

# Ethically Aligned Opportunistic Scheduling for Productive Laziness

Han Yu<sup>1,2,3</sup>, Chunyan Miao<sup>1,2,3</sup>, Yongqing Zheng<sup>4,5</sup>, Lizhen Cui<sup>4</sup>, Simon Fauvel<sup>2</sup> and Cyril Leung<sup>6</sup>

<sup>1</sup>School of Computer Science and Engineering, Nanyang Technological University (NTU), Singapore

<sup>2</sup>Joint NTU-UBC Research Centre of Excellence in Active Living for the Elderly (LILY), NTU, Singapore

<sup>3</sup>Alibaba-NTU Singapore Joint Research Institute, Singapore

<sup>4</sup>School of Software Engineering, Shandong University, Jinan, Shandong, China

<sup>5</sup>Shanda Dareway Software Co. Ltd, Jinan, Shandong, China

<sup>6</sup>Department of Electrical and Computer Engineering, The University of British Columbia, Vancouver, BC, Canada

Corresponding Authors: han.yu@ntu.edu.sg, clz@sdu.edu.cn

## Abstract

In artificial intelligence (AI) mediated workforce management systems (e.g., crowdsourcing), long-term success depends on workers accomplishing tasks productively and resting well. This dual objective can be summarized by the concept of productive laziness. Existing scheduling approaches mostly focus on efficiency but overlook worker wellbeing through proper rest. In order to enable workforce management systems to follow the IEEE Ethically Aligned Design guidelines to prioritize worker wellbeing, we propose a distributed *Computational Productive Laziness* (CPL) approach in this paper. It intelligently recommends personalized work-rest schedules based on local data concerning a worker's capabilities and situational factors to incorporate opportunistic resting and achieve superlinear collective productivity without the need for explicit coordination messages. Extensive experiments based on a real-world dataset of over 5,000 workers demonstrate that CPL enables workers to spend 70% of the effort to complete 90% of the tasks on average, providing more ethically aligned scheduling than existing approaches.

## Introduction

In today's world, artificial intelligence (AI) has been employed to manage large-scale workforces such as crowdsourcing systems (Miao et al. 2016; Michelucci and Dickinson 2016; Pan et al. 2016). For example, in DiDi Chuxing, AI technologies are deployed to dynamically match drivers to tasks in order to enhance operation efficiency (Xu et al. 2018). Human workers become fatigued or bored over long sessions of work, which can cause inefficacy (Leiter and Maslach 2015). A recent study found that short breaks significantly improve workers' motivation while maintaining the quality of work (Dai et al. 2015). However, existing AI approaches for workforce management in crowdsourcing mostly do not explicitly incorporate rest into their recommendations (Chai et al. 2017).

From an ethically aligned design perspective (The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems 2018), it is desirable to incorporate breaks into scheduling approaches in order to prioritize workers' wellbeing. Nevertheless, as highlighted in (Yu et al. 2018), ethically aligned AI approaches need to be designed such that

they balance the concern for human wellbeing while still achieving their design objectives. In this paper, we set out to address this challenge by proposing an ethically aligned opportunistic scheduling approach that can achieve "productive laziness" - *Computational Productive Laziness* (CPL).

Originally conceptualized in social sciences literature, productive laziness (Whillans et al. 2017) is a rule-of-thumb guideline on how workers should approach their work in order to achieve work-life balance. The general idea is to work when you are highly efficient, and rest otherwise. CPL coordinates workers to work when situational factors induce high efficiency or demand working, and rest when they do not. Thus, it is important to identify factors influencing workers' productivity and the urgency of work. We formulate the problem of achieving productive laziness in large-scale systems as a multi-objective constrained optimization problem. Following the Lyapunov optimization-based framework (Yu et al. 2016), we analyse the interaction dynamics involved in a workforce management system and derive a novel *Work-Rest Index (WRI)* which expresses the interplay among four aspects affecting worker effort output:

1. *Situational factors*: the current workload pending in each worker's backlog, and how long they have been pending (as tracked by the system);
2. *Worker performance*: each worker's productivity based on past observation data;
3. *System-level preference*: the emphasis, given by the system operators, on achieving high collective productivity compared to allowing for workers to be more rested; and
4. *Personal preference*: each worker's mood at a given time which is used to infer the level of productivity that can be expected from the worker in the immediate future following the happy-productive worker thesis (Oswald, Proto, and Sgroi 2015).

Based on this index, CPL dynamically determines the timing and amount of work a worker should perform in order to conserve the collective effort output while maintaining a high level of collective productivity in the system. It can be implemented in a distributed manner as a personal scheduling agent with a computational time complexity of  $O(1)$ . This enables it to effectively support the need for real-time scheduling for large-scale workforce management.

CPL allows human values to be codified and algorithmically guide the recommendations by the scheduling agent to balance worker wellbeing and system throughput. Through extensive numerical experiments based on a large-scale real-world dataset containing over 5,000 workers' performance characteristics, CPL is shown to significantly outperform alternative approaches, consistently achieving superlinear collective productivity (Sornette, Maillart, and Ghezzi 2014).

## Related Work

Existing AI-powered workforce management approaches (Lee et al. 2015) can be divided into two main categories: (1) the direct approaches and (2) the indirect approaches.

