EFFECTS OF COMMUNITY SIZE AND CONTACT RATE ON SYNCHRONOUS SOCIAL Q&A

Ryen W. White  
Microsoft Research

Matthew Richardson  
Microsoft Research

Yandong Liu  
Carnegie Mellon University
Question Answering (Q&A)

- People have questions, want answers
- Automatic question answering not yet practical
  - Complex questions
  - Opinion questions
  - Knowledge that is not written down
- Solution: get others to help you out…
Social Question Answering

• Also known as “Community Question Answering”
• Ask people for help
  • Send email to mailing list
  • Use web forum
  • Answers service (Yahoo! Answers)
• Downsides:
  • Spams a lot of people (mailing lists)
  • Slow response (web forums)
• Solution: use instant messaging…
Synchronous Social Question Answering

• Users ask a question using instant messaging (IM)
• System forwards question to users likely to know answer
  • Forwards to a few at a time
• Once a willing answerer is found, asker and answerer engage in dialog

• Example systems
  • Aardvark: Deployed on the Web (contacts friends, FoF, etc.)
  • IM-an-Expert: Built and deployed within Microsoft (contacts “experts”)
  • Others in CSCW and CHI community
• This paper uses IM-an-Expert for experiments
  • But similar results are expected for Aardvark or other systems
**IM-an-Expert**

- Facilitates question answering via real-time IM dialog
- All users are “experts” - can ask and answer questions
- IM-an-Expert finds answerers, connects askers to them, and mediates dialog:
  1. Asker poses question via IM or a Web page
  2. IM-an-Expert finds best available answerer
  3. Proxies IM conversation between asker and answerer
IMX Sample Conversation (1 of 2)
IM an Expert
○ Glenn (gmetzler) says: I just a video classes from Dietel (sp??)
○ Glenn (gmetzler) says: two DVDs worth of sessions
○ Glenn (gmetzler) says: returned it a couple of weeks back

Ryen White
○ what is dietel? Is that in the library?

IM an Expert
○ Glenn (gmetzler) says: he's an author

Ryen White
○ oh, I see. I'll take a look. I always forget to use the library :)

IM an Expert
○ Glenn (gmetzler) says: writes many books on programming languages
○ Glenn (gmetzler) says: 😊

Ryen White
○ cool, thanks.

IM an Expert
○ Glenn (gmetzler) says: sure

Ryen White
○ bye

IM an Expert
○ You have finished the conversation. How helpful was the answer? Please provide a rating between 1(not helpful at all) - 5(very helpful):

Ryen White
○ 5

IM an Expert
○ You rated the answer as 5 and I've sent along the text if answerer's window is still open. Thank you. I hope I've been helpful. Bye!

Last message received on 3/31/2011 at 7:37 AM.
Expert Finding

• Sources of user information
  • Implicit
    • Emails sent to internal distribution lists
  • Explicit
    • User-provided keywords and URLs about themselves or their interests

• TF.IDF ranking with temporal decay to balance questions

• Profile page where users can also:
  • Set question limits
  • Tune privacy settings
  • Suspend or disable the service
Dialog Management

- Coordinates flow of messages between askers/answerers
- Contacts top-$k$ experts
  - $k$ is “contact rate”
  - Only asks those who are Available
  - Availability set from calendars and users could set manually
- If answerer doesn’t respond in 60 seconds or types “no”, then contact next user in list
- Once answerer accepts, other invitations are canceled
- All IM dialog logged
Asker and Answerer needs in IMX

- In IM-an-Expert, all users can ask and answer questions

<table>
<thead>
<tr>
<th>Askers want</th>
<th>Answerers want</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low time-to-answer</td>
<td>Few interruptions</td>
</tr>
<tr>
<td>Quality answers</td>
<td>Relevant questions</td>
</tr>
</tbody>
</table>

- Needs are in tension
  - E.g., to get low time to answer may need to interrupt many users
- Investigate effect of community size and contact rate on the extent to which these needs can be satisfied

- This can help us:
  - Understand the impact of these factors in synchronous Q&A
  - Design better social Q&A systems
User Study: Participants

- Participants and Recruitment
  - Redmond-based MSFT employees w/ mailing-list based profiles ≥ 1kb
  - Users required to be available for two-week study duration
  - 402 volunteers in total, users were highly familiar with IM (4.5/5)

- Experimental Groups:
  - 6 groups, varying both community size (n) and contact rate (k)
  - Group members didn’t know about the other groups

<table>
<thead>
<tr>
<th>$k$</th>
<th>Group</th>
<th>$n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>100</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>100</td>
</tr>
</tbody>
</table>
User Study: Methodology

