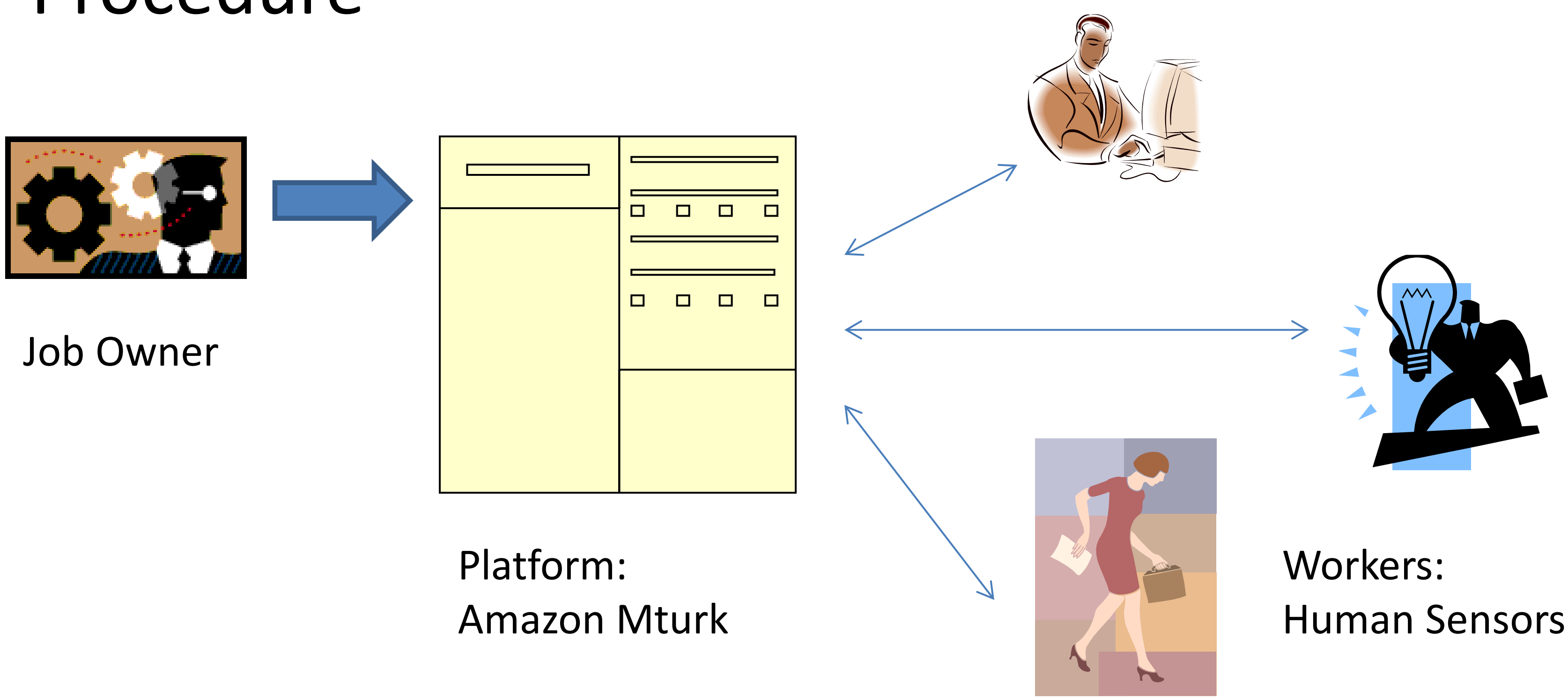
Correlated Friends’ Impacts in Social-crowdsensing

Wei Chang, Department of Computer Science,
Saint Joseph’s University
Wei-Shih Yang, Department of Mathematics,
Jie Wu, Department of Computer and Information Sciences,
Temple University

Background

- Social sensing is a typical application of the crowdsourcing system.
- Procedure



