Twitch Crowdsourcing: Crowd Contributions in Short Bursts of Time

Rajan Vaish, Keith Wyngarden, Jingshu Chen, Brandon Cheung, Michael S. Bernstein

twitch@cs.stanford.edu
Crowdsourcing campaigns that cannot motivate participants will fail.

Lack of time is a primary reason experts do not contribute to sites such as Wikipedia.

Even small time requirements can deter users.

Ross, L. and Nisbett, R.E. The person and the situation: Perspectives of social psychology. Mcgraw-Hill, 1991
Downtime? Check your phone.

People spend short bursts of time habitually checking their phones.

Oulasvirta, A., et.al. Interaction in 4-Second Bursts: The Fragmented Nature of Attentional Resources in Mobile HCI. CHI ’05,
Can we mobilize crowd participation in short bursts of time?
Twitch Crowdsourcing

Quick volunteer crowd contributions in short bursts of time via mobile phone
Unlock: A captcha for your phone

Standard slide-to-unlock

Twitch Crowdsourcing
Mobile crowdsourcing

Applications such as mClerk allow the crowd to participate from their phones.

However, these tasks require roughly 30 seconds.

Possible Twitch applications

- Map human activity in the world
- Survey a population
- Micro-tutor a student
- Find the best images on the web
- Extract structure from the web
- Filter the news
Three design goals

- Understand human behavior
- Supplement artificial intelligence
- Capture subjective opinions
Twitch applications

Census

Structuring the Web

Photo Ranking
Twitch applications

Activity?

Census
Mapping human activities

Despite progress in understanding static elements of our physical world - we lack an understanding of human activity.

Census creates a human-centered equivalent of Google Street View, populated with peoples’ activities

Census: how many people around?
Twitch activity tasks
Instant feedback
Activity map

Cambridge, MA

Bay Area, CA
Activity map

World
Twitch applications

Structuring the Web (STW)
Factual queries require structured data

Structured data requires verification

Algorithms guess facts by reading the web, but they are noisy. Could people verify facts on topics they care about?
$ echo "Michael Cunningham is best known as author of The Hours" | ./reverb -q | tr '\t' '\n' | cat -n
1 stdin
2 1
3 Michael Cunningham
4 is best known as
5 author of The Hours
...
12 0.999999997341693
Structuring the Web

Source text: “Stanford has 27 ACM fellows”
Extraction format: {Subject} {Relationship} {Object}
Extraction: {Stanford} {has} {27 ACM fellows}
ReVerb Confidence Score: 0.94

Verifying a web extraction animates a new branch onto the knowledge graph
Source text: “Jr., who died of typhoid two months before his 16th birthday.”

Extraction format: {Subject} {Relationship} {Object}

Extraction: {Jr.} {died of} {typhoid two months}

ReVerb Confidence Score: 0.63

Rejecting the extraction shows how far the user is in the article.
The vision: to be able to eventually contribute towards factual query results.
Twitch applications

Photo Ranking
Computer vision and photography

Matchin: photo ranking

iStockphoto: stock photography


Photo Ranking demo

Rank the photo you like better, every time you unlock your phone
Photo Ranking

Today's Theme: Nature

65% of 401 people
Twitch applications

Census

Structuring the Web

Photo Ranking
Implementation

Twitch unlock screen is implemented as an Android home-screen replacement application. Twitch can replace or complement the user’s password screen.

Evaluation

Thesis: Twitch Crowdsourcing enables tasks that can be completed in 1-2 seconds.
Evaluation metrics

• Quick completion
• Low cognitive load
• Length of active engagement
Twitch deployment

• After initial trials, introduced to Stanford and MIT communities.
• Later deployed as a public app on Google Play, and collected data for a month.
Evaluation

We ran 3 different evaluation studies on:

- Naturalistic Usage
- Speed
- Cognitive Load
Over 1,000 downloads

- Our sample: 82 volunteer users from the first month
- 50% uninstall rate (standard for App store)
- 31 days of average active use for the remaining 50% of installs
- 11,000 Twitch tasks over one month (100,000 as of today)
- Users unlock their phone 37% of the time they press the unlock button
Control condition: slide-to-unlock

During weekends, users saw slide-to-unlock instead of twitch tasks.
Most twitch tasks are equally fast as slide to unlock

Field deployment Twitch unlock times (N=11,014)

Activities
- Slide-to-Unlock
- Census:people
- Census:activity
- Census:energy
- Census:attire

Friedman test revealed a significant effect of Twitch activity on unlock time: $\chi^2(4)=24.4$, $p<.001$.

Post-hoc paired Wilcoxon signed rank tests revealed that only Census:activity is slower than slide-to-unlock: $p<0.05$. 
Cognitive load study

Do Twitch tasks decrease performance on a working memory task?

Instrument: 2-back test

Cognitive load study: walkthrough

Twitch Study (round 1/6)

New task: remember the letter that was 2 letters ago

Next letter
Cognitive load study: walkthrough

Remember this letter.
Cognitive load study: walkthrough
Cognitive load study: walkthrough
Cognitive load study: walkthrough

Twitch Study (round 1/6)

Compare to 2 letters ago

Different!
Cognitive load study: walkthrough
Cognitive load study: walkthrough
Cognitive load study: walkthrough
Cognitive load study: walkthrough
Cognitive load study: walkthrough

Twitch Study (round 1/6)

Compare to 2 letters ago

Same  Different
Cognitive Load Study: Walkthrough

Twitch Study (round 1/6)
Compare to 2 letters ago

V

N

Same!

Twitch Study (round 1/6)
Compare to 2 letters ago

V

Same
Different
Cognitive load study: measures

Measured the effect of Twitch tasks on reaction time and accuracy

Twitch tasks increase reaction time no more than slide-to-unlock

Two-way ANOVA on Twitch activity and task type revealed a significant effect of Twitch activity on reaction time: $F(6, 5378)=25.8, p<.001$.

Post-hoc Tukey test: no significant difference between Twitch tasks and slide-to-unlock.
Twitch tasks do not impact accuracy

Logistic regression predicting accurate response: **no significant effect of Twitch activity.**
Will people continue completing twitch crowdsourcing tasks over a long period of time?
Discussion

Tasks become repetitive if the user is in the same location for a long period of time.
Future work

- End-user authored microtasks
- Intelligent task assignment
- Twitch as an experience sampling science platform
Conclusion

Twitch crowdsourcing tasks:

• Are fast,
• Incur low cognitive load, and
• Contribute to worthwhile goals.
twitch.stanford.edu
hci.stanford.edu