

Paid Crowdsourcing as a Vehicle for Global Development

William Thies
Microsoft Research India
thies@microsoft.com

Aishwarya Ratan
Microsoft Research India
aratan@microsoft.com

James Davis
UC Santa Cruz
davis@cs.ucsc.edu

ABSTRACT

By connecting remote workers to a global marketplace, paid crowdsourcing has the potential to improve earnings and livelihoods in poor communities around the world. However, there is a long way to go before realizing this potential. To date, most workers on microtasking platforms come from relatively well-off backgrounds, and there has been limited impact on low-income individuals.

In this position paper, we outline a research agenda to extend the benefits of informal, paid microtasking to low-income workers in developing countries. This goal will require research along multiple fronts, spanning the crowdsourcing platforms themselves, their impact upon users' livelihoods, and their scalability to large populations. While there are many challenges to overcome, the rewards are great. We believe that a new focus on low-income workers is critically important to unlock the potential scale and impact of paid crowdsourcing platforms.

INTRODUCTION

There are reasons to believe that paid crowdsourcing¹ could be of particular benefit to low-income workers in developing countries. Unlike most employment opportunities, online marketplaces such as Amazon Mechanical Turk (MTurk) do not require geographic co-location between employer and employee. There is no formal contract needed to commence work; the only criterion for employment is the ability to complete the task at hand. In addition, the working hours are completely flexible, allowing workers to earn money whenever they have extra time. Workers with access to a mobile computer may even be able to earn money during their daily commute, or when they are idle during other jobs (e.g., a driver could work while waiting for a client, or a shopkeeper could work while the store is empty). For all of these reasons, one would expect online microtasking services to lower the barrier to entry for employment in low-income settings, potentially boosting the socio-economic standing of otherwise disadvantaged populations.

¹In this paper, we consider *paid crowdsourcing* to be the provision of small tasks in exchange for monetary payment via websites such as Amazon Mechanical Turk, CloudCrowd, ShortTask, etc. While we are agnostic as to the platform used, we restrict our attention to the informal setting in which workers choose to work on one task at a time, rather than having a long-term contract with a company.

But has paid crowdsourcing delivered on its potential to benefit those at the bottom of the economic pyramid? The answer today is both “yes” and “no”. On the positive side, it has been documented that over a third of workers on MTurk are based in India [4, 5], suggesting that the service can make inroads in a developing-country context. And in our own survey of 200 Indian Turkers [5], we uncovered several cases in which MTurk had fundamentally changed the livelihood of our respondents. For example, a 26-year old college graduate from Kolkata, India, describes how he came to earn \$1860 per year on Mechanical Turk:

I'm from a middle class family. After completing my degree I looked for job everywhere but failed. But when I found MTurk, it changed my life. It helped me a lot.

However, despite such anecdotes, the impact of paid crowdsourcing in developing countries remains limited. While many Turkers are based in India, in our survey we found that these workers have a relatively high socio-economic standing: 80% have a Bachelor's degree or higher, 92% have a PC and Internet connection in their home, and their median annual income is \$2700. This is substantially better than the position of the average Indian. Only 6% of India's workforce has a Bachelor's degree, only 6% of households have a computer and Internet access, and average per capita income is \$1100 per year [5]. To have an impact on poor communities, systems such as Mechanical Turk will need to be accessible to those with lower formal education (secondary and higher secondary schooling) and only basic skills with computers and the English language. Today, the benefits of paid crowdsourcing are largely out of reach for this population.

In this paper, we outline an agenda to shape the evolution of paid crowdsourcing into a vehicle for socio-economic development. Our agenda is heavily influenced by our recent study, which examined the usability of Mechanical Turk for low-income workers in India [5]. The study found that while there are tasks on MTurk that are not beyond the cognitive capabilities of low-income workers, the language and user interface represent significant barriers for this population to earn money on MTurk. By simplifying the user interface and task instructions, and translating all content into the local language, we demonstrated that low-income workers can achieve much higher rates of task completion. Still, it remains challenging for such workers to reliably earn money on MTurk, due to costs of accessing computers and other barriers. Many questions also remain in scaling up our re-

sults from a laboratory environment to impact large numbers of users across India.

We structure the remainder of this paper around three complementary research directions: improving the crowdsourcing platforms themselves, impacting the livelihoods of end users, and enabling such impact to scale to large numbers of people. We believe that progress along all these fronts will be needed for paid crowdsourcing to fulfill its potential as an accelerator of global development.