Defects

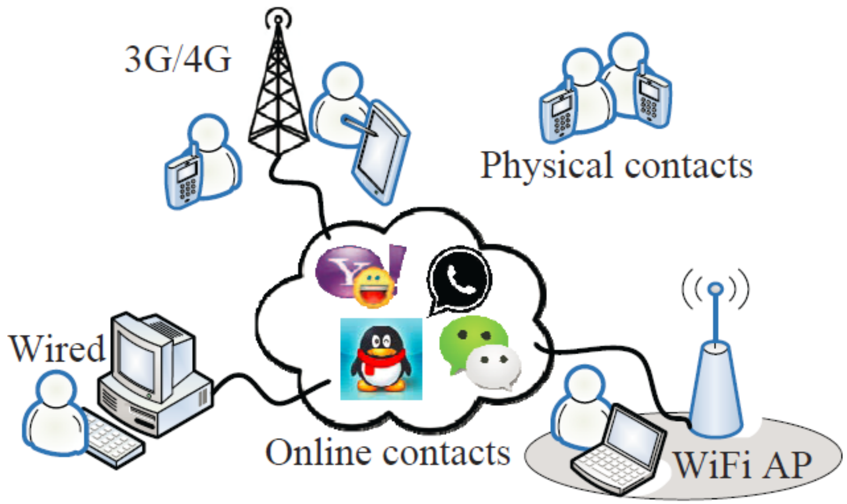
- There are 3 defects with the existing system:
  1. Lacks a timely advertising mechanism to recruit works;
     - New task: tempting payment
     - Offline workers are not aware of its existence
  2. Centralized and platform-specified;
     - Malfunctions / Unavailable
  3. Someone has to pay a fee for using the platform
     - Mturk collects a 10% commission on top of a task's total payment

Social-crowdsensing

- A self-organized, distributed, and cross-platform crowdsensing system
- Main Idea:
  - To create a multilayered outsourcing structure via any stochastic social contacts, including physical encounters between friends and virtual contacts via any online-chatting system

Social-crowdsensing

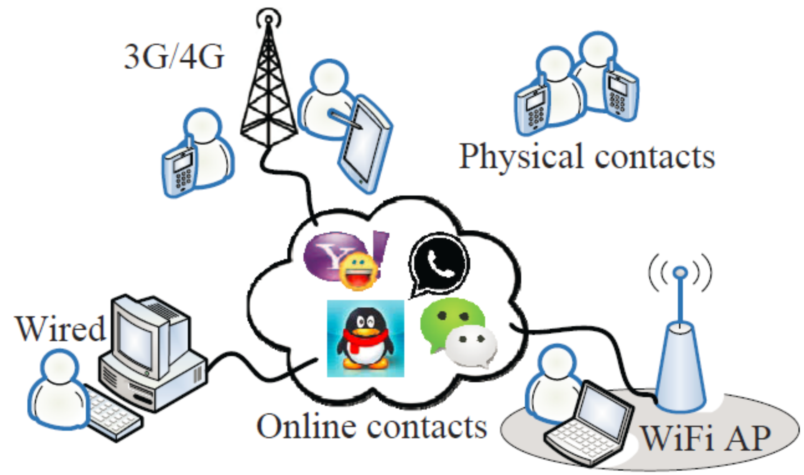
- General procedure:
  - A job owner first creates an SC task (first worker)
  - Any worker can further recruit new workers via stochastic social contacts (multi-hop relays)

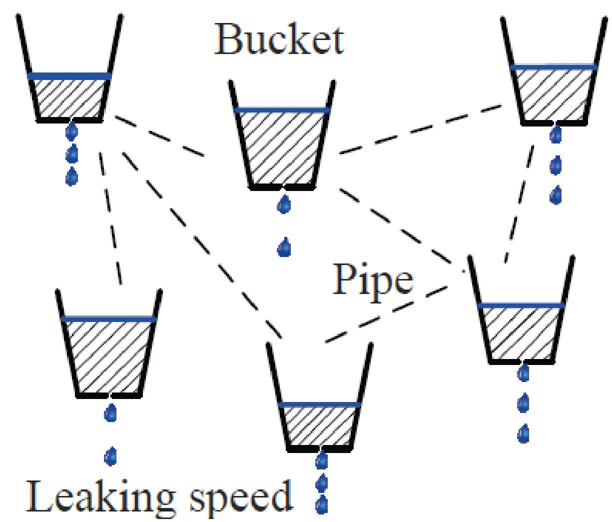


  - Sensing results are returned via the Internet or physical encounters-based multi-hop relays.