Some direct approaches seek to balance the division of labour among workers in a situation-aware manner through data-driven deliberation (Mason and Watts 2012; Yu et al. 2013a; Dai et al. 2013; Tran-Thanh et al. 2014b; Yu et al. 2015; 2016; Grossmann, Brienza, and Bobocel 2017; Yu et al. 2017a). Others design reputation and/or incentive mechanisms to motivate workers to work harder (Yu et al. 2011; Mao et al. 2013; Yu et al. 2013b; Tran-Thanh et al. 2014a; Miao et al. 2016; Zeng, Tang, and Wang 2017). They generally leave it up to the workers to plan their rest breaks. There are approaches which implicitly limit how long a worker can continuously work by setting a budget on the number of tasks they are allocated (Chen et al. 2015; Zenonos, Stein, and Jennings 2016). However, these approaches do not opportunistically take advantage of periods when a worker is highly productive (e.g., periods of good mood in our approach). Such ad hoc planning may not help the system maintain high collective productivity. With CPL, a desirable balance between worker wellbeing and collective productivity can be achieved.

Indirect approaches (Morris, Dontcheva, and Gerber 2012) seek to induce good mood among workers in order to improve their productivity. This is based on the happy-productive worker thesis (Oswald, Proto, and Sgroi 2015), which suggests that good (bad) mood improves (hampers) productivity. They also do not explicitly recommend resting to workers.

A recent work (Yu et al. 2017b) is starting to explore how to opportunistically schedule rest. However, it does not account for how long tasks have been pending in workers' backlog and can lead to schedules which do not make business sense. Without considering the urgency of the tasks, it is vulnerable to workers misreporting their mood. With the conceptual queue technique developed in this paper, CPL considers workers' wellbeing, system objectives and situational factors in a more holistic manner, and is better able to deal with misbehaving workers trying to game the system.

## The Proposed CPL Scheduling Approach

The system model in this paper consists of a set of  $N$  workers and a set of  $M$  tasks at any given time slot  $t$ .

- *Workers* are associated with personal profiles. Each profile contains information on a worker  $i$ 's task backlog queue  $q_i(t)$  and his maximum productivity  $\mu_i^{\max}$  which indicates the maximum workload he can complete in a

given time slot  $t$ . The granularity of  $t$  can be adjusted to fit any given system's requirements. The ground truth of  $\mu_i^{\max}$  may not be directly observable. With analytics tools such as Turkalytics (Heymann and Garcia-Molina 2011), it can be tracked and estimated statistically.

- *Tasks* are associated with task profiles. The most important information for our purpose is the task deadline which is expressed in terms of number of time slots since task creation before which a given task must be completed.

The number of workers and tasks available in a given system at different time slots may differ.

We model the dynamics of a worker  $i$ 's task backlog queue following previous research (Yu et al. 2013a; 2015):

$$q_i(t+1) = \max[0, q_i(t) + \lambda_i(t) - \mu_i(t)] \quad (1)$$

where  $\lambda_i(t)$  is the number of new tasks delegated to worker  $i$  during time slot  $t$ ; and  $\mu_i(t)$  is the number of tasks completed by worker  $i$  during time slot  $t$ . Based on the 'happy-productive worker' thesis which suggests that productivity is positively correlated to mood (Oswald, Proto, and Sgroi 2015),  $\mu_i(t)$  can be expressed as a function of  $i$ 's current mood and how much effort he spends on the tasks:

$$\mu_i(t) = \mu(\xi_i(t), m_i(t)) = [\xi_i(t)m_i(t)\mu_i^{\max}] \quad (2)$$

where  $\xi_i(t) \in [0, 1]$  is the normalized effort worker  $i$  spends during time slot  $t$ .  $m_i(t) \in [0, 1]$  is worker  $i$ 's mood during time slot  $t$ , where 1 denotes the most positive mood. It can be self-reported by a worker using tools such as *MoodPanda* (<http://moodpanda.com/>), or through facial image analytics (e.g., in-vehicle monitoring of drivers' mood (Zimasa, Jamson, and Henson 2017)). Here, we treat  $m_i(t)$  as an external parameter to CPL and do not assume that the system is capable of predicting future mood values.

## Deriving the Work-Rest Index

In order to take task pending time into account, we propose a *conceptual queue* technique. Let  $Q_i(t)$  denote a conceptual queue which is being managed by a CPL agent on behalf of a worker  $i$ . The queuing dynamics of this conceptual queue is defined as:

$$Q_i(t+1) = \max[0, Q_i(t) + \mu_i^{\max}1_{[q_i(t)>0 \& \mu_i(t)=0]} - \mu_i(t)] \quad (3)$$

where  $1_{[\text{condition}]}$  is an indicator function. Its value is 1 if and only if [condition] is satisfied; otherwise, it evaluates to 0. When first created, a conceptual queue is empty. The conceptual queue starts from 0 at  $t = 0$  (i.e.  $Q_i(0) = 0$ ). Its size increases by  $\mu_i^{\max}$  if and only if worker  $i$ 's pending task queue is not empty at time slot  $t$  and worker  $i$  rests during this time slot. Its size decreases in the same way as  $q_i(t)$ .

