Secret-Sharing-Based Secure User Recruitment Protocol for Mobile Crowdsensing

Mingjun Xiao [a], Jie Wu [b],
Sheng Zhang [c], and Jiapeng Yu [a]

[a] University of Science and Technology of China,
[b] Temple University, [c] Nanjing University

Outline

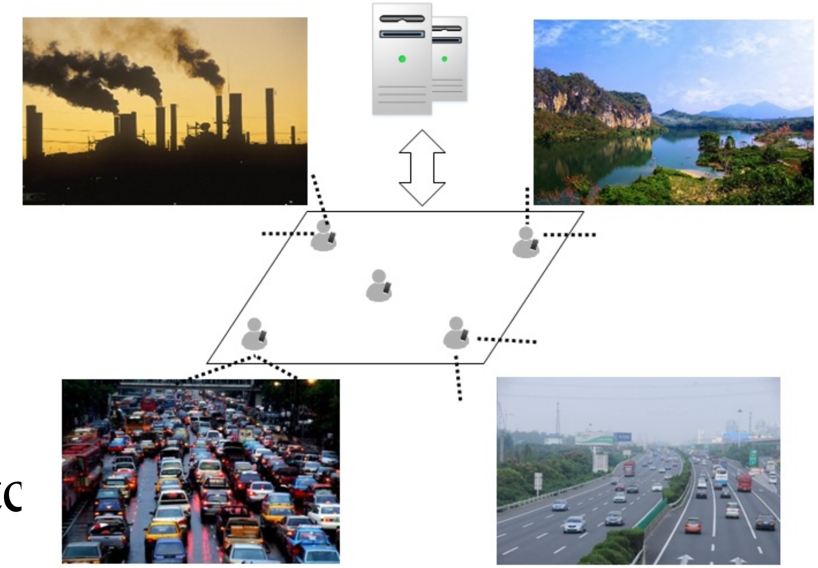
- Motivation
- Model & Problem
- Solution
- Extension
- Evaluation
- Conclusion

Motivation

Mobile Crowdsensing

- **Flexible Sensing Paradigm**

  A requester can recruit a group of mobile users via a platform and coordinate them to perform some sensing tasks by using their smart phones