Question 1

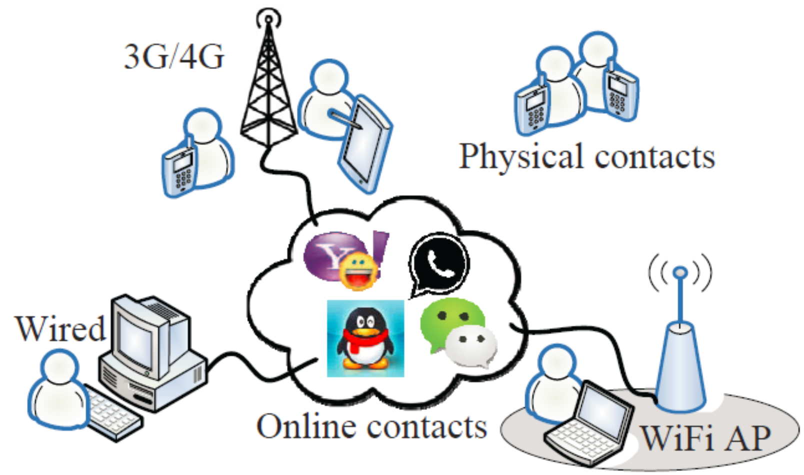
- Given that SC tasks initiate at random nodes, by what workload allocation strategy can the entire task's completion time be minimized?
- Minimizing the completion time = Maximize a system's utilization
- Estimate an overall computing capacity

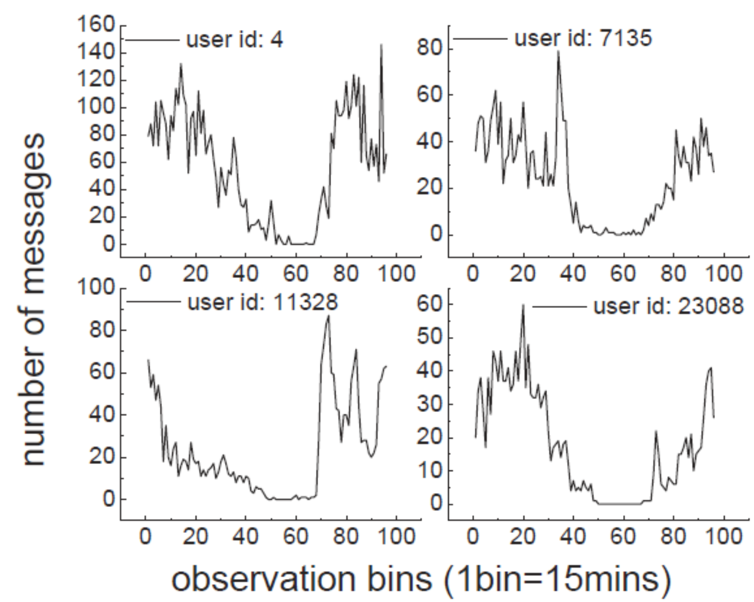




Question 2

- Friends are correlated.
- For example, colleagues have similar working hours, and therefore, their times of unavailability for participating in the sensing tasks are not independent.
- Incorporate correlations into the decision-making phase of workload allocations
- How to model correlation





Correlations Modeling

- Based on a user's real-time contacting frequency, we associate a virtual state to each user, called active level. $\sigma_i = \{-k, \cdots, -2, -1, 1, 2, \cdots, k\}$
- Active levels of all workers at any moment:

$$\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_i, \ldots, \sigma_m)$$

- The probability distribution of workers' active levels

$$P(\sigma) = \frac{1}{Z} e^{\sum_i X_i \sigma_i + \sum_{i \neq j} X_{ij} \sigma_i \sigma_j + \sum_{i \neq j \neq k} X_{ijk} \sigma_i \sigma_j \sigma_k} \ldots$$