- Study lasted two weeks
1. Asked participants to take a pre-experiment survey
2. Randomly-assigned participants to experimental group
3. Asked participants to visit their profile page and provide keywords and URLs describing interests and expertise
   - Re-indexed daily to capture any profile updates
4. Participants asked to consider using IM-an-Expert as resource for answering questions for study duration
5. Two weeks from start date, study ended and participants completed post-experiment survey
   - 70% of all participants did so
   - Attrition was spread evenly across groups
Findings: General Usage

- Around 50% of participants asked and answered questions in the two-week study (35% of users did both)

- 25% of participants asked/answered half the questions

- Dialogs:
  - Lasted around six minutes
  - Comprised around 10 dialog turns
  - Turns evenly distributed between askers and answerers
Recall from earlier

<table>
<thead>
<tr>
<th>Askers want</th>
<th>Answerers want</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low time-to-answer</td>
<td>Few interruptions</td>
</tr>
<tr>
<td>Quality answers</td>
<td>Relevant questions</td>
</tr>
</tbody>
</table>

- We’re going to look at each of these needs in more detail
### Findings: Asking – Time to Answer

<table>
<thead>
<tr>
<th>Measures</th>
<th>Comm. size (n)</th>
<th>Contact rate (k)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Average time to answer</td>
<td>25</td>
<td>4m 31s</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>3m 58s</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>3m 12s</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>3m 37s</td>
</tr>
</tbody>
</table>

**Key takeaways:**
- Doubling group size leads to 30s reduction in time to answer
- Higher contact rate leads to lower time to answer
Findings: **Asking – Answer Ratings**

- Askers rate answers on a scale from 1-5 at end of dialog

<table>
<thead>
<tr>
<th>Comm. size (n)</th>
<th>Contact rate (k)</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>25</td>
<td>3.33</td>
<td>3.25</td>
</tr>
<tr>
<td>50</td>
<td>3.50</td>
<td>3.32</td>
</tr>
<tr>
<td>100</td>
<td>3.61</td>
<td>3.52</td>
</tr>
</tbody>
</table>

- Key takeaways:
  - Larger group size leads to higher answer ratings (more expertise)
  - Higher contact rate leads to lower answer ratings
    - Less expert answerers may respond before more expert answerers
Findings: **Answering – Interruptions**

- Median number of users interrupted per question = 6

<table>
<thead>
<tr>
<th>Measures</th>
<th>Comm. size (n)</th>
<th>Contact rate (k)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>% of group interrupted</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>14.5%</td>
<td>25.2%</td>
</tr>
<tr>
<td>50</td>
<td>9.6%</td>
<td>17.4%</td>
</tr>
<tr>
<td>100</td>
<td>7.7%</td>
<td>14.2%</td>
</tr>
<tr>
<td>All</td>
<td>10.4%</td>
<td>17.7%</td>
</tr>
<tr>
<td>Felt unbothered by questions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>3.59</td>
<td>3.34</td>
</tr>
<tr>
<td>50</td>
<td>3.92</td>
<td>3.78</td>
</tr>
<tr>
<td>100</td>
<td>4.23</td>
<td>4.09</td>
</tr>
<tr>
<td>All</td>
<td>4.20</td>
<td>4.00</td>
</tr>
</tbody>
</table>

- Key takeaways:
  - Larger community size, less % interrupted + answerers less bothered
  - Higher contact rate, more % community interrupted + more bothered
Findings: **Answering – Relevance**

- **Asked answerers:**
  - Approximately what percentage of questions asked were relevant to you? (0, 1-10%, 11-20%, etc.)
  - $k=2$ more relevant than $k=5$
  - No differences from community size

- **Reasons for not answering:**
  - Question wasn’t relevant to me (~25%)
  - I didn’t know the answer (~50%)
  - Expertise **level** is important in addition to having expertise
Findings: Overall Perceptions

- $k=5$ meant more answers and more timely answers, but ...
- $k=2$ was more useful
  - Users may wait longer for **better** answers, dislike interruptions
Conclusions

- Investigated impact of community size and contact rate on the effectiveness of synchronous social Q&A

- As community size grew, system performance increased

- Contact rate:
  - Askers prefer $k$ with timely answers ($k=5$), high quality answers ($k=2$)
  - Answerers prefer $k$ with relevant questions, few interruptions ($k=2$)

- To satisfy most users, synchronous social Q&A systems should use low contact rates and large communities

- More research is needed on the answer quality vs. timeliness tradeoff e.g., ceiling effects as community size grows