- **Applications**

  Urban WiFi characterization,

  Traffic information mapping,

  Noise pollution monitoring, etc

Motivation

Mobile Crowdsensing

- **Typical system**

  a platform on the cloud

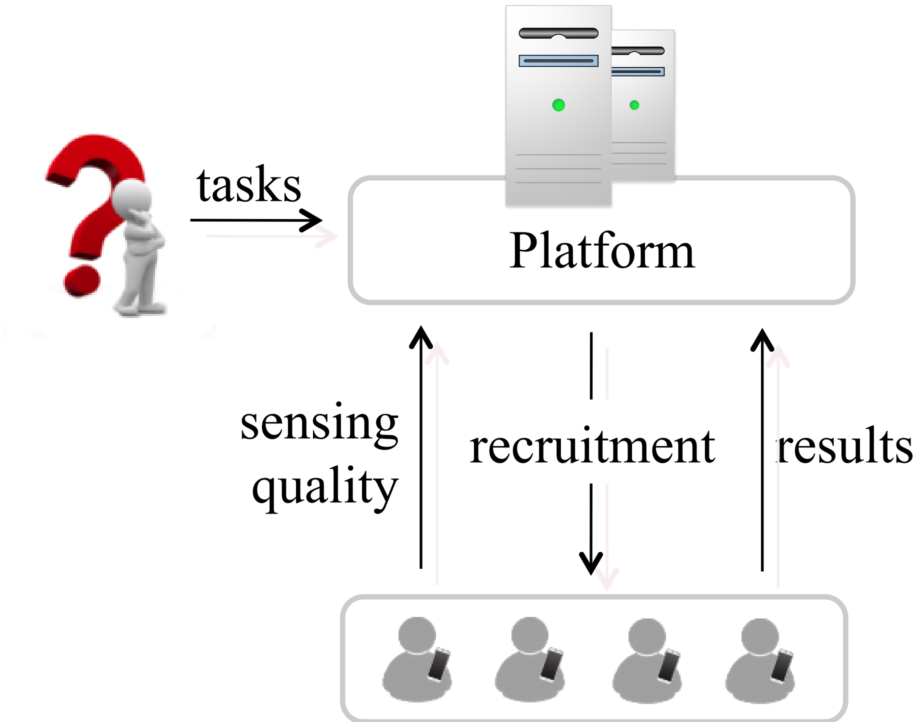
  a group of mobile users

  requesters

- **User recruitment**

  minimum users/cost

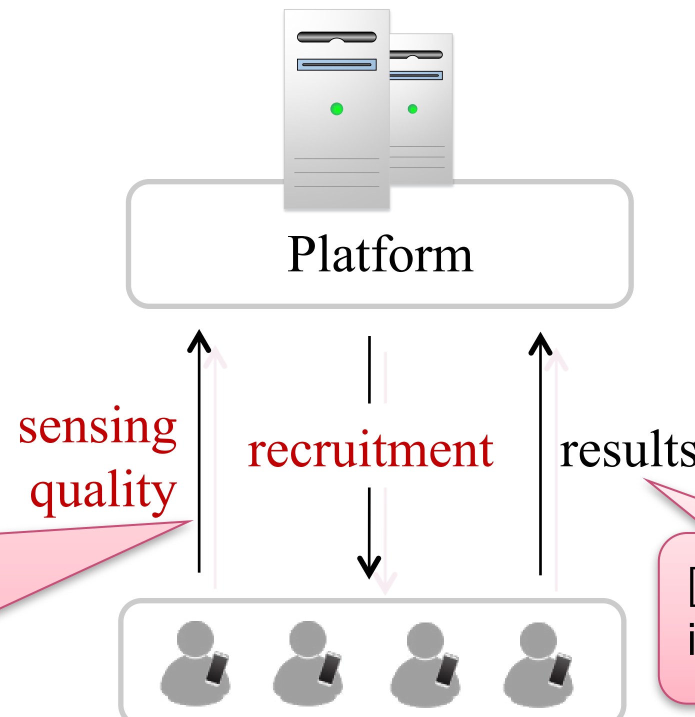
  enough sensing quality



Motivation

Protecting users' privacy

- **Sensing results**
- **Privacy-preserving user recruitment**



Imply which locations each user might visit, and the corresponding frequency, distance, time, etc.

[Rescuedp, infocom16]

Model

Quality-sensitive Crowdsensing Model

- m sensing tasks: $S = \{s_1, \ldots, s_m\}$
- n mobile users: $U = \{u_1, \ldots, u_n\}$
- each user might perform one or more tasks
- sensing quality: $q_{i,j} \in Z_p$, $1 \leq i \leq n$, $1 \leq j \leq m$
- $q_{i,j} = 0$ means that user u_i cannot deal with task s_j

Semi-honest Security Model

- **the dishonest aspect:** each user will try to derive the extra information from the received data
- **the honest aspect:** the user will also follow the whole user recruitment protocol, so as to benefit from the crowdsensing

Problem

Privacy–preserving User Recruitment Problem

- **Objective:** Securely recruit some users to perform all tasks so that we can minimize the number of recruited users, while ensuring that the total sensing quality of each task is no less than a given threshold θ.
- **Formalization**

$$\begin{aligned} &Minimize: && |\Phi| \\ &Subject\ to: && \Phi \subseteq U \\ & && Q_j \geq \theta, \quad 1 \leq j \leq m \\ & && \text{Security under the semi-honest model} \end{aligned}$$

Total sensing quality of task s_j : $Q_j = \sum_{u_i \in \Phi} q_{i,j}$

Solution

Problem Hardness Analysis

- **Theorem 1:** The user recruitment problem is NP-hard.

Basic idea

- **Basic User Recruitment (BUR) protocol**

↓ **Secret sharing scheme**
Secure multi-party computation

- **Secure User Recruitment (SUR) protocol**

Basic User Recruitment Protocol

Utility Function

- **Utility function $f(\Phi)$** indicates the total sensing qualities of all tasks in S contributed by the users in set Φ, until they reach the threshold θ:

$$f(\Phi) = \sum_{j=1}^{m} \min\{Q_j, \theta\} = \sum_{j=1}^{m} \min\{\sum_{u_i \in \Phi} q_{i,j}, \theta\}$$

