Group-Oriented Task Allocation for Crowdsourcing in Social Networks

Jiuchuan Jiang, Bo An, Yichuan Jiang, *Senior Member, IEEE*, Chenyan Zhang, Zhan Bu, and Jie Cao

***Abstract*—Previous crowdsourcing studies often adopted the individual-oriented approach that outsources a task to an individual worker or team formation-based approach that outsources a task to an artificially formed team of workers. Nowadays, workers are often naturally organized into groups through social networks. To address such common issue of grouped workers in real crowdsourcing systems, this article explores a novel crowdsourcing paradigm in which the task allocation targets are naturally existing worker groups but not individual workers or artificially formed teams as before. Because a natural group might not possess all required skills and needs to coordinate with other groups in the social network contexts for performing a complex task, a concept of contextual crowdsourcing value is presented to measure a group's capacity to complete a task by coordinating with its contextual groups, which determines the priority that the group is assigned the task; then, the task allocation algorithms, including the allocations of groups and the workers actually participating in executing the task, are designed. The experiments on a real-world dataset show that our presented group-oriented approach can nearly always achieve better synergy performance, consistency performance, conflict performance, adaptability, and effectiveness on reducing costs, as compared with previous benchmark individual-oriented and team formation approaches.**

***Index Terms*—Context-aware, crowdsourcing, group, social networks, task allocation.**

I. Introduction

MANY task allocation approaches in crowdsourcing often care about how to satisfy the skill requirements of tasks, since skill requirement viewpoint is a generally used method to frame tasks [6], [40]. Individual-oriented crowdsourcing approach has been widely used in previous studies, in which the personal skills of individual workers are cared about and each assigned worker can complete the task individually and independently [1]–[3], [41]. To facilitate cooperation among workers to perform complex tasks, the team formation approach, in which the requester seeks a team of workers to perform a complex task that requires various skills, has recently been explored [4], [5]. However, each time a complex task is published, a new team should be formed from scratch to satisfy the skill requirements of the task; thus, such a mechanism will bring heavy task allocation costs in crowdsourcing markets in which tasks are numerous and dynamic [23], [27].

In reality, people are often naturally organized into groups through social connections [3], [6], [10], [20], [32]. The natural group is different from the artificially formed team because the former does naturally exist in social networked crowds, but the latter is formed artificially and transiently for a special task. The phenomenon of grouped workers is common in real crowdsourcing systems. For example, at `www.upwork.com`, we find that the workers affiliated with any groups (i.e., agency freelancers) can constitute 47% of the total workers; at the Github website, there were 226 449 groups registered from January 1, 2016 to June 30, 2016.

To address the above common issue, this article presents a novel group-oriented crowdsourcing paradigm, in which the task allocation targets are naturally existing worker groups but not individual workers or artificial teams. With this crowdsourcing paradigm, members in the assigned group will cooperate to perform the complex task. Because naturally organized groups do not need to be formed from scratch, they can better fit the crowdsourcing markets with numerous tasks than can teams tailored for specific tasks.

However, because groups are organized naturally, an allocated group might occasionally not possess the complete skills of the task. Moreover, workers are often connected through social networks [7], [8]. Thus, the workers of one group can coordinate with other workers of the contextual groups for assistance [9], in which the context of a group primarily means the counterpart groups interacting with this group through the social network. For example, at GitHub, an assigned group lacks HTML5 skill required by a task "design a responsive website;" thus, the group needs to coordinate with another contextual group possessing HTML5 skill.

To consider the coordination between the assigned groups and their contextual groups, this article presents a

Manuscript received August 19, 2018; revised November 28, 2018, March 25, 2019, and June 7, 2019; accepted July 19, 2019. Date of publication August 28, 2019; date of current version June 16, 2021. This work was supported in part by the National Natural Science Foundation of China under Grant 61932007, Grant 61472079, Grant 61806053, and Grant 61807008, and in part by the Natural Science Foundation of Jiangsu Province of China under Grant BK20171363, Grant BK20180356, and Grant BK20180369. This article was recommended by Associate Editor E. Chen. *(Corresponding author: Jiuchuan Jiang.)*

J. Jiang, Z. Bu, and J. Cao are with the Jiangsu Provincial Key Laboratory of E-Business, College of Information Engineering, Nanjing University of Finance and Economics, Nanjing 210003, China (e-mail: jiangjiuchuan@163.com).

B. An is with the School of Computer Science and Engineering, Nanyang Technological University, Singapore 639798.

Y. Jiang is with the School of Computer Science and Engineering, Southeast University, Nanjing 211189, China.

C. Zhang is with the School of Computer Science and Engineering, Southeast University, Nanjing 211189, China, and also with the Co-Innovation Center of Shandong Colleges and Universities for Future Intelligent Computing, Shandong Technology and Business University, Yantai 264005, China.

