable using them. As mentioned earlier only about 1% of the Indian population have access to computers. Hence, though this solution would work well for people who own computers but is not useful for low-income workers who do not know how to work on computers or do not have access to it.

reCAPTCHA is a revolutionary technology currently being used in digitizing thousands of books. However, reCAPTCHA requires that a high percentage of words from the scanned document are already recognized. Hence, it can be difficult to use it for handwritten text. Moreover, reCAPTCHA is a proprietary technology owned by google currently being used to digitize books as part of the Google Book Project.
However, crowdsourced micro work has been used to produce accurate results for even blurry text. In [8], the authors compared two task flows: a parallel work flow in which different micro workers do the same OCR tasks independently, and an iterative work flow where each subsequent worker is shown answers from the previous workers. The authors reported that using iterative task flow it is possible to achieve 71% accuracy for even very blurry text.

III. HUMAN OCR USING MOBILEWORKS

MobileWorks does human OCR by splitting up documents into small chunks and sending them over data networks to the phones of micro workers in India. The workers type out the text of the document and send it back to the system. MobileWorks then puts these chunks back together and creates a digital version of the scanned document. The system consists of three main components: First, a preprocessing module that takes scanned paper documents and breaks it down to generate small OCR tasks associated with each document. Secondly, a mobile web application that enables workers to perform the OCR. Third, a post-processing stage that reassembles translated pieces and conducts quality control.

A. Preprocessing Documents

Preprocessing of the documents consists of breaking a scanned document into smaller chunks which are easily readable on small screens. This means breaking down a page of text into its individual words. Paper forms which have standard format and broken down into individual fields of those forms. For example, a bank check is only scanned for its individual fields of names, amount, account number and date instead of the standard details like name of the bank and address of the bank etc. (See Figure 2)

Because of the limited screen size of mobile devices, and candy-bar phones in particular, documents have to be chopped into small pieces of one or two words each. Division of document can be done using open source OCR software like tesseract[9]. For the sake of simplicity we used manual processing to divide a document into individual words. It is also possible to extract different words from a document using human computation over a platform such as Mechanical Turk[10].

The information about different chunks of documents (or different words) was stored in a database. This database was then made available to the web application so that user can log in and digitize these words.

B. MobileWorks Web Application

The web application for showing these tasks is written in standard XHTML Mobile Phone profile and WCSS (Wireless Cascading Style Sheet)[11]. This makes the web application accessible on almost any mobile phones including simple candy-bar phones. To maintain compatibility with most phones, no client side processing is done using client side scripting language such as javascript and all the processing is done on the server side which was built using the python language over the Django framework[12].

Since the application should work on small screens of the candy-bar phones, the UI needs to be minimalistic. As soon as a the user logs into the web application, she can start doing the work. The only UI that user sees is the image of the words and the text box to put the word into. The user also sees the total money earned in Indian Rupees. The user types in the words and her earned money goes up instantly. Thus, she gets immediate feedback on her income. This system has an advantage that it gives her immediate feedback on her earnings. Since, the user always knows how much money they are earning this keeps the users motivated to do as much work as fast as they can. However, this also means that the user even gets paid for erroneous entries and might not have any incentive to do quality work. A way to solve this problem is to model per-task payment for the worker on the basis of their accuracy. Hence, the user starts at the middle tier of payments. As her accuracy goes up, her per task payment increases. On the other hand, if her accuracy goes down, her per task payment decreases.

C. Document Post Processing

Different workers digitize the pieces using the Mobileworks Web application and submit it to the server.
These smaller pieces are put together to create a digitized copy of the document. Quality is maintained by checking each worker’s result against another worker’s answer. If the two answers are different, the task is sent to a third worker and compared again until there is a consistent result between two results. Once this agreement occurs, the answer is considered correct and is the accuracy factor of the users who entered the correct answers go up and the quality factor of the users who gave contradictory answers goes down.

IV. PILOT STUDY

For the pilot study, we recruited 10 workers from Mumbai and a semi-rural town near New Delhi. Our test set of data entry tasks consisted of 1000 auto-generated images in a font similar to handwriting. Additionally, we included 70 images with handwritten text that we considered as hard to infer. The pilot was carried out over a period of two months and a set of 1,070 images were digitized by the group of workers. Each image was digitized three times by each workers and created a total set of 30,000 completed tasks. Among the users were 3 housewives, 5 students and 2 full-time workers. Most of the users were not active computer users. Only 2 students had any experience working on computers out of which 1 person was an active computer user.