- **Marginal utility**

$$\Delta_i f(\Phi) = f(\Phi \cup \{u_i\}) - f(\Phi)$$

Greedy User Recruitment Strategy

$$\Phi \Leftarrow u_i : \arg\max_{u_i \in U \setminus \Phi} \Delta_i f(\Phi)$$

Basic User Recruitment Protocol

The Detailed BUR Protocol

Protocol 1 The BUR Protocol

Input: $\mathcal{U}$, $\mathcal{S}$, $\{q_{i,j} | u_i \in \mathcal{U}, s_j \in \mathcal{S}\}$, θ

Output: Φ, $b_1, \cdots, b_n$

Phase 1: the requester publishes $\mathcal{S}$ to $\mathcal{U}$ via the platform;

Phase 2: users input their sensing quality values;

for $i = 1$ **to** n **do**

  user u_i sends $\{q_{i,1}, \cdots, q_{i,m}\}$ to the platform;

Phase 3: the platform makes the decision of user recruitment;

$\Phi = \emptyset$; $f(\Phi) = 0$;

while $f(\Phi) < m\theta$ **and** $|\Phi| < n$ **do**

  Select a user $u_i \in \mathcal{U} \backslash \Phi$ to maximize $\Delta_i f(\Phi)$;

  $\Phi = \Phi \cup \{u_i\}$;

Phase 4: the platform returns the results to users;

for $i = 1$ **to** n **do**

  if $u_i \in \Phi$ **then**

    the platform returns $b_i = 1$ to user u_i;

  else

    the platform returns $b_i = 0$ to user u_i;

Basic User Recruitment Protocol

Correctness and Approximation Ratio of BUR

- **Theorem 2:** $f(\Phi)$ is an increasing function with $f(\emptyset)=0$.
- **Theorem 4:** $f(\Phi)$ is a submodular function.
- **Theorem 5:** the BUR protocol is correct.
- **Theorem 6:** BUR can produce a $(1+\ln\gamma)$-approximation solution, where $\gamma = \max_{u_i \in U} f(\{u_i\})$

Secure User Recruitment Protocol

Secret shares

the shares of a secret s among n users are denoted as

$$[s] \equiv (s[1], \cdots, s[i], \cdots, s[n])$$

$s[i]$ is the i-th user's share.

Shamir's secret sharing scheme

Let p be an odd prime and Z_p be a prime field. To share a secret s ($s \in Z_p$) among n users ($n < p$), Shamir's scheme determines a random polynomial function

$$g_s(x) = s + \alpha_1 x + \alpha_2 x^2 + \ldots + \alpha_k x^k \bmod p$$

with randomly chosen for $\alpha_i \in Z_p$ for $1 \leq i \leq k$, $k \leq n$. Then, the share of the i-th user is $s[i] = g_s(i)$.

Secure User Recruitment Protocol

The Building Blocks

- **Secure multi-party addition/subtraction operation**

$$[z_1] \leftarrow \textit{SecAdd}([x], [y]), \qquad [z_2] \leftarrow \textit{SecSub}([x], [y])$$

$$[x] \equiv (x[1], \cdots, x[i], \cdots, x[n])$$

$$[y] \equiv (y[1], \cdots, y[i], \cdots, y[n])$$

$$[z_1] \equiv (z_1[1], \cdots, z_1[i], \cdots, z_1[n]) \qquad [z_2] \equiv (z_2[1], \cdots, z_2[i], \cdots, z_2[n])$$

$z_1[i]=x[i]+y[i] \bmod p$

$z_2[i]=x[i] - y[i] \bmod p$

Secure User Recruitment Protocol

The Building Blocks

- **Secure multi-party multiplication/comparison operation**

$[z_3] \leftarrow \textit{SecMulti}([x],[y]), \qquad [z_4] \leftarrow \textit{SecCmp}([x],[y])$

$$[x] \equiv (x[1], \cdots, x[i], \cdots, x[n])$$

$$[y] \equiv (y[1], \cdots, y[i], \cdots, y[n])$$

$$[z_3] \equiv (z_3[1], \cdots, z_3[i], \cdots, z_3[n]) \qquad [z_4] \equiv (z_4[1], \cdots, z_4[i], \cdots, z_4[n])$$

$$z_3 = xy \bmod p$$

$$z_4 = \begin{cases} 1 \bmod p, & x \le y \\ 0 \bmod p, & x > y \end{cases}$$

Secure User Recruitment Protocol

The Building Blocks

- **Secure multi-party max/min operation**

$$[z_5] \leftarrow \boldsymbol{SecMax}([x], [y]), \qquad [z_6] \leftarrow \boldsymbol{SecMin}([x], [y])$$