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TSMC.2019.2933327.

Digital Object Identifier 10.1109/TSMC.2019.2933327

context-aware approach for group-oriented task allocation in crowdsourcing, in which a group candidate's self-situation and its contextual-situation in the social network are considered when the requester wishes to assign a task to such group. Generally, the main challenges in this context-aware approach include: 1) how to measure the priority of a group being allocated a task by considering its contextual groups and 2) how an assigned group coordinates with other contextual groups in performing the task.

To solve the main challenges stated above, at first we present a metric of contextual crowdsourcing value of a group to measure the group's capacity to perform a task by coordinating with its contextual groups; the higher a group's contextual crowdsourcing value is, the more preferentially that the groups is assigned the task. Then, we present a method to model the coordination among groups for performing the task. Finally, we present the task allocation algorithms that include: 1) assigning the task to a principal group according to the candidate groups' contextual crowdsourcing values; 2) allocation of assistant groups if the principal group cannot complete the task along by itself; and 3) selection of workers actually participating in executing task from the principal and assistant groups.

Finally, the experiments are conducted on real-world datasets extracted from GitHub by comparing with previous benchmark individual-oriented and team formation approaches, which show that our presented approach can nearly always achieve better synergy performance, consistency performance, conflict performance, adaptability, and effectiveness on reducing costs.

As far as we know, this article is the first work that investigates how to outsource a complex task to a natural worker group. In summary, the main originalities and contributions of this article are shown as follows.

1) A novel group-oriented crowdsourcing paradigm is presented, in which the task allocation targets are naturally existing worker groups but not individual workers or artificial teams.
2) The context-aware task allocation problem in group-oriented crowdsourcing is formally defined and proven to be NP-hard.
3) A heuristic context-aware task allocation approach is presented that is based on the concept of contextual crowdsourcing value and can be realized within a limited time complexity.
4) A modeling method for the natural worker groups in crowdsourcing is presented, which include the one for the groups with leaders and the one for the groups without leaders.
5) Comprehensive experiments are conducted and show the advantages of our presented group-oriented approach as compared with previous benchmark approaches.

II. Related Work

A. Individual-Oriented Crowdsourcing

Many previous studies adopted an individual-oriented crowdsourcing approach; thus, each worker can complete the allocated task individually and independently. Generally, there are two types of tasks in crowdsourcing; one is *micro-task*, and the other is *complex task*.

The micro-tasks are atomic computation operations and can be completed in minutes by nonprofessional individual workers [11], [12]. In fact, many traditional crowdsourcing systems, such as Amazon's Mechanical Turk, are often designed for micro-task markets [26]. In the crowdsourcing of micro-tasks, each task is redundantly allocated to more than one individual worker to improve accuracy, and each worker executes the task independently; finally, the requester will select the correct result from the multiple answers from the redundant allocated individual workers [11].

Complex tasks are involved in many computation operations and need multiple skills. To implement complex tasks in those micro-task-suited crowdsourcing systems, a popular method is to decompose each complex task into a flow of simple subtasks and then combine the partial results of subtasks together to obtain the final answer [2]. Therefore, an efficient task decomposition method is necessary for the crowdsourcing of complex tasks [1]. A representative work is that Tran-Thanh *et al.* [2] proposed the BudgetFix algorithm to solve complex tasks that involve different types of interdependent micro-tasks structured into complex workflows.

In summary, the allocation targets of tasks in many traditional crowdsourcing systems are oriented to individual workers; then, each allocated worker executes the micro-task (or the micro-task decomposed from a complex task) independently. In comparison, the approach in this article directly allocates the complex task to a group, and the workers in the allocated group cooperate to complete the complex task.

B. Team Formation-Based Crowdsourcing

Team formation is a new method for crowdsourcing of complex tasks, in which individuals with different skills form a team to complete tasks together. In many existing studies, the team formation of workers is controlled by the requester, in which interested candidate workers advertise their skills and bid a price for their participation into the team. Liu *et al.* [4] presented an efficient method that is implemented through some profitable and truthful pricing mechanisms. Kargar *et al.* [13] presented a team formation method to satisfy the two objectives in social networks: finds a team of experts that covers all the required skills of tasks and minimizes the communication cost between workers in the team.