The interface was minimal as shown in the Figure 2. The left-hand image illustrates the application login. The user enters a user name and password and can immediately start working. The image on the right shows what a typical task looks like for a user. In this case, the worker transcribes the picture and then can easily move onto the next one by clicking the “Next” button. The cumulative amount each worker has earned is recorded after each task is complete and displayed.

A. Accuracy

For test purposes, we had 5 workers access the mobile works platform via their cell phone and 5 workers access it via a computer. The analysis of a set of 5,000 samples revealed high accuracy rates of approximately 98% from both worker groups for the auto-generated images (see figure 3), whereas we did not find a statistically significant difference in performance for any of the worker group (cell phone or computer). We consider a submitted result as correct when the sequence of characters in the submission was equal to the sequence on the image. Hence, we disregarded if a worker included spaces in between characters or submitted the results all in upper case/lower case. We account these differences due to the lack of guidelines for the workers how to deal with ambiguity regarding the format of the submission.

For the 70 “hard” handwriting tasks, we took a sample of 700 results (every worker completed each handwriting once). All workers completed these images through their cell phone. We found that the overall accuracy of the workers without considering multiple entry error detection was about 89% comparing it to a gold standard created from our own responses. Though we did not explicitly test the accuracy using quality control entry checks, we can make an estimation using basic probability principals. The majority of the errors consisted of small misspellings. In order to test the accuracy of the workers’ results, we chose several particularly difficult examples (see figure 4).

For some tasks however, most of the workers submitted incorrect answers which suggests a high covariance between wrong answers and specific tasks. Based on the analysis, we observed that once the difficulty of a task passes a certain threshold, almost all answers were submitted wrong. However, in case we considered an image “less hard” almost all workers submitted correct results. Based on the analysis, we observed that once the difficulty of a task passes a certain threshold, almost all answers were submitted wrong. However, in case we considered an image “less hard” almost all workers submitted correct results. On that basis, we can assume accuracy rates of about 95% - 99% when using double entry for images with a degree of hardness that is below the described threshold. For images above this threshold, we can assume lower accuracy rates that are less likely to improve with double or triple entry. In order to flag hard tasks, we suggest to include a “this image is too hard” option in the user interface. We can conclude that when using multiple entry it is possible to provide very high quality crowdsourced human OCR work up until a certain threshold. Additionally, guidelines that define the format of the submission for the workers are helpful to reduce the number of dissimilar results.

Figure 4: Example of data entry tasks from the 1000 auto-generated images

B. Efficiency

In our pilot we compared the efficiency of a worker on a computer and on a cell phone. Figure 3 illustrates how sub-
mitted results are distributed over time for both computer and cell phone workers. The box plot shows a distribution for computer workers that is centered around 9 seconds (median) with a comparably small interquantile range indicating that 50% of the data entry tasks were completed within around 7-16 seconds. The analysis of the distribution for cell phone workers shows a distribution that with high variances between 3-92 seconds and larger interquartile range of 22 to 41 with a median at 30 seconds. Besides actual delays in data entry, we believe that network connectivity problems for cell phone workers account for the heavy tail on both graphs.

Based on the graph analysis, we found out that a worker on a computer were approximately three times faster than a worker on a cell phone. Additionally, cell phone workers were able to complete around 120 tasks per hour. In our survey we found out that in their regular jobs these workers were earning 20 - 25 Indian Rupees an hour (USD 0.55). On the basis of this efficiency, the workers should be paid about 0.18 to 0.20 Indian Rupees per task to match their regular wages. We predict that the efficiency of the workers would increase as they become more expert at doing these tasks on their mobile phones. Moreover with improvement in wireless infrastructure, the data transfer rate would increase, decreasing the Round Trip Time of each task increasing the efficiency.