$$[x] \equiv (x[1], \cdots, x[i], \cdots, x[n])$$

$$[y] \equiv (y[1], \cdots, y[i], \cdots, y[n])$$

$$[z_5] \equiv (z_5[1], \cdots, z_5[i], \cdots, z_5[n])$$

$$z_5 = \max\{x, y\} \bmod p$$

$$SecMax([x],[y]) \equiv SecAdd([x], SecMulti(SecCmp([x],[y]), SecSub([x],[y])))$$

$$[z_6] \equiv (z_6[1], \cdots, z_6[i], \cdots, z_6[n])$$

$$z_6 = \min\{x, y\} \bmod p$$

$$SecMin([x],[y]) \equiv SecAdd([x], SecMulti(SecSub(1 - SecCmp([x],[y])), SecSub([x],[y])))$$

Secure User Recruitment Protocol

From BUR to SUR

- **Inputs**

  For each user's sensing quality: $q_{i,j} \rightarrow [q_{i,j}]$

- **Outputs**

  User recruitment result: $\Phi \rightarrow (b_1, \cdots, b_n)\,;\, u_i \in \Phi \rightarrow [b_i] = [1]$

- **Securely compute marginal utility**

$$\Delta_i f(\Phi) = f(\Phi \cup \{u_i\}) - f(\Phi)$$

$$= \sum_{j=1}^{m} \min\{q_{i,j},\ \theta - Q_j\}$$

$$[\Delta_i f] \leftarrow SecAdd_{j=1}^{m} : SecMin([q_{i,j}], SecSub(\theta, [Q_j]))$$

Secure User Recruitment Protocol

From BUR to SUR

- **Securely determine the recruited user**

$$[\Delta_{\max} f] \leftarrow SecMax([\Delta_1 f], \cdots, [\Delta_n f])$$

$$\text{for } i = 1 \rightarrow n \text{ do}$$

$$[z] \leftarrow SecCmp([\Delta_{\max} f], [\Delta_i f]);$$

$$[b_i] \leftarrow SecAdd([b_i], SecMulti(SecSub([1], [b_i]), [z]));$$

- **Securely update the total sensing quality in each round**

$$\text{for } j = 1 \rightarrow m \text{ do}$$

$$[\delta] \leftarrow SecMin([q_{i,j}], SecSub(\theta, [Q_j]));$$

$$[Q_j] \leftarrow SecAdd([Q_j], SecMulti([z], [\delta]));$$

Secure User Recruitment Protocol

The Detailed SUR Protocol

Protocol 2 The SUR Protocol

Input: $\mathcal{U}$, $\mathcal{S}$, $\{q_{i,j} | u_i \in \mathcal{U}, s_j \in \mathcal{S}\}$, θ

Output: $b_1, \cdots, b_n$

Phase 1: the requester publishes $\mathcal{S}$ to $\mathcal{U}$ via the platform;

Phase 2: users input their sensing quality vectors;

for $i=1$ **to** n **do**
  user u_i determines the sensing qualities $q_{i,1}, \cdots, q_{i,m}$;
  for $j=1$ **to** m **do**
    user u_i generates the polynomial sharing $[q_{i,j}]$;
    user u_i sends the share $q_{i,j}[i']$ to user $u_{i'}$;

Phase 3: users jointly make the decision of user recruitment;