There are also self-organized team formation studies, in which some workers in the crowd organize a team to bid for the task. Lykourentzou *et al.* [5] presented a self-organized team formation strategy where the workers can select the teammates by themselves. Rokicki *et al.* [14] explored a strategy for team formation in which workers decide by themselves on which team they want to participate. In this strategy, each worker initially forms a one-man team and becomes its administrator.

Note that Chamberlain [15] presented the concept of groupsourcing in which the task is allocated to a group of people of varying expertise connected through a social network.

TABLE I
NOTATIONS

Notation	Definition	Notation	Definition
SN	A worker social network	$v_x^{G_i}(t)$	The self-crowdsourcing value of w_{ix} in G_i (without leader) for task t
W	A crowd of workers	$av_x^{W_{G_i}(t)}$	The crowdsourcing value of w_{ix} for assisting $W_{Gi}(t)$ to perform task t
E	The set of social connections among workers	$v_{G_i}(t)$	The self-crowdsourcing value of G_i for task t
G_i	Worker group i	$Cv_{G_i}(t)$	The contextual crowdsourcing value of G_i for task t
w_{ix}	Worker x in group G_i	$v_{G_j}(G_i-t)$	The assistance value of G_j for G_i on executing t
w_{il}	The leader worker in group G_i	$c_x^{G_i}$	The centrality of w_{ix} in G_i
$C_{ix,iy}$	Communication cost between w_{ix} and w_{jy}	$C_{ix,WGi}$	The communication cost between w_{ix} and $W_{Gi}(t)$
$C_{Gi,Gj}$	The communication cost between groups G_i and G_j	$n_{Gj\leftarrow Gi}$	The historical number of G_i's providing real assistance for G_j's executing tasks
S_{ix}	The set of skills possessed by w_{ix}	$c_{G_i}(\leftarrow G_j)$	The credit of G_i paid by G_j
t	Task t	$m_{G_i\rightarrow G_j}(t)$	The possible monetary reward paid by G_i to G_j
b_t	The budget for task t	$c_{Gi\rightarrow Gj}(t)$	The credit paid by G_i to G_j for executing task t
S_t	The set of skills required by task t	Occupancy rate of groups with leaders	The ratio of groups with leaders to the total groups
$\overline{S_t}$	The set of skills for t that are currently lacking	$d(x, y)$	The shortest path distance between vertices x and y in a graph
$W_{Gi}(t)$	The set of workers in G_i actually participating in executing t	$s(x, y)$	The pairwise synergy value between vertices x and y in a task-based relationship graph
γ_{ix}	The reservation wage of w_{ix}	$S(A)$	The synergy value of the worker set A
$C(W_{G_i}(t))$	The total communication costs between all workers in $W_{Gi}(t)$ and w_{il}	$C(A)$	The consistency of the worker set A
$v_x^{G_{i,l}}(t)$	The self-crowdsourcing value of w_{ix} within group G_i (whose leader is w_{il}) for task t	$APL(A)$	The Harmonic Mean of Average Path Length (HMAPL) of the worker set A in social networks

However, the task allocation mechanism to groups and the task coordination among groups were not systematically investigated in [15].

III. Motivation, Problem Description, and Complexity Analyzes

A. Motivation

Previous studies in economics and management science found that group is a general subsystem in any organization, in which the members in a group work together to do the tasks [39]. Moreover, one of the popular characteristics of social networks is that allowing people to initialize different kinds of groups [34], thus groups are often seen in social networks.

After analyzing the data from some popular crowdsourcing platforms, we find that the phenomenon of people forming groups is common in real crowdsourcing systems. We collected data from www.upwork.com, in which there are two types of freelancers, independent freelancers and agency freelancers. An independent freelancer is an individual worker, whereas an agency freelancer means a group that includes several managers and developers. After randomly counting 9018 workers, we find that there are 4018 workers affiliated with at least one group, i.e., 45% workers are organized into groups. Moreover, we find that there were 226 449 groups registered at https://github.com from January 1, 2016 to June 30, 2016. Although the groups may not be obvious at some crowdsourcing platforms, the existence of grouping can also be shown by analyzing the data of the platforms. For example, at www.mturkforum.com, we can find obvious community structures by constructing the social networks among the active users; thus the workers within a community can also be considered as a group.

These data denotes that groups are quite common at many crowdsourcing websites. To address this common issue of workers forming groups through social networks, a group-oriented approach needs to be explored, which can utilize the cooperation of workers within the groups to perform the complex tasks. Table I summarizes the notations used in this article.