IMPROVING THE PLATFORM

The first category of challenges pertains to the crowdsourcing platform itself – that is, the website (such as Mechanical Turk) on which jobs are requested and completed.

Enhancing Usability

One of the principal barriers that prevents low-income workers from utilizing sites such as Mechanical Turk is the complexity of the user interface. As detailed in our prior work [5], even tasks as simple as image labeling (drawing a bounding box around a given object) require complex navigation and sequencing that frequently confuse non-expert computer users. Moreover, the instructions for completing tasks are written in complex and ad-hoc language, making them incomprehensible for many first-time workers. For the task of image labeling, we demonstrated that redesigning the user interface and simplifying the instructions (coupled with translation into the local language) can boost task completion from 0 to 66% for low-income users in India. We believe that our design recommendations can be generalized and extended to encompass a broad array of tasks on MTurk.

Allowing Self-Evaluation

In addition to improving the tasks themselves, crowdsourcing platforms need to provide users with better tools to understand their own productivity and potential on the platform. For example, prior to accepting any task, users should be able to evaluate their capability to perform a task by attempting a prior instance and checking their answer against a previously-accepted input. For some tasks (such as drawing a bounding box), the system can also automatically check a worker's input for known training examples. In addition, users should be able to visualize their hourly earnings over time, compare those earnings with peers, and search for new tasks that are likely to be a good match for them, given their history of work.

Providing Authoring Tools

One factor that contributes to the ad-hoc task instructions on MTurk is that each requester is asked to formulate their own instructions from scratch. This is especially difficult in cases where the task requires interaction with the user; for example, for users to draw a bounding box on the screen, requesters need to author an embedded Javascript or Flash plugin that relays results to a third-party site. The resulting pages are heterogeneous, non-standardized, and difficult to navigate and understand.

We believe that it would be useful to develop a set of minimal primitives, analogous to a “human instruction set”, that are integrated into the platform and can be assembled by requesters in order to accomplish a complex task. For example, primitive display and annotation of images should be standardized yet extensible plugins that are supported by the platform, as should input and output of text, audio, and references to websites. Each of these interface widgets can be carefully designed for usability, and instructions for using them can be translated into local languages.

Given an instruction set, it becomes an interesting research challenge to decompose complex tasks into a fine-grained sequence of instructions. It will be valuable to develop semi-automatic tools that can map high-level tasks into the primitives available on a given platform, while being conscious of cost, accuracy, and other factors.

Regulating Requesters

One of the most frustrating aspects of MTurk for users is the erratic and unpredictable behavior of those who post jobs. Currently, workers are given no guarantees as to the timeframe for payment, and are frequently confused regarding the exact criterion by which their work is evaluated. Requesters should be required to pay on a fixed schedule, and should be regulated to be more transparent in evaluating results. Following online systems such as eBay, it might make sense to give experienced workers the benefit of the doubt, paying up front and only retracting payment if a problem is detected. Such an ecosystem would also benefit from a reputation system for requesters and workers as well, in order to evaluate their reliability and diligence and to hold them accountable for each transaction. This accountability is especially important given the absence of labor laws to govern such marketplaces.

IMPACTING LIVELIHOODS

Even if one builds the perfect website, additional research is needed to enable these improvements to translate to improved earnings and livelihoods for low-income workers in developing countries.

Lowering the Cost of Access

As low-income workers are unlikely to have a computer at home, they will need to establish an alternative and low-cost means for performing work online. In our study, the median earning rate for low-income workers on a bounding box task (during their first attempt) was \$0.41 per hour [5], which is comparable to the hourly charges at many Internet cafes in India. Thus, it is possible for the costs of computer and Internet access to significantly offset or even outweigh the earnings accrued from the site.

One potential avenue of research is to explore which (if any) tasks can be completed from a mobile phone, as they are much more common than computers in the developing world. However, phone ownership in the low-income segment is generally limited to basic handsets or feature phones, which may have graphics and video but are lacking a touch screen. This makes it difficult to perform graphical tasks (such as

drawing bounding boxes). While low-end handsets could be used for text-based tasks, in our experience textual tasks were much more difficult for respondents, not least because they often require subtle manipulation of the English language [5].

Perhaps one approach to saving costs could be to utilize a more expensive feature-rich phone, but to share it amongst many workers in an urban or rural community. This could redefine the role of the “village phone lady”, who once made house visits to enable residents to make voice calls [6]; in the future, such visits could perhaps be associated with earning opportunities.