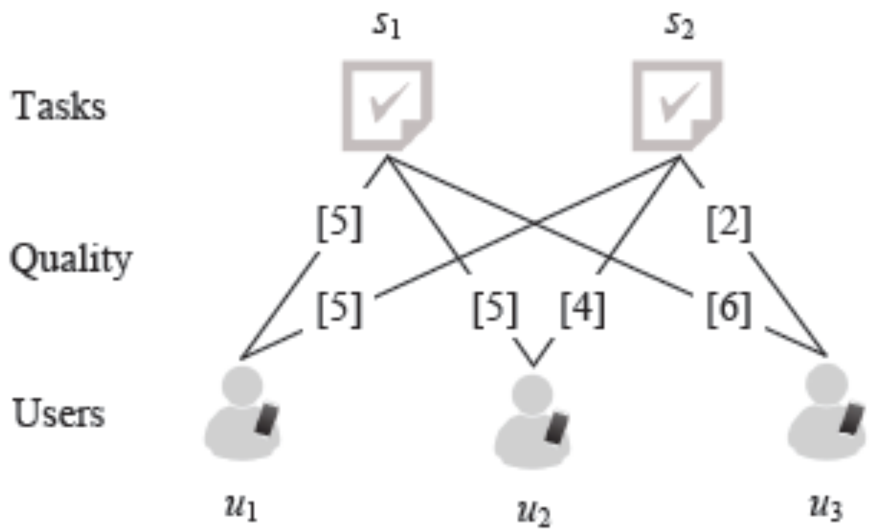
for $i=1$ **to** n **do**
  $[b_i] \leftarrow [0]$;
for $j=1$ **to** m **do**
  $[Q_j] \leftarrow [0]$;
for $round=1$ **to** n **do**
  for $i=1$ **to** n **do**
    $[\Delta_i f] \leftarrow [0]$;
    for $j=1$ **to** m **do**
      $[\delta] \leftarrow SecMin([q_{i,j}], SecSub(\theta, [Q_j]))$;
      $[\Delta_i f] \leftarrow SecAdd([\Delta_i f], [\delta])$;
    $[\Delta_i f] \leftarrow SecMulti([\Delta_i f], SecSub([1], [b_i]))$;
  $[\Delta_{max} f] \leftarrow SecMax([\Delta_1 f], \cdots, [\Delta_n f])$;
  for $i=1$ **to** n **do**
    $[z] \leftarrow SecCmp([\Delta_{max} f], [\Delta_i f])$;
    $[b_i] \leftarrow SecAdd([b_i], SecMulti(SecSub([1], [b_i]), [z]))$;
    for $j=1$ **to** m **do**
      $[\delta] \leftarrow SecMin([q_{i,j}], SecSub(\theta, [Q_j]))$;
      $[Q_j] \leftarrow SecAdd([Q_j], SecMulti([z], [\delta]))$;

Phase 4: the users reconstruct the results;

for $i=1$ **to** n **do**
  user u_i collects all shares of $[b_i]$;
  user u_i derives $b_i = \sum_{j=1}^{m} b_i[j]$;

Secure User Recruitment Protocol

Example



(a) Users, tasks and sensing qualities

round 1	round 2
$[\Delta_1 f] = [10]$	$[\Delta_1 f] = [0]$
$[\Delta_2 f] = [9]$	$[\Delta_2 f] = [6]$
$[\Delta_3 f] = [8]$	$[\Delta_3 f] = [5]$
$[b_1] = [1]$	$[b_1] = [1]$
$[b_2] = [0]$	$[b_2] = [1]$
$[b_3] = [0]$	$[b_3] = [0]$
$[Q_1] = [5]$	$[Q_1] = [8]$
$[Q_2] = [5]$	$[Q_2] = [8]$

(b) Intermediate results

The total sensing quality threshold θ=8

Secure User Recruitment Protocol

Performance Analysis

O(mn^2) invocations of secure multiplication operations

O(mn^4l) bit-operations per user ($l = \lceil \log_2 p \rceil$)

O(mn^2l) rounds of communication

Correctness and Approximation Ratio

- **Theorem 7:** SUR is correct, and it can also produce a (1+lnγ)-approximation solution, where $\gamma = \max_{u_i \in U} f(\{u_i\})$

Secure User Recruitment Protocol

Security of SUR

- **Theorem 8:** SUR can protect the sensing qualities of each user from being revealed to any κ semi-honest adversaries and the platform, even if they might collude, where κ (i.e., the degree of polynomial sharing) may be any integer less than n.

Extension

Extension: the total sensing quality function

- $Q(\bullet)$ becomes a general function about $q_{i,j}$

$$Q_j(\Phi) \equiv Q(q_{i,j}\,|_{u_i \in \Phi})$$

- **Example**

The sensing quality $q_{i,j}$ represents the probability of successful sensing

$Q(\bullet)$ may be defined as the joint successful probability

$$Q_j(\Phi) = 1 - \prod_{u_i \in \Phi} (1 - q_{i,j})$$

Extension

Extension

- **Theorem 9:** When $Q_j(\Phi)$ is a trivial function that can be securely computed by using the secure multi-party computation operations in SUR, SUR will still be secure.