B. Problem Description

Definition 1: Worker groups in a social network. Let there be a worker social network $SN = (W, E)$, where vertices W are workers, and E is the set of social connections among workers. The workers are organized into several groups, $W = \{G_i\}$, each edge $(w_{ix}, w_{jy}) \in E$ represents a social connection between worker w_{ix} (who is attributed to group G_i) and w_{jy} (who is attributed to group G_j); each edge is associated a weight $C_{ix,jy}$ denoting the communication cost between workers w_{ix} and w_{jy}. Each worker $w_{ix} \in W$ possesses a set of skills $\mathcal{S}_{ix}$.

Because groups are organized naturally rather than tailored for any tasks, some problems need to be addressed when we allocate a task to a natural group. First, how can the appropriate group that can mostly match the skill requirements of the task be found. Second, how can the appropriate group that requires fewer payments be found, to save more budgets for the requester and retain more funds to seek the assistance of other groups. Third, how can the appropriate group with lower communication costs be found, because previous benchmark studies have shown that the communication costs between allocated workers will significantly influence the performance in completing the outsourced task [13], [33].

Therefore, group-oriented crowdsourcing should find a group to optimize the following three factors: 1) the degree to which the skills of the workers in the group satisfy the necessary skills required by the task; 2) the ratio of the wages of the workers in the group to the task's budget; and 3) the communication costs among the workers in the group to execute the task. A group might not have the complete set of skills to implement the allocated task, thus the group needs to coordinate with other contextual groups in the social network to obtain the lacking skills. Therefore, the contextual groups' situations should be considered.

Let a crowd of workers be organized into n groups, $G = \{G_i | 1 \le i \le n\}$. Given a complex task t with a budget b_t, the set of skills required by t is S_t. Let the set of workers in G_i actually participating in executing the task be $W_{Gi}(t)$, $W_{Gi}(t) \subseteq G_i$. We use γ_{ix} to denote the reservation wage of w_{ix}, where w_{ix} is a worker in G_i; $C_{ix,iy}$ denotes the communication cost between workers w_{ix} and w_{iy}, and $C_{Gi,Gj}$ denotes the communication cost between groups G_i and G_j in the social network.[1] Let α and 1-α denote the relative importance between a group and its contexts, and β_1, β_2, and β_3 denote the relative importance of the three factors. If we want to focus on one aspect of task assignment for groups, we can set the importance of other aspects to the minimum. The context-aware task allocation objective in group-oriented crowdsourcing for t can be formalized as selecting a group, G_*, which can satisfy (1) and (2), as shown at the top of the next page.

With the above objective, when a task is outsourced to a group, the group's self-situation, the contextual groups' situations, and the communication costs between them should all be considered. Therefore, the priority of a group to be assigned a task is determined not only by the group itself but also by its contextual groups. If a group has higher values for the three factors, it has higher priority to be assigned the task. However, a group may also have higher priority to be assigned the task even it has lower values for the three factors but its contextual groups have higher values for the three factors.

In summary, *the problems of group-oriented crowdsourcing when contexts are aware* can be described as follows.

1) The situations of the contextual groups can influence the performance of a group in performing the task. Thus, the crowdsourcing value of G_i is influenced by G_i' contextual groups in the social network as well as G_i itself. Moreover, the communication costs between G_i and G_i' contextual groups in the social network will also influence the performance. Therefore, how to measure the crowdsourcing value of G_i by considering G_i' contexts is a problem to be investigated (*modeling the groups*).
2) In fact, a group often needs to coordinate with its contextual groups in the social network to execute complex tasks. Then, how can an efficient coordination mechanism between groups to ensure that the task can be completed with higher efficiency be designed (*coordination among groups)*?
3) The execution of a complex task needs the skills of more than one group. Then, how can an efficient method to select assistant groups in the contexts as well as an efficient allocation mechanism to allocate the task to a principal group be designed (*allocation of groups*)?

[1] Workers can self-report their social connections to other workers [7], [8], and workers can construct an underlying social network by following other workers at some platforms, such as Github; therefore, the crowdsourcing system knows the distance between two workers in a social network, i.e., the connection hops between them in the social network. The communication cost between two workers can be calculated as a monotone ascending function of the communication distance between them; the communication cost between two groups is the minimum communication cost of all workers between the two groups.

C. Complexity Analyzes

Theorem 1: The context-aware task allocation problem in group-oriented crowdsourcing satisfying the objective in (1) is NP-hard.