For simplicity of notation, we denote  $\mu_i^{\max}1_{[q_i(t)>0 \& \mu_i(t)=0]}$  as  $x_i(t)$ . Then, equation (3) can be re-expressed as:

$$Q_i(t+1) \geq Q_i(t) + x_i(t) - \mu_i(t). \quad (4)$$

By re-arranging the above inequality, we have:

$$Q_i(t+1) - Q_i(t) \geq x_i(t) - \mu_i(t). \quad (5)$$

Summing both sides of the above inequality over all  $t \in \{0, \dots, T-1\}$  yields:

$$\sum_{t=0}^{T-1} [Q_i(t+1) - Q_i(t)] \geq \sum_{t=0}^{T-1} [x_i(t) - \mu_i(t)]. \quad (6)$$

Thus, we have:

$$Q_i(T) - Q_i(0) \geq \sum_{t=0}^{T-1} [x_i(t) - \mu_i(t)]. \quad (7)$$

Since  $Q_i(0) = 0$ , the above inequality is simplified as:

$$\frac{Q_i(T)}{T} \geq \frac{1}{T} \sum_{t=0}^{T-1} x_i(t) - \frac{1}{T} \sum_{t=0}^{T-1} \mu_i(t). \quad (8)$$

From equation (8), it can be observed that the effect of the conceptual queue is to signal the scheduling approach as to when the need to reduce pending workload shall take precedence over helping workers plan their rest. By ensuring that the computed  $\mu_i(t)$  values satisfy the queueing stability requirement of  $\frac{1}{T} \sum_{t=0}^{T-1} \mu_i(t) \geq \frac{1}{T} \sum_{t=0}^{T-1} x_i(t)$  for the conceptual queue, a scheduling approach will ensure that tasks do not stay pending indefinitely.

By jointly considering  $q_i(t)$  and  $Q_i(t)$ , the *Lyapunov function* (Neely 2010) which measures the overall concentration of work among workers in a given system during time slot  $t$  is:

$$L(t) = \frac{1}{2} \sum_{i=1}^N [q_i^2(t) + Q_i^2(t)]. \quad (9)$$

A large value of  $L(t)$  indicates that tasks are highly concentrated among a small number of workers and/or that tasks have been pending for a long period of time. Both of these scenarios are undesirable from a system productivity perspective and shall be avoided as much as possible. The constant term  $\frac{1}{2}$  is included to simplify subsequent derivations without affecting the physical meaning of the formulation. We adopt the time-averaged *Lyapunov drift*,  $\Delta = \frac{1}{T} \sum_{t=0}^{T-1} [L(t+1) - L(t)]$ , to measure the changes in the degree of seriousness of these two scenarios over time.

We formulate a joint  $\{\text{effort output} + \text{drift}\}$  optimization objective function as:

$$\phi \mathbb{E}\{\tilde{\xi}(t)|\tilde{q}(t), \tilde{m}(t)\} + \Delta \quad (10)$$

which shall be minimized.  $\phi > 0$  controls the emphasis placed on conserving worker effort compared to getting more work done. A larger  $\phi$  value can be interpreted as stronger emphasis on allowing workers to rest more. This value can be set by the system operators to express system-level preference on how to utilize the collective productivity of the workers.  $\tilde{\xi}(t)$ ,  $\tilde{q}(t)$  and  $\tilde{m}(t)$  are vectors containing the workers' effort output values, the backlog queue sizes, and the self-reported mood during time slot  $t$ , respectively.

Based on equation (1) and equation (9), the time-averaged

Lyapunov drift can be expressed as:

$$\begin{aligned} \Delta &= \frac{1}{T} \sum_{t=0}^{T-1} \sum_{i=1}^N \left[ \left( \frac{1}{2} q_i^2(t+1) \right. \right. \\ &\quad \left. \left. - \frac{1}{2} q_i^2(t) \right) + \left( \frac{1}{2} Q_i^2(t+1) - \frac{1}{2} Q_i^2(t) \right) \right] \\ &= \frac{1}{T} \sum_{t=0}^{T-1} \sum_{i=1}^N \left[ \left( \frac{1}{2} \max[0, q_i(t) + \lambda_i(t) - \right. \right. \\ &\quad \left. \left. \mu_i(t)]^2 - \frac{1}{2} q_i^2(t) \right) + \left( \frac{1}{2} \max[0, Q_i(t) + \right. \right. \\ &\quad \left. \left. \mu_i^{\max} 1_{[q_i(t)>0 \& \mu_i(t)=0]} - \mu_i(t)]^2 - \frac{1}{2} Q_i^2(t) \right) \right] \quad (11) \\ &\leq \frac{1}{T} \sum_{t=0}^{T-1} \sum_{i=1}^N \left[ \left( q_i(t)[\lambda_i(t) - \mu_i(t)] \right. \right. \\ &\quad \left. \left. - \mu_i(t)\lambda_i(t) + \frac{1}{2} [\lambda_i^2(t) - 2\lambda_i(t)\mu_i(t) + \mu_i^2(t)] \right) \right. \\ &\quad \left. + \left( Q_i(t)[\mu_i^{\max} 1_{[q_i(t)>0 \& \mu_i(t)=0]} - \mu_i(t)] + \frac{1}{2} [(\mu_i^{\max})^2 1_{[q_i(t)>0 \& \mu_i(t)=0]} - \right. \right. \\ &\quad \left. \left. 2\mu_i^{\max} 1_{[q_i(t)>0 \& \mu_i(t)=0]} \mu_i(t) + \mu_i^2(t)] \right) \right]. \end{aligned}$$

This formulation enables simultaneous modelling of the absolute sizes of the real and the conceptual queues, the distribution of congestions among these queues, and the fluctuations of these queues over time. All three quantities should be minimized in order to optimize our design objectives. In this way, we translate system-level efficiency and worker wellbeing requirements into queueing system stability concepts. Then, through maintaining queue system stability, C-PL achieves these desirable objectives.