- **Theorem 10:** When $Q_j(\Phi)$ is an increasing submodular function with $Q_j(\Phi=\varnothing)=0$, we have: 1) the utility function $f(\Phi)$ is still submodular; 2) SUR can still produce a (1+ lnγ)-approximation solution, where $\gamma = \max_{u_i \in U} f(\{u_i\})$

Evaluation

Evaluate the User Recruitment Performance

- **Compared Protocols**

  MCUR: the user who can perform the most tasks is recruited first

  MQUR: the user who performs tasks with the most sensing qualities is recruited first

- **Simulation Settings**

  Synthetic traces

  Metric: the number of recruited users

Evaluation

Evaluate the User Recruitment Performance

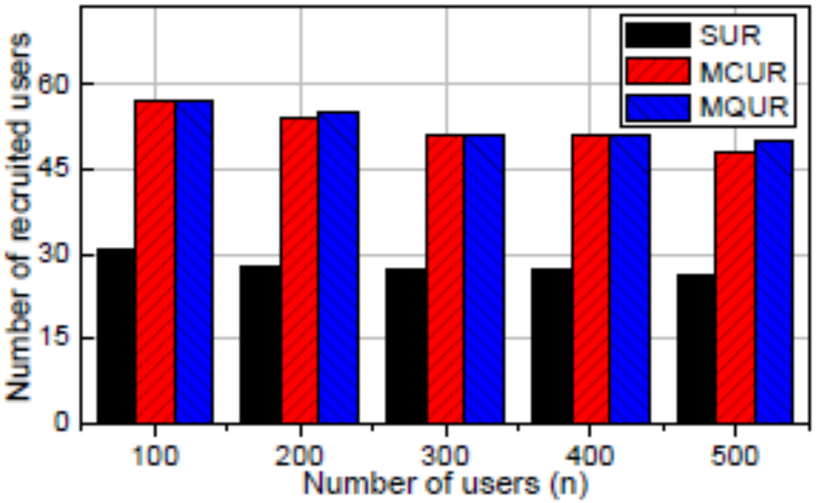
- Simulation Settings

Parameter name	default	range
number of users n	200	100-500
number of tasks m	100	50-250
average sensing quality p	30	10-90
variance of sensing qualities σ	0.4	0.2-1.0
sensing quality threshold θ	100	20-250
largest number of tasks per user ρ	20	15-35