Another barrier that prevents many people from earning money on sites such as Mechanical Turk is that a bank account is needed to receive payment. Formal banking remains out of reach for many poor communities, making this a difficult logistical hurdle. We believe it will be important to explore alternate means of payment, for example, by crediting a worker’s mobile phone with prepaid voice minutes. It could also be valuable to establish human intermediaries who can receive payment on behalf of a worker; for example, the owner of an Internet cafe could receive payment into his bank account and offer cash payouts to users who work from the premises.

Improving Earnings Stability

In low-income settings, one of the primary determinants of household welfare is not only how much money is earned, but also *when* that money is available. For households that are financially vulnerable, a reliable flow of wages offers significant consumption benefits over a volatile stream of high but unpredictable payouts [1]. Unfortunately, sites such as MTurk currently fall into the latter category. Every day, a worker is at the mercy of the system and the set of jobs that happen to be available. Even jobs that a worker has finished do not have a fixed schedule for payment, making it even more difficult to predict and smooth one’s income over time.

The desire for earnings stability is one of the primary benefits of a BPO (business process outsourcing) model, in which workers are employed full time and receive a regular salary, rather than being compensated on a task-by-task basis. While we believe that platforms such as MTurk may be more suitable for supplemental income generation as opposed to full-time employment, even in this scenario it will be important for workers to predict and stabilize their earnings over time. Technology could address part of this challenge, by tracking the availability of various tasks and allowing workers to anticipate (or perhaps even reserve) work of a given kind. Human or institutional intermediaries, like Kiva for instance, could also stabilize a worker’s earnings by offering short-term loans and savings against a worker’s record, thereby compensating for the varying availability of tasks suitable for a given worker.

Enabling Professional Growth

Paid crowdsourcing platforms offer the opportunity for workers to complete increasingly complex tasks, thereby increas-

ing their earnings and potentially contributing to their career advancement. This is unusual for low-income workers, who often are given a very specialized task with limited opportunities for advancement. For example, even contract employment at a BPO would likely require a worker to specialize in a single task with limited flexibility for trying new things.

However, more tools are needed to ensure that professional growth opportunities are recognized and exploited by workers on microtasking websites. Firstly, it would help to devise a recognition and reputation system for workers, rewarding them with a virtual badge (for example) whenever they do certain categories of tasks. Such rewards are commonplace in other crowdsourcing systems, for example, the elaborate sequence of badges on MathOverflow.net and StackOverflow.net. Other systems leverage reputation systems to reward positive contributions by members; examples include eBay, Amazon, and Yelp. A reputation system would also benefit requesters, as they could judge the accuracy and reliability of a worker based on feedback left by prior requesters.

It will also be important for sites to provide explicit guidance on how workers can advance their long-term growth. For example, sites could illustrate exemplary “career paths” which would help workers acquire complementary skills and apply them towards increasingly high-level tasks. Sites could also integrate computer learning software that tracks workers’ progress and suggests personalized exercises for them to improve their skills. It could even be possible to utilize a worker’s accomplishments on the site to produce a certification that they have a given skill; for example, a worker that successfully wrote several reviews or blog posts could use this as evidence of their English skills when they apply for jobs in the real world.

ENABLING SCALABILITY

Even if we can learn how to positively impact the livelihoods of low-income workers within our reach, it is a different matter entirely to create an environment where the system can spread virally and grow to impact large numbers of people. In this section we outline some steps that will be needed to achieve scale.

Understanding and Increasing the Market Size

It has been estimated that Mechanical Turk has a payout of \$2,000 per day [3]. While this represents a significant number of tasks, it pales in comparison with the 2.6 billion people living on less than \$2 per day [2]. Does there exist a scalable source of tasks that can enable paid crowdsourcing to make a significant dent in the face of global poverty? We do not know the answer to this question, but ultimately it may prove to be the most pivotal question in this research agenda.

One challenge facing the informal crowdsourcing market is that it seems best suited for “odd jobs” that have only a limited number of instances. Once a given task appears in very large numbers, it becomes profitable for a company to specialize in catering to those customers alone – as is frequently

done for transcription, translation, data entry, business listings verification, and other services. If paid crowdsourcing positions itself primarily as a tool for unusual tasks, or requests from individuals and small companies who do not have the volume needed to engage with a specialized firm, it is not clear if it will ever have the volume needed to impact developing regions at scale.