Proof Sketch: We consider a simple and special version of this problem; if the simple version is NP-hard, then our problem is also NP-hard. The simple version is that there is only one group in the social network, whether we can find a set of workers from the group that satisfies

$$\max\left(1/\left(\overbrace{\sum_{\forall w_{ix}\in W_{G_i}(t)}\gamma_{ix}}^{\text{Wages}}\cdot\overbrace{\sum_{\forall w_{ix},w_{iy}\in W_{G_i}(t)}C_{ix,iy}}^{\text{Communication costs}}\right)\right).$$

We prove the decision version of the problem is NP-hard. The decision problem asks whether a set of workers exists from the group that satisfies

$$1/\left(\sum_{\forall w_{ix}\in W_{G_i}(t)}\gamma_{ix}\cdot\sum_{\forall w_{ix},w_{iy}\in W_{G_i}(t)}C_{ix,iy}\right) = q$$

where q is a constant number.

We prove the theorem by a reduction from the set cover problem which is well known to be NP-hard [16]. An instance of the set cover problem consists of a universe U of n elements, a collection of subsets of $U, S = \{S_1, \ldots, S_m\}$, and a cost function b: $S \to Q^+$, $\{b_1, \ldots, b_m\}$. Given a constant t, the decision problem asks whether we can find a sub collection of S that covers all elements of U, and the total cost of them is t.

We transform the instance of the set cover problem to an instance of our problem as follows. Every worker corresponds to a subset x, worker w_{ix}'s skill set is $S_{ix} = S_x$, his wage is $\gamma_{ix} = b_x$, the communication cost between worker w_{ix} and worker w_{iy} is $C_{ix,iy} = b_x + b_y$, $q = 1/((k-1)t^2)$, and k is the number of the selected subsets.

We shall show that the total cost is t if and only if the group satisfies $1/(\sum_{\forall w_{ix}\in W_{G_i}(t)}\gamma_{ix}\cdot\sum_{\forall w_{ix},w_{iy}\in W_{G_i}(t)}C_{ix,iy}) = q$.

Suppose that the total cost is t and the number of subsets selected is k. Thus, the number of selected workers is k, their

$$G_* = \arg\max_{\forall G_i \in G} \left(\alpha \cdot \underbrace{\left(\overbrace{\beta_1 \cdot \left| \left(\bigcup_{\forall w_{ix} \in G_i} S_{ix} \right) \cap S_t \right| / |S_t|}^{\text{Skills}} / \left(\overbrace{\beta_2 \cdot \sum_{\forall w_{ix} \in W_{G_i}(t)} \gamma_{ix}}^{\text{Wages}} \overbrace{\cdot \beta_3 \cdot \sum_{\forall w_{ix}, w_{iy} \in W_{G_i}(t)} C_{ix,iy}}^{\text{Communication costs}} \right) \right)}_{\text{Group}} + (1-\alpha) \cdot \underbrace{\sum_{\forall G_j \in (G - \{G_i\})} \left(\frac{\left| \left(\bigcup_{\forall w_{jy} \in G_j} S_{jy} \right) \cap \left(S_t - \bigcup_{\forall w_{ix} \in G_i} S_{ix} \right) \right|}{\left| S_t - \bigcup_{\forall w_{ix} \in G_i} S_{ix} \right|} \cdot \frac{1}{C_{G_i, G_j}} \right)}_{\text{Group's contexts}} \right) \tag{1}$$

$$\text{s.t.} \sum_{\forall w_{ix} \in W_{Gi}(t)} \gamma_{ix} \le b_t \tag{2}$$

communication cost $(k-1)t$, and

$$1/\left(\sum_{\forall w_{ix} \in W_{G_i}(t)} \gamma_{ix} \cdot \sum_{\forall w_{ix}, w_{iy} \in W_{G_i}(t)} C_{ix,iy} \right) = 1/\left((k-1)t^2 \right).$$

Suppose the group we find satisfies $1/(\sum_{\forall w_{ix} \in W_{G_i}(t)} \gamma_{ix} \cdot \sum_{\forall w_{ix}, w_{iy} \in W_{G_i}(t)} C_{ix,iy}) = q$, and the number of selected workers is k.

Then $q = 1/((k-1)(\sum_{\forall w_{ix} \in W_{G_i}(t)} \gamma_{ix})^2)$, and the total cost of selected subsets is t. The decision version of the simple version of our problem is now proved NP-hard, so we can obtain Theorem 1. □

To solve the NP-hard problem, we present a heuristic approach that can be realized within a limited time complexity. In the approach, we define a function of crowdsourcing value that combines the factors in (1) to measure the priority of a group being selected to participate in a task. Then, the group with higher crowdsourcing value can be preferentially selected to participate in performing the task.

IV. Modeling the Groups in Crowdsourcing

In general, there are two typical types of groups, one is the groups with leaders, and the other is the groups without leaders [17]. The former indicates that the persons in one group are all coordinated by the leader; thus, the communication costs are largely determined by the communication between the leader and all common persons in the group. The latter form indicates that the persons in the group can coordinate with one another; thus, the communication costs are largely determined by the communication among the persons in the group.