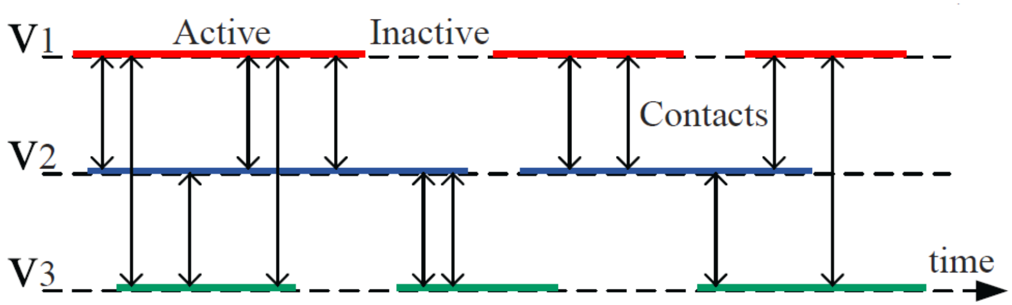
$$\approx \frac{1}{Z} \times e^{\sum_{i \neq j} J_{ij} \sigma_i \sigma_j + \sum_i h_i \sigma_i}$$

Simulate Stochastic Contacts

- We build a discrete-time Markov Chain to simulate the changes of users' active levels.
- Assume that, at any moment, at most one worker changes its active level.
- Transition probability $P(\sigma) = \sum_{\sigma'} P(\sigma')P(\sigma', \sigma)$

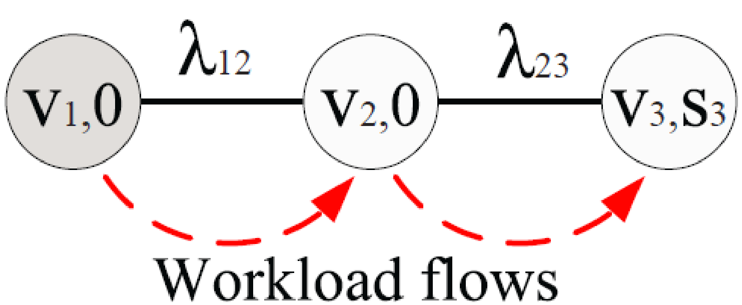
$$P(\sigma, \sigma') = \begin{cases} \frac{1}{m} \cdot \frac{e^{\sum_{i \neq j} J_{ij}\sigma_i\sigma'_j + h_j\sigma'_j}}{\sum_{\sigma_j=\pm 1} e^{\sum_{i \neq j} J_{ij}\sigma_i\sigma_j + h_j\sigma_j}} & \text{if} \quad \sigma' \sim_j \sigma \\ \frac{1}{m} \cdot \frac{e^{\sum_{i \neq j} J_{ij}\sigma_i\sigma_j + h_j\sigma_j}}{\sum_{\sigma_j=\pm 1} e^{\sum_{i \neq j} J_{ij}\sigma_i\sigma_j + h_j\sigma_j}} & \text{if} \quad \sigma' = \sigma \\ 0 & \text{if} \quad \text{others} \end{cases}$$

- Simulate



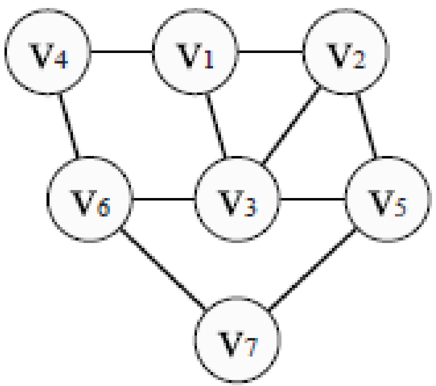
Workload Allocation

- The capacity *q* reflects the amount of workload that a worker and his friends can process within one unit of time.
- Basic Rule: $(w_1^*, w_2^*) = \left[\frac{(w_1 + w_2)q_1}{q_1 + q_2}, \frac{(w_1 + w_2)q_2}{q_1 + q_2}\right]$
- The overall capacity cannot be a fixed-value