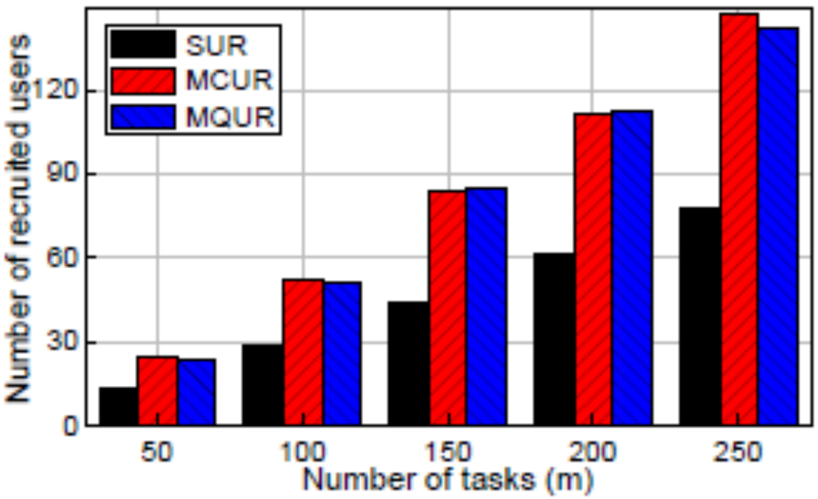
Evaluation

Evaluate the User Recruitment Performance

- Evaluation Results



(a) $m = 100$



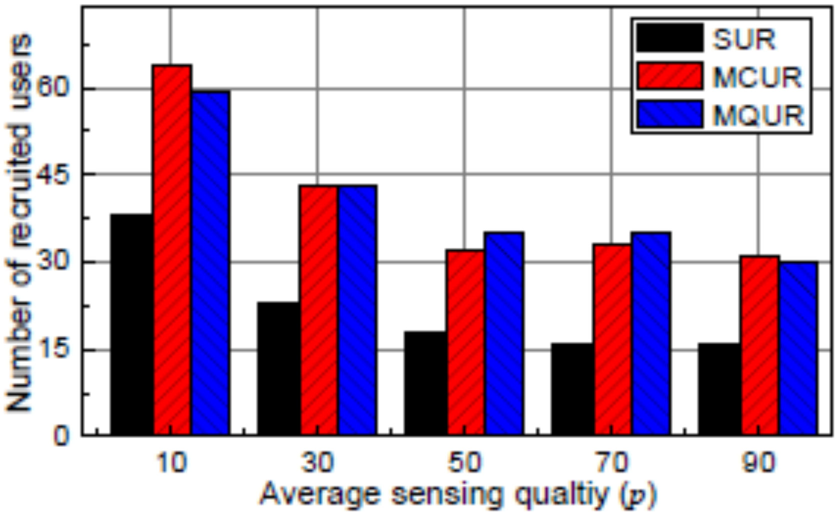
(b) $n = 200$

Number of recruited users vs. number of users and tasks

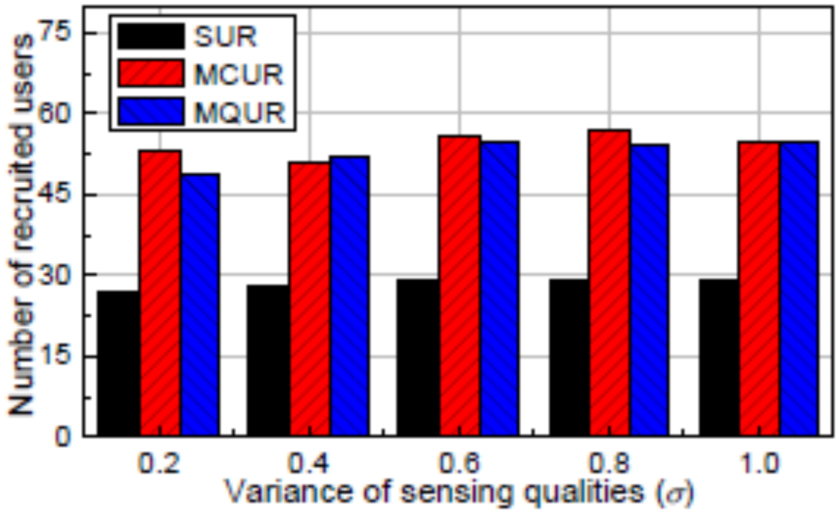
Evaluation

Evaluate the User Recruitment Performance

- Evaluation Results



(a) $\sigma = 0.4$



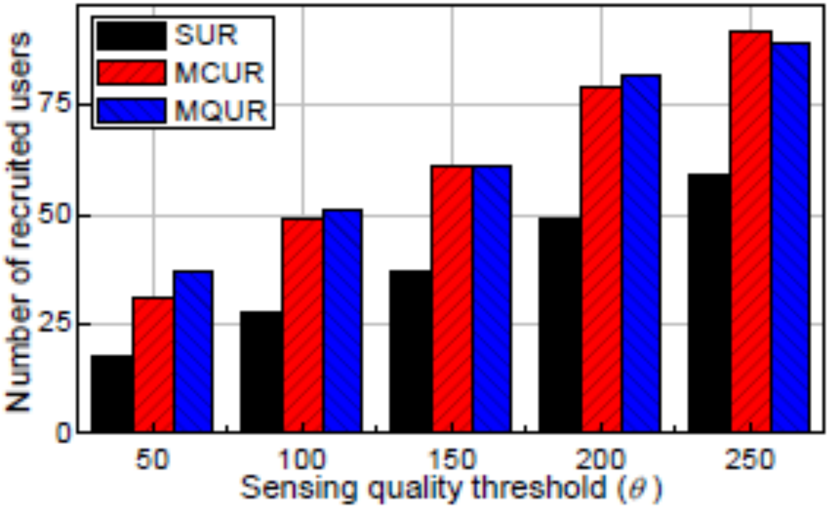
(b) $p = 30$

Number of recruited users vs. average sensing quality and variance

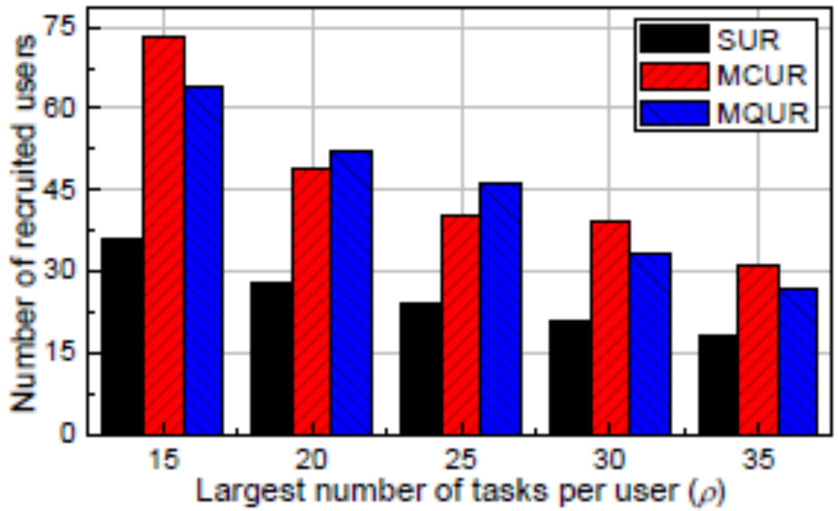
Evaluation

Evaluate the User Recruitment Performance

- Evaluation Results



(a) $\rho = 20$



(b) $\theta = 100$

Number of recruited users vs. sensing quality threshold and largest number of tasks performed by each user

Evaluation