Since neither  $\lambda_i(t)$  nor  $\mu_i(t)$  can be infinite in real-world systems, we can simplify  $\Delta$  as:

$$\begin{aligned} \Delta &\leq \frac{1}{T} \sum_{t=0}^{T-1} \sum_{i=1}^N \left[ \left( q_i(t)[\lambda_i(t) - \mu_i(t)] \right. \right. \\ &\quad \left. \left. - \mu_i(t)\lambda_i(t) + \frac{1}{2} [\lambda_{\max}^2 + \mu_{\max}^2] \right) \right. \\ &\quad \left. + \left( Q_i(t)[\mu_{\max} 1_{[q_i(t)>0 \& \mu_i(t)=0]} - \mu_i(t)] + \frac{1}{2} [\mu_{\max}^2 1_{[q_i(t)>0 \& \mu_i(t)=0]} + \mu_{\max}^2] \right) \right]. \quad (12) \end{aligned}$$

where  $\lambda_{\max} \geq \lambda_i(t)$  and  $\mu_{\max} \geq \mu_i(t)$  for all  $i$  and  $t$  are constant values. As we only aim to influence  $\mu_i(t)$  with recommendations, we only retain terms containing  $\mu_i(t)$ . In this way, by substituting equation (12) into equation (10), we obtain the following objective function:

Minimize:

$$\frac{1}{T} \sum_{t=0}^{T-1} \sum_{i=1}^N \xi_i(t)[\phi - (q_i(t) + Q_i(t))m_i(t)\mu_i^{\max}] \quad (13)$$

Subject to:

$$0 \leq \xi_i(t) \leq 1, \forall i, \forall t \quad (14)$$

$$0 \leq \mu(\xi_i(t), m_i(t)) \leq \mu_i^{\max}, \forall i, \forall t \quad (15)$$

By minimizing equation (13) subject to Constraints (14) and (15), we simultaneously minimize the time-averaged total worker effort output while maximizing the time-averaged collective productivity in a given system. To minimize equation (13), at each time slot  $t$ , we need to compute the values of the expression  $[\phi - (q_i(t) + Q_i(t))m_i(t)\mu_i^{\max}]$  for all  $i$ . For simplicity of notation, we denote  $[\phi - (q_i(t) + Q_i(t))m_i(t)\mu_i^{\max}]$  as the *Work-Rest Index (WRI)*,  $\Psi_i(t)$ . This index enables a scheduling agent to efficiently search through a very large solution space to determine if the current situation is more suitable for work or rest.

### Opportunistic Work-Rest Scheduling

Algorithm 1 presents a distributed implementation of the CPL approach as a scheduling agent for each worker. For a worker  $i$ , if  $\Psi_i(t) < 0$ , it sets  $\xi_i(t) = \min\left[1, \frac{q_i(t)}{m_i(t)\mu_i^{\max}}\right]$  and computes the corresponding  $\mu_i(t)$  value; otherwise, it advises the worker to rest for the current time slot. The computational time complexity of Algorithm 1 is  $O(1)$ , making it highly scalable. The algorithm implements computationally the intuition that *the more pending tasks in a worker's backlog queue, the longer these tasks have been pending, the more emphasis on worker wellbeing by system operators, and the higher the worker's current mood, the more effort should be expended towards completing tasks (subject to the physical limitations of the worker's effort output per time slot)*.

---

#### Algorithm 1 CPL

---

**Require:**  $\phi, q_i(t), \mu_i^{\max}$  and  $m_i(t)$ .

- ```

1: if  $\Psi_i(t) < 0$  then
2:    $\xi_i(t) = \min\left[1, \frac{q_i(t)}{m_i(t)\mu_i^{\max}}\right]$ ;
3:   Compute  $\mu_i(t)$  according to equation (2);
4: else
5:    $\xi_i(t) = 0$ ;
6:    $\mu_i(t) = 0$ ;
7: end if
8: return  $\mu_i(t)$ ;

```
- 

WRI incorporates the ethical considerations (Yu et al. 2018) by prioritizing worker wellbeing considerations through recommending opportunistic rest breaks, and allowing stakeholders to influence the AI recommendations by expressing their preferences through  $\phi$  and  $m_i(t)$ . The recommendations from CPL are given to a worker in the form of the number of tasks he should complete over a given time slot (a 0 value indicates that the worker should rest) so as to make it actionable enough for the worker to follow.

### Experimental Evaluation

To evaluate the performance of CPL under realistic settings, we compare it against four alternative approaches

through extensive simulations. The characteristics of worker agents in the simulation are derived from the *Tianchi* dataset (<http://dx.doi.org/10.7303/syn7373599>) released by Alibaba. This real-world dataset contains information regarding 5,547 workers' reputation (i.e. quality of work) and maximum productivity (i.e. the maximum number of tasks a worker can complete per time slot). This dataset allows us to construct realistic simulations.