The workers can cooperate to complete a complex task. Generally, intragroup cooperation is much easier than intergroup cooperation [19]. Therefore, without loss of generality, this article makes the following assumptions: 1) a worker will definitely accept a cooperation request within the same group and 2) a group will accept a cooperation request outside the group only if the group's threshold can be satisfied, e.g., the group's desire for momentary award or other factors can be satisfied.

A. Groups With Leaders

In a group with a leader, all workers will communicate with the leader, and only the leader communicates with the requester. Let there be a group of workers, $G_i = \{w_{ix}\}$, whose leader is w_{il}, $w_{il} \in G_i$. When the leader wants to select a worker to participate in the execution of the task, the following three aspects of the worker will be considered: 1) the degree to which the worker's skills satisfy the current lacking skills required by the task; 2) the ratio of the worker's wage to the task's budget; and 3) the communication cost between the leader and the worker. The priority of a worker being selected by the leader within the group to participate in the execution of task is determined by the self-crowdsourcing value of the worker within the group.

Definition 2: Self-crowdsourcing value of a worker within the group with a leader. Given a budget b_t for an outsourced task t, the necessary skills to complete t is S_t; let $\overline{S_t}$ be the set of skills for t that are currently lacking. The self-crowdsourcing value of w_{ix} within group G_i (whose leader is w_{il}) for executing t is as follows:

$$v_x^{G_{i,l}}(t) = \frac{|S_{ix} \bigcap \overline{S_t}| / |\overline{S_t}|}{(\gamma_{ix}/b_t) \cdot C_{ix,il}} \tag{3}$$

where S_{ix} and γ_{ix} denote the skills and reservation wage of w_{ix}; $C_{ix,il}$ is the communication cost between w_{ix} and w_{il} in the social network.

We assume that a task, t, is now allocated to a group with a leader, and the leader will select the workers within the group that actually participate in executing the task according to the worker's self-crowdsourcing value defined in Definition 2, shown as Algorithm 1.

Algorithm 1: Selection of Workers Within a Group (With a Leader) for Participating in Executing Task (t, G_i)

```
b = 0; S̄_t = S̄_t − S_il; W_Gi(t) = {w_il}; Temp_G_i = G_i − {w_il};
While (S̄_t ≠ φ) and (b == 0))
    ∀w_ix ∈ Temp_G_i:
        calculate the self-crowdsourcing value of w_ix for
        satisfying current S̄_t according to Definition 2;
    w_i* = arg max_{∀w_ix ∈ Temp_G_i}(v_x^{Gi,l}(t));
    Temp_G_i = Temp_G_i − {w_i*};
    If S_i* ⋂ S̄_t ≠ φ:
        S̄_t = S̄_t − S_i*;
        W_Gi(t) = W_Gi(t) ⋃ {w_i*};
    If Temp_G_i == φ: b = 1;
Output (W_Gi(t));
End.
```

Lemma 1: Let there be two workers in group G_i, w_{ix}, and w_{iy}. If $v_x^{Gi,l}(t) > v_y^{Gi,l}(t)$, it is more probable that worker w_{ix} rather than w_{iy} will be selected by group leader w_{il} to participate in executing task t. Moreover, the union of w_{il} and w_{ix} will satisfy the task allocation objective in crowdsourcing for t more significantly than the union of w_{il} and w_{iy}.

Proof Sketch: 1) Now $v_x^{Gi,l}(t) > v_y^{Gi,l}(t)$, assuming that w_{ix} is not selected by w_{il} but w_{iy} is selected by w_{il} to execute task t, which denotes that the worker with the higher self-crowdsourcing value is not selected but the one with the lower self-crowdsourcing value is selected. In Algorithm 1, the selection of the real participating worker is implemented by selecting one with the maximum self-crowdsourcing value in the crowd of candidates within the group. Therefore, such an assumption cannot occur when Algorithm 1 is used.