Pairwise-Capacity Estimation

- Approximate it by 2-hop social information
- Without Correlation



Algorithm 1: Pairwise Overall Speed

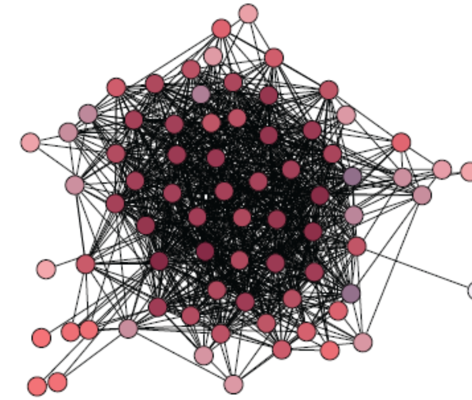
/*Suppose that v_u is contacting with $v_{u'}$*/
Eliminate node $v_{u'}$ from social contact graph G
Initialization $q_u \leftarrow s_u$ /*Local contribution*/
for $i \in N(u)$ **do**
  $q_u \leftarrow q_u + \lambda_{ui} s_i$ /*1-hop contribution*/
  for $j \in N(i)$ **do**
    $q_u \leftarrow q_u + \lambda_{ui} \lambda_{ij} s_j$ /*2-hop contribution*/
    if $j \in N(u)$ **then**
      $q_u \leftarrow q_u - \lambda_{ui} \lambda_{ij} \lambda_{uj} s_j$ /*Double counted*/
Return $q^{u}_{uu'} \leftarrow q_u$

- With Correlation $\lambda_{ij}(\sigma_i, \sigma_j)$
  - It gives the average contacting frequency for each type of active level combination.

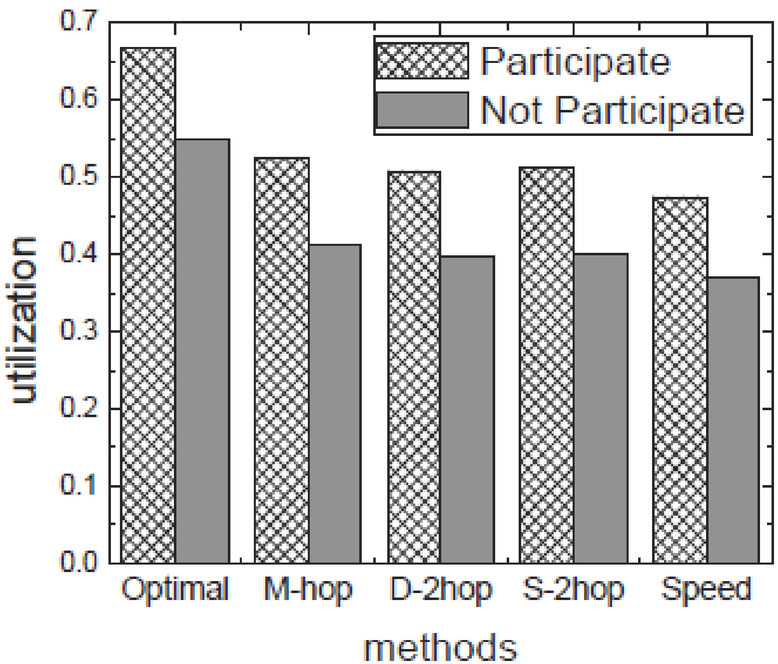
Evaluation

System Utilization

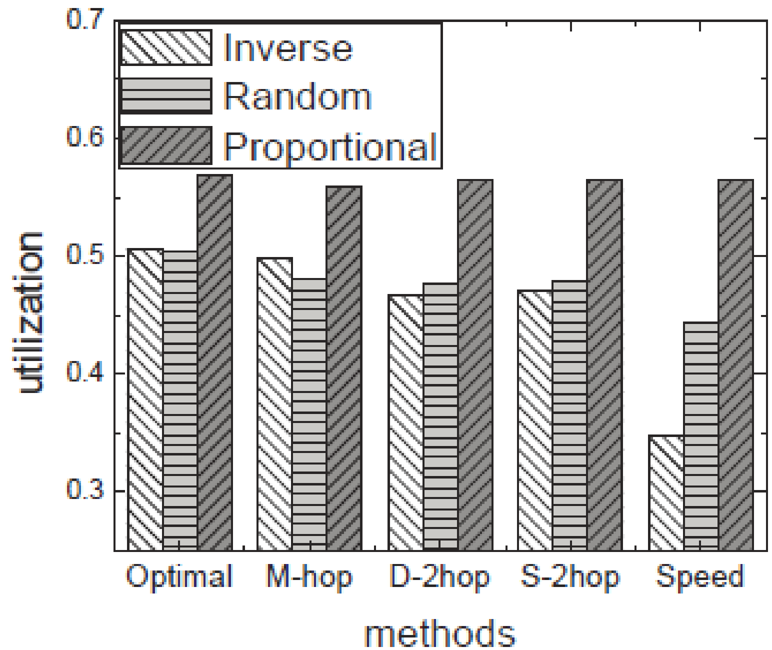
$$U = \frac{\sum_{i=1}^{m} s_i \int_0^{t^*} \delta_i(t)\, dt}{\sum_{i=1}^{m} s_i t^*}$$