### Experiment Settings

The five comparison approaches are:

1. The *Max Effort (ME)* approach: under this approach, a worker agent  $i$  always works as long as there are tasks in its backlog queue regardless of its mood.
2. The *Mood Threshold (MT)* approach: under this approach, a worker agent  $i$  works whenever  $m_i(t) \geq \theta_1$  and there are tasks in its backlog queue, where  $\theta_1 \in [0, 1]$  is a pre-determined mood threshold used by MT.
3. The *Mood and Workload threshold (MW)* approach: this approach jointly considers a worker agent  $i$ 's current mood and workload to determine how much effort to exert. Whenever  $q_i(t)\mu_i(1, m_i(t)) \geq \mu_i^{\max}\mu_i(1, \theta_2)$ , worker  $i$  exerts up to the maximum effort subject to there being enough tasks in its backlog queue, where  $\theta_2 \in [0, 1]$  is a predetermined mood threshold used by MW.
4. The *Affective Crowdsourcing (AC)* approach (Yu et al. 2017b) which is similar to CPL but does not take task pending time into account.
5. The *CPL* approach proposed in this paper.

The approach used to delegate tasks to worker agents under all comparison approaches is SMVM (Yu et al. 2017a). It dynamically distributes tasks among workers in a situation-aware manner in order to avoid over concentration of workload. At each time slot, SMVM determines how many tasks to delegate to each worker agent  $i$  in the simulations (i.e. SMVM computes  $\lambda_i(t)$  for all  $i$  and  $t$ ) based on its current reputation and workload. The principle implemented by SMVM is that the higher a worker agent's reputation and the lower its current workload, the more tasks should be delegated to it. SMVM can be replaced by any other approach (Ho, Jabbari, and Vaughan 2013; Basu Roy et al. 2015) as long as such an approach can determine the values of  $\lambda_i(t)$  for all  $i$  and  $t$ .

In order to create different experiment conditions, we vary the value of  $\phi$  between 5 and 100 in increments of 5. The values of  $\theta_1$  and  $\theta_2$  are varied between 0.05 and 1 in increments of 0.05. The system workload is measured in relation to the maximum collective productivity of the worker agent population,  $\Omega = \sum_{i=1}^N r_i \mu_i^{\max}$ . In this equation,  $r_i$  is a worker agent  $i$ 's reputation and  $N = 5,547$ . We adopt the concept of load factor (LF) from (Yu et al. 2016; 2017a) to denote the overall workload placed on the system. It is computed as the ratio between the number of new tasks delegated to the worker agents during time slot  $t$ ,  $W_{req}(t)$ , and the maximum collective productivity  $\Omega$  of the system (i.e.  $LF = \frac{W_{req}(t)}{\Omega}$ ). We vary LF between 5% to 100% in 5% increments.

Throughout the experiments, the mood for each worker agent  $i$  during time slot  $t$ ,  $m_i(t)$ , is randomly generated in the range of  $[0, 1]$  following a uniform distribution. This eliminates the possibility for any of the comparison approaches to predict a worker agent's future mood based on previous observations, thereby focusing the experimental comparisons on the effectiveness of the scheduling strategies. Under each LF setting, the simulation is run for  $T = 10,000$  time slots. All tasks must be completed within 3 time slots after they have been delegated.

## Evaluation Metrics

The performances of the five approaches in the experiments are compared using the following metrics:

1. The time-averaged worker effort output,  $\bar{\xi} = \frac{1}{TN} \sum_{t=0}^{T-1} \sum_{i=1}^N \xi_i(t)$ . The smaller the  $\bar{\xi}$  value, the better the performance of an approach.
2. The time-averaged task expiry rate,  $\bar{e} = \frac{1}{TN} \sum_{t=0}^{T-1} \sum_{i=1}^N \frac{n_i^{(e)}(t)}{q_i(t)}$ , where  $n_i^{(e)}(t)$  is the number of tasks in  $i$ 's backlog which passed their deadlines during a given time slot  $t$ . The smaller the  $\bar{e}$  value, the better the performance of an approach.
3. The time-averaged task completion rate,  $\bar{\mu} = \frac{1}{T} \sum_{t=0}^{T-1} \frac{\sum_{i=1}^N \mu_i(t)}{N_{\text{total}}(t)}$ , where  $N_{\text{total}}(t)$  is the total number of tasks waiting to be completed in the system during a given time slot  $t$ . The larger the  $\bar{\mu}$  value, the better the performance of an approach.

Since worker agents under the ME approach consistently expend the most effort and achieve the highest task completion rate, we use ME as the baseline for comparing the performance of other approaches under different LF,  $\phi$ ,  $\sigma$ ,  $\theta_1$  and  $\theta_2$  settings.

## Results and Discussions

Figures 1(a)–1(d) show the time-averaged task expiry rates achieved by MT, MW, AC and CPL respectively under various experiment settings. As MT uses mood as the threshold to control worker effort output, the changes in task expiry are directly related to mood values (Figure 1(a)). On average, 29% of the tasks under MT expire before they can be completed. MW is also a threshold-based approach. However, its threshold consists of a combination of workers' mood and their current workload. Thus, its task expiry rate increases with both mood and LF with the effect of mood being more significant (Figure 1(b)). On average, 30% of the tasks under MW expire before they can be completed. AC is not a threshold-based scheduling approach. A worker can indicate to AC his desire to rest by adjusting the value of the control variable  $\sigma$ . As  $\sigma$  and LF values increase, an increasing percentage of tasks expire under AC. On average, 7.5% of the tasks under AC expire before they can be completed, which is significantly lower than MT and MW. As AC only considers mood and workload when making work-rest recommendations, workers may fall into the condition in which their mood and workload trigger AC to recommend resting. However, their workload is also not low enough to

cause the task delegation approach to delegate new tasks to them. Therefore, AC continues to recommend resting until pending tasks pass their deadlines and become expired. This problem is addressed by CPL as it takes task pending time into account with the conceptual queue technique when optimizing work-rest scheduling. As shown in Figure 1(d), increases in  $\phi$  and LF values result in an increasing percentage of tasks expire under CPL. The task expiry rate under CPL is lower than that under AC. On average, 5.4% of the tasks under CPL expire before they can be completed, which is significantly lower than MT, MW and AC considering the scale of the experiment.