Evaluate the Time Efficiency

- **Compared Protocols**

  HEUR: Homomorphic-Encryption-based User Recruitment

  GCUR: Garbled-Circuit-based User Recruitment

- **Experiment Settings**

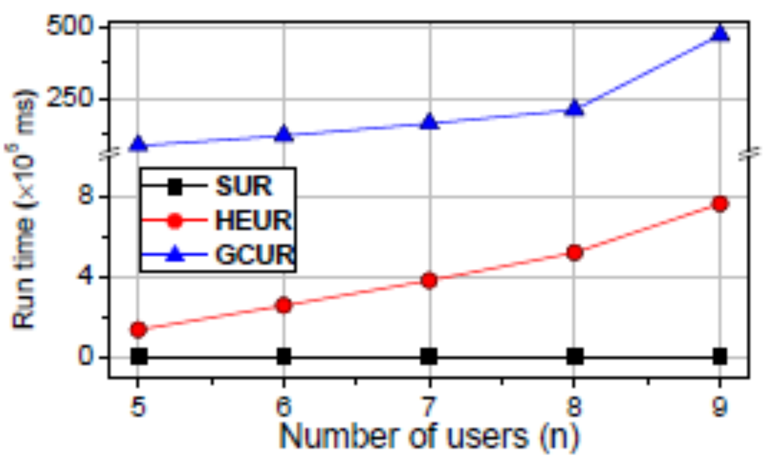
  2.0GB memory

  a processor of 4-core 2.2GHz plus 4-core 1.5GHz

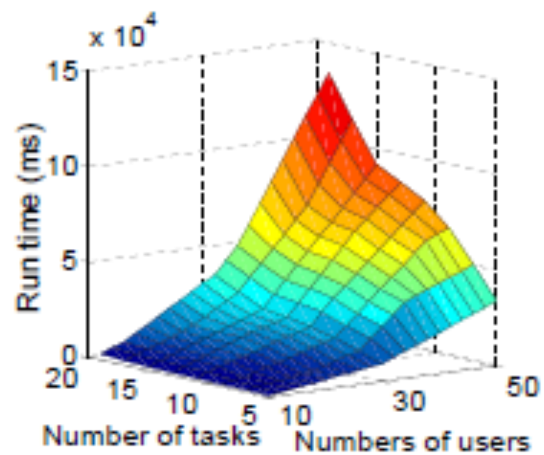
Evaluation

Evaluate the Time Efficiency

- Experiment Results



(a) Run time of three protocols



(b) Run time of SUR

Evaluation: run time vs. the number of users and tasks

Conclusion

- **SUR can produce a solution with a logarithmic approximation ratio**
- **SUR can protect the inputs of each user from being revealed to the platform or to other users, even if they might collude**
- **Simulation results show that SUR can work well in real smartphones**

Q&A

Thank You!