mClerk: Enabling Mobile Crowdsourcing in Developing Regions

Aakar Gupta*, William Thies#, Edward Cutrell#, Ravin Balakrishnan*

*University of Toronto
#Microsoft Research India
Paid Crowdsourcing has yet to Deliver on its Potential in Developing Regions

Prior efforts either have middle-class workers.. or see barriers to scale in low-income contexts
Overcoming the Barriers to Scalable Crowdsourcing in Dev. Regions

Lack of access to computers & Internet

Limited skills for tasks

Lack of payment mechanisms

Low awareness
Overcoming the Barriers to Scalable Crowdsourcing in Dev. Regions

Lack of access to computers & Internet

Limited skills for tasks

Lack of payment mechanisms

Low awareness

1. Send visual tasks via SMS
1. Send Visual Tasks via SMS

Nokia’s Smart Messaging

- Binary Images
- 74x28 pixels
- Same cost as 3 SMSs!
Overcoming the Barriers to Scalable Crowdsourcing in Dev. Regions

Lack of access to computers & Internet

Limited skills for tasks

Lack of payment mechanisms

Low awareness

1. Send visual tasks via SMS

2. Local-language digitization
2. Local-Language Digitization

**Goal:** Digitize paper documents in local language

(i) Document Segmentation

(ii) Crowdsourced Digitization

(iii) Response Verification
(i) Document Segmentation

Clean, Segment, Binarize, Resize
(ii) Crowdsourced Digitization

1. Send worker a word image via picture SMS

2. Worker replies with text SMS

   - But local language fonts difficult or unsupported!
   - Solution: back and forth transliteration

User sends English transliteration

murthy

Server transliterates back

মুর্থি
(iii) Response Verification

Verify agreement of transliterated text

moorthi \equiv murthy

Accepted
Overcoming the Barriers to Scalable Crowdsourcing in Dev. Regions

Lack of access to computers & Internet

Limited skills for tasks

Lack of payment mechanisms

Low awareness

1. Send visual tasks via SMS
2. Local-language digitization
3. Pay with mobile airtime
3. Pay With Mobile Airtime

- Manual recharge via mobile shop
Overcoming the Barriers to Scalable Crowdsourcing in Dev. Regions

Lack of access to computers & Internet
1. Send visual tasks via SMS

Limited skills for tasks
2. Local-language digitization

Lack of payment mechanisms
3. Pay with mobile airtime

Low awareness
4. Incentivize viral spread
4. Incentivize Viral Spread

(i) Referral system

→ Each worker earns 10% of their referrals’ earnings

(ii) Leaderboard messages

(iii) Feedback & motivational messages
Field Study

Will users adopt the system and use it willingly in a real-world setting?
Field Study

• 5 Week study, divided into 2 phases
  • Phase 1 (3 Weeks): Paid INR 0.5 / task (~1 ¢)
  • Phase 2 (2 Weeks): Paid INR 0.2 / task

• Semi-urban location, 4 hours from Bangalore

• Language: Kannada
Results: Diffusion Network

We contacted only 10 users. Five weeks later:

239 Users, 64000 Responses, 25000 Digitized Words
Results: Group Effects

Social interactions drive usage
Qualitative Themes

Time Pass

“I have to wait 20mins for bus. I stand and *do at bus-stop*. I have stopped going to the recharge shop, I get enough.”

Flip side

“We sit at back in class and *message during lecture*.”

“Earlier we [friends] used to message poetry, jokes etc. Now no one does that. Everyone is busy with this.”

Skepticism

“It is like some *code sending*. What do you do using this?”

“*Is it legal?* What’s your profit? I don’t want any trouble.”
Performance: Accuracy

- Fraction of words digitized correctly: 90.1%
- Improvement of Accuracy
Performance: Agreements

Avg. no. of responses per word

Word Length (px) before scaling the length

Screen Width (74px)
Business Sustainability

<table>
<thead>
<tr>
<th></th>
<th>mClerk</th>
<th>Market</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Costs</strong></td>
<td><strong>Phase 1: 2.4 ¢/ word</strong></td>
<td><strong>2 ¢/ word</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Phase 2: 1 ¢/ word</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Accuracy</strong></td>
<td><strong>90%</strong></td>
<td><strong>97%</strong></td>
</tr>
</tbody>
</table>

- With some optimization, could be market viable
  - Partner with telcos to decrease payment overhead
  - Identify more accurate workers to improve accuracy
Conclusions

• **mClerk** enables scalable crowdsourcing in developing regions by:
  1. Sending visual tasks via SMS
  2. Leveraging local language skills
  3. Paying with mobile airtime
  4. Incentivizing viral spread

• Future opportunities in
  - Optimizing accuracy and costs
  - Finding more tasks amenable to picture-SMS
Thank You

To Nithya Sambasivan, Richard T. Guy, James Davis, Indrani Medhi

To all of our users, as well as their..

Collaborators - “Coming back from college in the bus, all of us do messaging and ask each other meanings of the words for fun. One time no one knew so we thought we’ll ask the Kannada lecturer in college and if he does not know that will be fun.. but he knew.”

Competitors - “All my friends have become leaders [at least once]. Now I sleep at 12, so that I can do fast messages at night.”

Family members - “I gave my phone to my wife. She is free at home. She can do more SMS. I take it in evening when I get free with friends”

Contact: aakar@cs.toronto.edu