The time-averaged effort output  $\bar{\xi}$  achieved by all five approaches is shown in Figure 1(e). Compared to ME, all other approaches achieved significant savings in effort as LF increases. This is partially due to the SMVM task delegation approach used in the simulations. When LF is low, tasks are mostly concentrated on worker agents with good reputation. In this case, the task backlogs of individual worker agents who have been delegated tasks tend to be relatively high, which makes scheduling approaches allocate less time for these workers to rest in order to meet task deadlines. As LF increases, the workload is spread more evenly among a larger segment of the worker agent population, creating more opportunities for scheduling approaches to slot in rest breaks. The  $\bar{\xi}$  values achieved by MT, MW and AC stabilize between 20% and 40% while that of CPL stabilizes around 60%. MW achieves the lowest worker agent effort output. The time-averaged task completion rates  $\bar{\mu}$  achieved by all five approaches are shown in Figure 1(f). It can be observed that CPL achieves the highest  $\bar{\mu}$  values which stabilize around 85% and are higher than AC, MT and MW by 22%, 90% and 108%, respectively.

Figure 1(g) shows the performance landscape of MT, MW, AC and CPL as a percentage of ME under different control parameter settings. MW and MT both use mood thresholds ( $\theta_1$  and  $\theta_2$ , respectively) to control effort output. The higher the mood threshold values, the lower the effort output (and hence the task completion rates) achieved by MW and MT. On the other hand, mood serves as one of the inputs to AC and CPL. Both AC and CPL allow a worker agent to specify a general emphasis on conserving effort using variables  $\sigma$  and  $\phi$ , respectively. The larger the values of these variables, the more emphasis is placed on effort conservation. By varying the  $\phi$  value in CPL, we can control the trade-off between worker effort output and task completion rate (from spending 92% of the ME effort output and achieving 99% of the ME task completion rate, to spending 53% of the ME effort output and achieving 78% of the ME task completion rate). CPL consistently and significantly outperforms both MT and MW, conserving significant worker effort while achieving high task throughput. In effect, CPL limits the range of trade-off between work and rest achieved by AC based on consideration of an additional situational factor – the task pending time – in order to achieve better collective performance. In the worst case scenario in which  $\theta_1$  and  $\theta_2$  are set to 1, indicating that workers are unwilling to work under any mood condition, the worker effort output and the task completion rates achieved by both MT and MW

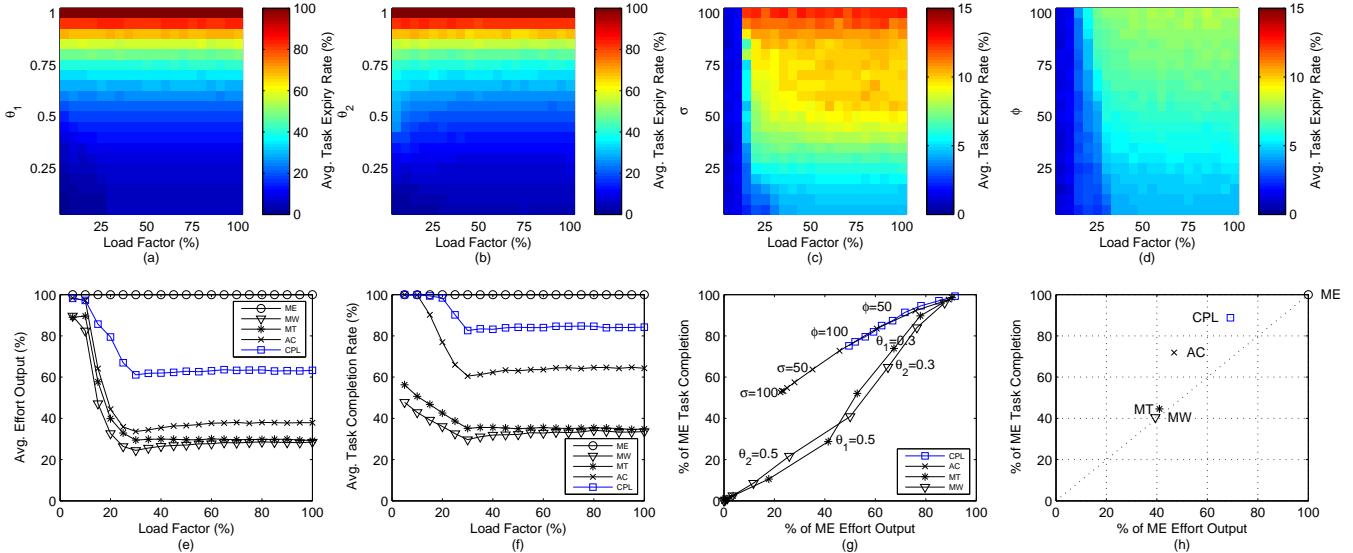


Figure 1: Experiment results: **(a)** the time-averaged task expiry rates achieved by MT under various  $\theta_1$  and LF settings; **(b)** the time-averaged task expiry rates achieved by MW under various  $\theta_2$  and LF settings; **(c)** the time-averaged task expiry rates achieved by AC under various  $\sigma$  and LF settings; **(d)** the time-averaged task expiry rates achieved by CPL under various  $\phi$  and LF settings; **(e)** comparison of the time-averaged effort output achieved by various approaches under different LF settings; **(f)** comparison of the time-averaged task completion rates achieved by various approaches under different LF settings; **(g)** the time-averaged task completion rates vs. the time-averaged effort output achieved by various approaches under different parameter settings as a percentage of those achieved by ME; **(h)** the time-averaged task completion rate vs. the time-averaged effort output achieved by various approaches averaged over different parameter settings as a percentage of those achieved by ME.