One opportunity in this vein is to explore if we can increase the number of requesters by providing more intuitive authoring tools, as described earlier. Moreover, it will prove critical to develop “killer apps” that leverage the unique capabilities of paid crowdsourcing – including low latency, low cost, highly individualized instructions and tasks, the ability to reach a large and diverse set of workers, and the relative anonymity of requesters – to distinguish it from more formal business contracts and increase the overall paying capacity of the system.

Fostering Peer-to-Peer Skills Transfer

Grassroots organizations such as Samasource have demonstrated that training programs can improve the skills of low-income workers in completing tasks online [9]. But how can such training be implemented at scale? To this end, we think that it will be vital to foster peer-to-peer skills transfer amongst workers in online microtasking systems. In prior work, we observed how low-income workers can acquire basic IT skills by collectively experimenting with a publicly-available computer in an office environment [8]. One of the first mechanisms for demystifying the technology and gaining the confidence needed to experiment by oneself is to look over the shoulder of a more experienced peer in an informal environment.

In the context of paid crowdsourcing, perhaps similar skill transfers could be nurtured by assigning mentors and mentees in the system. An apprentice worker could observe a more experienced peer, perhaps via a screencast as a task is completed. Such interactions could also take place in a physical space, for example, by designating a collaborative “Turking Hour” in a local Internet cafe. The reputation systems suggested earlier could also play a role in mentorship and skills transfer, as they help to provide role models and also provide an avenue for providing credit (reputation points) for offering help to a peer.

Building Awareness

While Mechanical Turk has received a lot of attention in our research community, it remains somewhat obscure to the general population, especially in developing regions. Our hypothesis is that there are a large number of potential Turkers – for example, the 2.4 million unemployed college graduates in India [7] – who are completely unaware of its existence. It represents an interesting challenge to understand the perceptions of online microtasking sites and ways in which any misconceptions could be addressed. For example, Mechanical Turk is relatively unique as a site that requests the details of your bank account and then legitimately deposits money into that account; it could easily be perceived as a

scam in many communities. Simple awareness campaigns and tutoring programs at colleges or work placement agencies could offer large returns in publicizing Mechanical Turk to low-income communities.

CONCLUSIONS

We believe that systems such as MTurk have the potential for much broader scale and impact than they have achieved to date. We framed this paper in terms of the potential impact on low-income workers in developing countries. However, it is important to recognize that such impact will also benefit other stakeholders in the ecosystem; for example, requesters could find lower-cost labor, platform owners could benefit from larger volumes, and the crowdsourcing paradigm as a whole could scale up to address much grander problems.

To achieve these goals, we believe that progress is needed on multiple fronts, spanning the crowdsourcing platforms themselves, their impact on users’ livelihoods, and their scalability to large populations. We are eager to advance this agenda.

REFERENCES

1. A. Banerjee and E. Duflo. The economic lives of the poor. *Journal of Economic Perspectives*, 21(2):141–167, 2007.
2. S. Chen and M. Ravallion. The Developing World Is Poorer Than We Thought, But No Less Successful in the Fight against Poverty. Policy Research Working Paper 4703, The World Bank, Development Research Group, Aug 2008.
3. P. Ipeirotis. Mechanical turk: Profitable or not? <http://behind-the-enemy-lines.blogspot.com/2009/03/mechanical-turk-profitable-or-not.html>, Mar 2009.
4. P. Ipeirotis. Demographics of Mechanical Turk. NYU Working Paper No. CEDER-10-01, Mar 2010.
5. S. Khanna, A. Ratan, J. Davis, and W. Thies. Evaluating and Improving the Usability of Mechanical Turk for Low-Income Workers in India. In *ACM Symposium on Computing for Development*, 2010.
6. Grameen Foundation. Village phone replication manual. <http://www.infodev.org/en/Document.14.pdf>, Oct 2005.
7. TeamLease Services. Indian Labour Report 2007: The Youth Unemployability Crisis. http://www.teamlease.com/images/reports/Teamlease_LabourReport_2007.pdf, 2007.
8. A. L. Ratan, S. Satpathy, L. Zia, K. Toyama, S. Blagsvedt, U. S. Pawar, and T. Subramaniam. Kelsa+: Digital Literacy for Low-Income Office Workers. In *Information and Communication Technologies and Development*, 2009.
9. Samasource website. <http://www.samasource.org/>.