2) The objective of task allocation in (1) includes three parts: 1) skills; 2) wages; and 3) communication costs. Now, we can find that the first and second factors in Definition 2 fully correspond to the first and second parts in (1). In fact, the communications costs of executing t by G_i are determined by the following: the communication cost between w_{il} and the requester, and the one between w_{il} and the selected worker within G_i. Now, the former is fixed; thus, the actual communication costs are influenced by the latter. Accordingly, the third factor in Definition 2 factually decides the third part in (1). Therefore, if $v_x^l(t) > v_y^l(t)$, according to Definition 2, the comprehensive value of the three factors $(|S_{ix} \bigcap \overline{S_t}|/|\overline{S_t}|, \gamma_{ix}/b_t, d_{ix,il})$ of w_{ix} is higher than that of w_{iy}. Because the situation of w_{il} is fixed, we can conclude that the union of w_{il} and w_{ix} will satisfy the task allocation objective in crowdsourcing for t more significantly than the union of w_{il} and w_{iy}. □

From Algorithm 1, the final set of workers within group G_i that will actually execute t is $W_{Gi}(t)$. Because each worker in $W_{Gi}(t)$ will communicate with the leader w_{il} for executing task t, the total communication costs between all workers in $W_{Gi}(t)$ and w_{il} is

$$C\big(W_{G_i}(t)\big) = \sum_{\forall w_{ix} \in W_{Gi}(t)} C_{ix,il}. \tag{4}$$

To satisfy the task allocation objective, when the requester decides whether to allocate a task to group G_i, the requester will consider the following three factors of G_i: 1) the satisfaction degree of all workers' skills in the group for the skills required by the task; 2) the ratio of the wages of the workers in $W_{Gi}(t)$ to the task's budget; and 3) the total communication costs between the leader and all workers in $W_{Gi}(t)$. Therefore, the priority of a group being allocated the task can be determined by the self-crowdsourcing value of the group.

Definition 3: Self-crowdsourcing value of a group with a leader. Given a budget b_t for a task t, the necessary skills to complete t are S_t, and the self-crowdsourcing value of group G_i for task t is as follows:

$$v_{G_i}(t) = \frac{\beta_1\left(\left|\left(\bigcup_{\forall w_{ix} \in W_{G_i(t)}} S_{ix}\right) \bigcap S_t\right| / |S_t|\right)}{\beta_2\left(\sum_{\forall w_{ix} \in W_{G_i}(t)} \gamma_{ix}/b_t\right) + \beta_3\left(\sum_{\forall w_{ix} \in W_{G_i}(t)} C_{ix,il}\right)}. \tag{5}$$

The weights, β_1, β_2, and β_3, are applied to reflect the relative importance among the three factors. Definition 3 shows that the self-crowdsourcing value can fully correspond to the task allocation objective. Therefore, the group with a leader that is allocated according to its $v_{Gi}(t)$ can effectively approach the task allocation objective.

B. Groups Without Leaders

In a group without leaders, workers can coordinate with each other autonomously. In the social network, the locality of each worker can be measured by the following.

Definition 4: Centrality of a worker within a group. Let there be a group, G_i; the centrality of a worker, w_{ix}, in G_i, is determined by the reciprocal of the ratio of the worker's total communication costs with others to the average total communication costs of all workers with others in G_i:

$$c_x^{G_i} = 1/\frac{\sum_{\forall w_{iy} \in G_i} C_{ix,iy}}{\left(\sum_{\forall w_{iy} \in G_i} \sum_{\forall w_{iz} \in G_i} C_{iy,iz}\right)/|G_i|} \tag{6}$$

where $C_{ix,iy}$ is the communication cost between workers w_{ix} and w_{iy}, and $|G_i|$ is the number of workers in G_i.

The more central a worker is in the group, the fewer communication costs are needed by such a worker to coordinate with other workers within the group.

In the group without leaders, each worker has a priority of being allocated the complex task, which can determined by the self-crowdsourcing value of a worker in the group and is influenced by the following three factors: 1) the satisfaction degree of the worker's skills for the skills required by the task; 2) the ratio of the worker's wage to the task's budget; and 3) the centrality of the worker in the group.

Definition 5: The self-crowdsourcing value of a worker in a group without leaders. Given a budget b_t for a task t, the necessary skills to complete t are S_t, and the self-crowdsourcing value of w_{ix} in group G_i for task t is as follows:

$$v_x^{G_i}(t) = \frac{\left(|S_{ix} \bigcap S_t|/|S_t|\right) \cdot c_x^{G_i}}{\gamma_{ix}/b_t}. \tag{7}$$

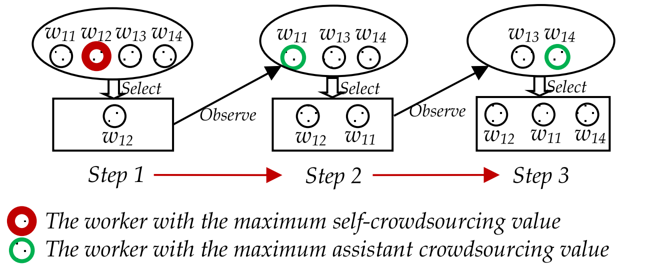


Fig. 1. Iterative selection of participating workers in a group without leaders.