are 0. This is expected as in MT and MW, mood is used as the threshold to control effort output. However, under AC, mood is only one of the situational factors considered by the approach. CPL adds in task pending time on top of the situational factors used by AC to make scheduling decisions. Even under scenarios in which  $\phi$  is set to 100 (indicating that workers place very high emphasis on rest), CPL is still able to maintain a long-term average effort output of 53% taking advantage of favourable working conditions whenever possible to achieve average task completion rates of about 78%. This is more advantageous from a business perspective compared to the worst case performance by AC of about 50%.

When we compute the averages of the results shown in Figure 1(g) over their respective setting variables (i.e.  $\phi$ ,  $\sigma$ ,  $\theta_1$  and  $\theta_2$ ), we obtain Figure 1(h) showing the overview of their performances. The diagonal dotted line represents linear productivity, meaning that an increase in effort output results in a directly proportional increase in collective productivity. It can be observed that ME, MT and MW all fall on the linear productivity line, whereas AC and CPL are significantly above this line in a region of superlinear productivity (Sornette, Maillart, and Ghezzi 2014). Under AC and CPL, an increase in effort output results in a disproportionately larger increase in collective productivity, indicating that the collective productivity achieved is larger than the sum of individual workers' productivity. Overall, CPL achieves 89% of ME task completion rate with 69% of the ME worker effort output, which is the most desirable work-rest trade-off among the five approaches from a system perspective.

## Conclusions and Future Work

Improving collective productivity is an important problem facing many social and economic systems. How to dynamically adapt workers' work-rest schedules in response to changing situations in order to maintain a high level of productivity and worker wellbeing remains an open research question. The proposed CPL approach translates considerations on workers' mood, workload and pending time of the tasks in their backlog into actionable personalized work-rest schedules. It establishes a framework to model complex relationships between work and rest, and helps workers optimize the balance between work and rest in order to achieve superlinear collective productivity. Taking into account its polynomial time complexity, CPL is an effective and scalable approach to help workers benefit from opportunistic rest. By nudging workers to be 'lazy' at opportune times, CPL achieves collective productivity which is larger than the sum of individual productivity. It provides a way to design ethically aligned workforce management systems that sustain long-term effective participation by promoting productive laziness among workers.

In future research, we plan to testbed CPL in a crowdsourcing platform (Pan et al. 2016) to reach out to more diverse users and study how to improve the approach in the presence of various behaviour patterns and how to foster trust (Shen et al. 2011) with user by explaining the rationale behind the recommendations.

## Acknowledgments

This research is supported, in part, by Nanyang Technological University, Nanyang Assistant Professorship (NAP) and the National Research Foundation, Prime Minister's Office, Singapore under its IDM Futures Funding Initiative.