C. Survey
We conducted a survey at the end of the experiment to establish the usability of the system. Overall, the response of the workers was positive. When asked to rate the usability of the system on a Likert Scale, 100% of the users rated the usability above 4. Moreover, 100% of the users said that they were more than likely to recommend the system to their friends and family. The users commented that the biggest advantage of the system was that the work could be done anytime of the day: “I could do the work and earn money while traveling to my regular job or even while watching television” one of the participants in the study claimed. We also conducted user interviews with the selected group of 4 workers from the 8 responding users. These users were selected to represent the two major groups of user house-wives and students. We conducted the interviews over the telephone and they were semi-structured.

A. Context of use
One of the major advantages according to the users was the flexibility of working hours. “You can do these tasks anytime, I used to do them while watching television.” , one user said. Many users had set target for themselves while working on the system. One of the housewives mentioned that she “wanted to work for 3-4 hours and earn 25 Rupees/hour”. Another participant mentioned that she “tried to earn Rs.60 - Rs. 80 per day”. All the users considered this as a source of part-time income. One of the participants mentioned “I don’t think anyone would use this as full-time. However, if I get paid more per task I would increase the number of hours I spend on it.” Another said “I would like to use this system full time but I feel that I don’t need to since the working hours are so flexible.” Many users identified travelling as the time when they would do a task. One participant mentioned that she does the tasks while “taking a walk”. Another participant mentioned that she could do more tasks on Saturday night because “kids don’t have to go to school of Sunday so I don’t have to put them to sleep.” The time spent on the system also differed by religious belief. One of the participant claimed that she could do more work on Thursday since she doesn’t have to do laundry on Thursday since it’s bad luck to do laundry on Thursday.

B. Motivation
Money was the biggest motivation for doing the tasks. To gauge the motivation of the participants, we gave the participants a chance to participate in the 2nd round of pilot. We also asked them if they would be willing to earn less money per task if we help them develop their skills while doing these tasks and provide them certification of these skills. The 2 students participants were enthusiastic about this change. One of the participant said he would be interested if the tasks can be used to increase his “general knowledge”. “I would like to learn more about global

V. SURVEY AND USER INTERVIEWS

Figure 5: Example of a data entry tasks from the hard handwritten images
warming. Can you device tasks that give me more knowledge about this.” One of the participants mentioned that her typing speed on cellphones has gone up because of the pilot. “This is a big advantage. Whatever improve my skills is useful.” The other 2 participants were not happy with the changes. “I would just like to earn more money.” A participant proclaimed.

D. User Interface

The system was universally considered to be easy to use, even by participant who did not know how to use a computer. Most of the participant were shown how to open a browser and reach the website by someone else. But they did not need any instructions on how to do the tasks. “Students, housewives, office workers, anyone can do use the system.” One participant mentioned that she “never got bored with the tasks as I expected be.” In this case, it seems that the user gamified the system for themselves. As mentioned before some users had set target for themselves for the number of tasks they wanted to do per hour. They were constant trying to beat the clock on the number of tasks and hence were not bored even by repetitive tasks. Some participants were so engaged with the system that they felt “distressed” when the system was down. Because of the service provider, two participants had connectivity issues with the web application. “This was distressful. I am going to change my provider” said the participant. Hence, since the efficiency is such an important issue, the participants were sensitive to even slight change in the connectivity.

E. Impact

The users seemed positively impacted by the system. A recurring theme in the user interviews was the participants wanting to buy computers to improve their efficiency. One participant said “I am ready to buy and to learn to use a computer if I can earn more money per task. The computers can be bought on Easy Monthly Installments which I can pay by making money off of the system”.

Most participants wanted to recruit their friends into the system. One of the participants said “One of our neighbours is poor and unemployed. I have told him about the system and he would really like to work.” On being asked if they would recommend our system to anyone the response was a yes in 100% of the cases.

VI. CONCLUSION AND FUTURE WORK

MobileWorks is a mobile Web based crowdsourcing platform for human OCR. Future work will explore the possibility of doing other types of tasks, such as audio transcription, same language subtitling and local language translation. We would also like to explore the possibility of providing a real-time response system for microwork by creating a priority cue, in which organizations could expedite their tasks for an extra cost, and workers could choose to complete them for additional incentives. Since workers are using mobile phones we can send an SMS to the user
telling them about the presence of high value tasks. Thus using SMS workers can be recruited to do tasks on run-

time.

References