Impact of 0-speed workers

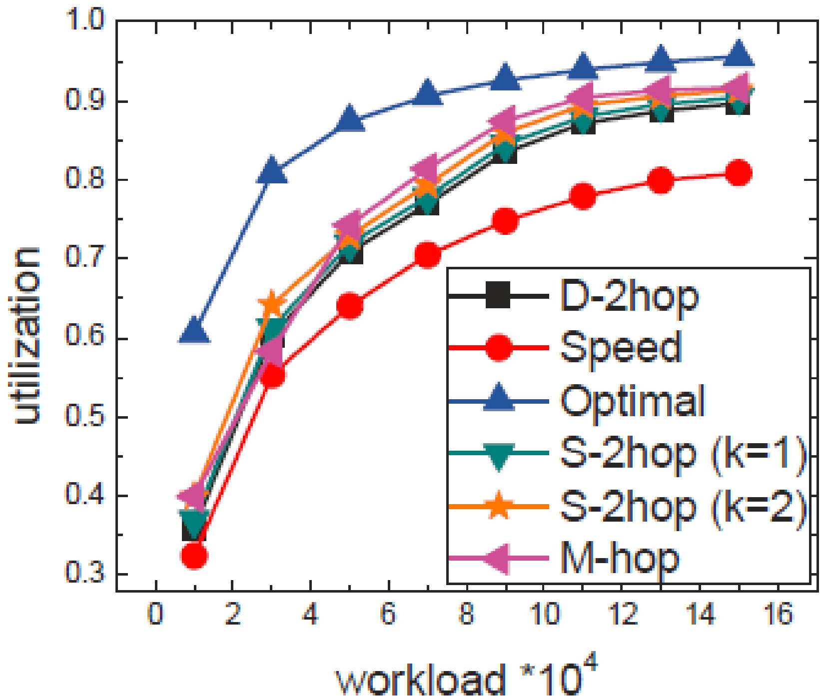


Impact of speed distribution

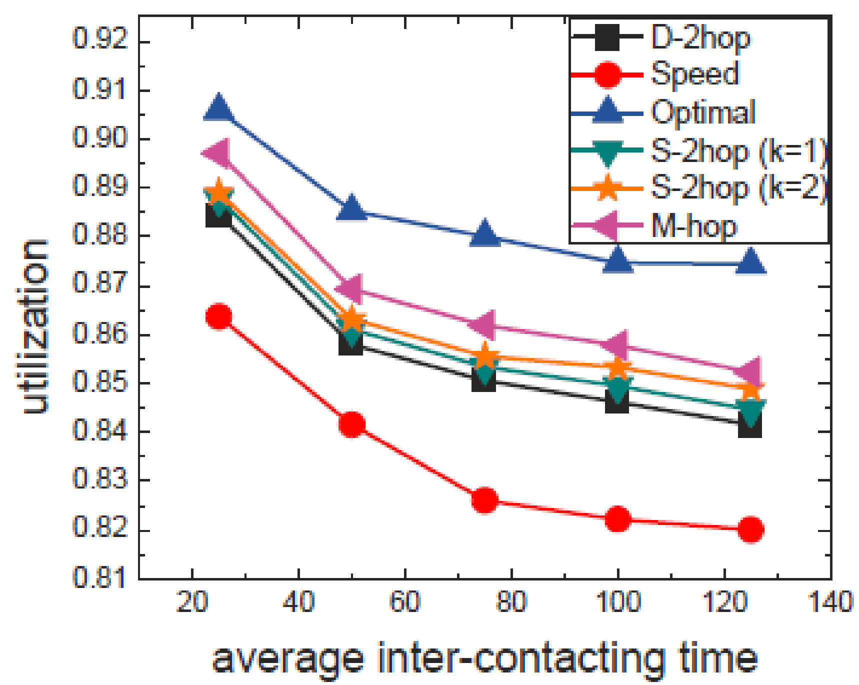


Evaluation

Workload's size



Ave inter-contacting intervals



Thank you !