## References

- Basu Roy, S.; Lykourentzou, I.; Thirumuruganathan, S.; Amer-Yahia, S.; and Das, G. 2015. Task assignment optimization in knowledge-intensive crowdsourcing. *VLDB Journal* 24(4):467–491.
- Chai, Y.; Miao, C.; Sun, B.; Zheng, Y.; and Li, Q. 2017. Crowd science and engineering: concept and research framework. *International Journal of Crowd Science* 1(1):2–8.
- Chen, C.; Cheng, S.-F.; Lau, H. C.; and Misra, A. 2015. Towards city-scale mobile crowdsourcing: Task recommendations under trajectory uncertainties. In *IJCAI*, 1113–1119.
- Dai, P.; Lin, C. H.; Mausam; and Weld, D. S. 2013. POMDP-based control of workflows for crowdsourcing. *Artificial Intelligence* 202:52–85.
- Dai, P.; Rzeszotarski, J. M.; Paritosh, P.; and Chi, E. H. 2015. And now for something completely different: Improving crowdsourcing workflows with micro-diversions. In *CSCW*, 628–638.
- Grossmann, I.; Brienza, J. P.; and Bobocel, D. R. 2017. Wise deliberation sustains cooperation. *Nature Human Behaviour* 1(0061):doi:10.1038/s41562-017-0061.
- Heymann, P., and Garcia-Molina, H. 2011. Turkalytics: Analytics for human computation. In *WWW*, 477–486.
- Ho, C.-J.; Jabbari, S.; and Vaughan, J. W. 2013. Adaptive task assignment for crowdsourced classification. In *ICML*, I-534–I-542.
- Lee, M. K.; Kusbit, D.; Metsky, E.; and Dabbish, L. 2015. Working with machines: The impact of algorithmic and data-driven management on human workers. In *CHI*, 1603–1612.
- Leiter, M. P., and Maslach, C. 2015. Conquering burnout. *Scientific American Mind* 26:30–35.
- Mao, A.; Kamar, E.; Chen, Y.; Horvitz, E.; Schwamb, M. E.; Lintott, C. J.; and Smith, A. M. 2013. Volunteering versus work for pay: Incentives and tradeoffs in crowdsourcing. In *HCOMP*, 94–102.
- Mason, W., and Watts, D. J. 2012. Collaborative learning in networks. *Proceedings of the National Academy of Sciences USA* 109:764–769.
- Miao, C.; Yu, H.; Shen, Z.; and Leung, C. 2016. Balancing quality and budget considerations in mobile crowdsourcing. *Decision Support Systems* 90:56–64.
- Michelucci, P., and Dickinson, J. L. 2016. The power of crowds. *Science* 351(6268):32–33.
- Morris, R. R.; Dontcheva, M.; and Gerber, E. M. 2012. Priming for better performance in microtask crowdsourcing environments. *IEEE Internet Computing* 16(5):13–19.
- Neely, M. J. 2010. *Stochastic Network Optimization with Application to Communication and Queueing Systems*. Morgan and Claypool Publishers.
- Oswald, A.; Proto, E.; and Sgroi, D. 2015. Happiness and productivity. *Journal of Labor Economics* 33(4):789–822.
- Pan, Z.; Yu, H.; Miao, C.; and Leung, C. 2016. Efficient collaborative crowdsourcing. In *AAAI*, 4248–4249.
- Shen, Z.; Yu, H.; Miao, C.; and Weng, J. 2011. Trust-based web-service selection in virtual communities. *Web Intelligence and Agent Systems* 9(3):227–238.
- Sornette, D.; Maillart, T.; and Ghezzi, G. 2014. How much is the whole really more than the sum of its parts?  $1 + 1 = 2.5$ : Superlinear productivity in collective group actions. *PLoS ONE* 9(8):doi:10.1371/journal.pone.0103023.
- The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. 2018. Ethically aligned design: A vision for prioritizing human well-being with autonomous and intelligent systems. Technical report, IEEE.
- Tran-Thanh, L.; Huynh, T. D.; Rosenfeld, A.; Ramchurn, S. D.; and Jennings, N. R. 2014a. BudgetFix: Budget limited crowdsourcing for interdependent task allocation with quality guarantees. In *AAMAS*, 477–484.
- Tran-Thanh, L.; Stein, S.; Rogers, A.; and Jennings, N. R. 2014b. Efficient crowdsourcing of unknown experts using bounded multi-armed bandits. *Artificial Intelligence* 214:89–111.
- Whillans, A. V.; Dunn, E. W.; Smeets, P.; Bekkers, R.; and Norton, M. I. 2017. Buying time promotes happiness. *Proceedings of the National Academy of Sciences USA* 114:8523–8527.
- Xu, Z.; Li, Z.; Guan, Q.; Zhang, D.; Li, Q.; Nan, J.; Liu, C.; Bian, W.; and Ye, J. 2018. Large-scale order dispatch in on-demand ride-hailing platforms: A learning and planning approach. In *KDD*, 905–913.
- Yu, H.; Shen, Z.; Miao, C.; and Tan, A.-H. 2011. A simple curious agent to help people be curious. In *AAMAS*, 1159–1160.
- Yu, H.; Miao, C.; An, B.; Leung, C.; and Lesser, V. R. 2013a. A reputation management model for resource constrained trustee agents. In *IJCAI*, 418–424.
- Yu, H.; Shen, Z.; Leung, C.; Miao, C.; and Lesser, V. R. 2013b. A survey of multi-agent trust management systems. *IEEE Access* 1(1):35–50.
- Yu, H.; Miao, C.; Shen, Z.; Leung, C.; Chen, Y.; and Yang, Q. 2015. Efficient task sub-delegation for crowdsourcing. In *AAAI*, 1305–1311.
- Yu, H.; Miao, C.; Leung, C.; Chen, Y.; Fauvel, S.; Lesser, V. R.; and Yang, Q. 2016. Mitigating herding in hierarchical crowdsourcing networks. *Scientific Reports* 6(4):doi:10.1038/s41598-016-0011-6.
- Yu, H.; Miao, C.; Chen, Y.; Fauvel, S.; Lesser, V. R.; and Yang, Q. 2017a. Algorithmic management for improving collective productivity in crowdsourcing. *Scientific Reports* 7(12541):doi:10.1038/s41598-017-12757-x.
- Yu, H.; Shen, Z.; Fauvel, S.; and Cui, L. 2017b. Efficient scheduling in crowdsourcing based on workers' mood. In *IEEE ICA*, 121–126.
- Yu, H.; Shen, Z.; Miao, C.; Leung, C.; Lesser, V. R.; and Yang, Q. 2018. Building ethics into artificial intelligence. In *IJCAI*, 5527–5533.
- Zeng, Z.; Tang, J.; and Wang, T. 2017. Motivation mechanism of gamification in crowdsourcing projects. *International Journal of Crowd Science* 1(1):71–82.
- Zenonos, A.; Stein, S.; and Jennings, N. R. 2016. An algorithm to coordinate measurements using stochastic human mobility patterns in large-scale participatory sensing settings. In *AAAI*, 3936–3942.
- Zimasa, T.; Jamson, S.; and Henson, B. 2017. Are happy drivers safer drivers? Evidence from hazard response times and eye tracking data. *Transportation Research Part F: Traffic Psychology and Behaviour* 46(A):14–23.