Based on the previous work [9], we present a mechanism for selecting participating workers to execute the outsourced task. After a worker is allocated a complex task, he will then seek another worker within the group to obtain the highest probability that can help him complete the task efficiently. Then, these two workers become the already allocated subgroup. The workers in the already allocated subgroup will then autonomously seek the next assistant worker to obtain the highest probability that can help complete the task efficiently. This iterative process will repeat until all skills required by the task can be satisfied or all workers in the group are observed. To measure the probability of a worker that can help other workers completing the task, we have the following definition.

Definition 6: Assistant crowdsourcing value of a worker perceived by other workers within the group. Let there be a subgroup of workers in group G_i that is allocated for task t, $W_{Gi}(t)$, $W_{Gi}(t) \subset G_i$. Let there be a worker, w_{ix}, $w_{ix} \in G_i \wedge w_{ix} \notin W_{Gi}$. The crowdsourcing value of w_{ix} for assisting $W_{Gi}(t)$ to perform task t is as follows:

$$av_x^{W_{G_i}(t)} = \frac{|S_{ix} \bigcap \overline{S_t}|/|\overline{S_t}|}{(\gamma_{ix}/b_t) \cdot C_{ix,W_{G_i}}} \tag{8}$$

where $C_{ix,WGi}$ is the communication cost between w_{ix} and $W_{Gi}(t)$, and $\overline{S_t}$ is the set of skills for t that are currently lacking by $W_{Gi}(t)$, which can be calculated as the following:

$$C_{ix,W_{G_i}(t)} = \frac{\sum_{\forall w_{iy} \in W_{G_i}(t)} C_{ix,iy}}{|W_{G_i}(t)|}, \overline{S_t} = S_t - \bigcup_{\forall w_{iy} \in W_{G_i}(t)} S_{iy}. \tag{9}$$

Fig. 1 is an example illustrating this process. First, w_{12} is selected because it has the maximum self-crowdsourcing value; then, w_{12} seeks another worker with the maximum assistant crowdsourcing value for $\{w_{12}\}$, which is w_{11}; then, w_{11} and w_{12} will seek another worker with the maximum assistant crowdsourcing value for $\{w_{12}, w_{11}\}$, which is w_{14}. The process can now be finished because the skill requirements of the task can be fully satisfied by $\{w_{12}, w_{11}, w_{14}\}$.

Finally, the process of selecting the workers actually participating in the execution of a task allocated to the group is shown as Algorithm 2.

Definition 7: Self-crowdsourcing value of the group without leaders. Given a budget B_t for a task t, the necessary skills to complete t are S_t, and the self-crowdsourcing value of group

Algorithm 2: Selection of Workers Within a Group (Without Leaders) to Actually Participate in Executing Task (t, G_i)

```
∀w_ix ∈ G_i: calculate the self-crowdsourcing values of w_ix for
  executing task t;
w_i* = arg max_{∀w_ix∈Gi}(v_x^{G_i}(t));
b = 0; S̄_t = S̄_t − S_i*; W_Gi(t) = {w_i*}; Temp_G_i = G_i − W_Gi(t);
While (S̄_t ≠ φ) and (b == 0))
    ∀w_ix ∈ Temp_G_i:
        calculate the assistant crowdsourcing value of w_ix for
        assisting W_Gi(t) to satisfy the current S̄_t according to
        Definition 6;
    w_i* = arg max_{∀w_ix∈Temp_G_i}(av_x^{W_{G_i}(t)});
    Temp_G_i = Temp_G_i − {w_i*};
    If S_i* ⋂ S̄_t ≠ φ:
        S̄_t = S̄_t − S_i*;
        W_Gi(t) = W_Gi(t) ⋃{w_i*};
    If Temp_G_i == φ: b = 1;
Output W_Gi(t);
End.
```

G_i (G_i has no leaders) for task t is as follows:

$$v_{G_i}(t) = \frac{\beta_1\left(\left|\left(\bigcup_{\forall w_{ix} \in W_{G_i}(t)} S_{ix}\right) \bigcap S_t\right| / |S_t|\right)}{\beta_2\left(\sum_{\forall w_{ix} \in W_{G_i}(t)} \gamma_{ix}/b_t\right) + \beta_3\left(\frac{1}{2}\sum_{\forall w_{ix}, w_{iy} \in W_{G_i}(t)} C_{ix,iy}\right)}. \tag{10}$$

The three weights, β_1, β_2, and β_3, are applied to reflect the relative importance among the three factors.

Lemma 2: Let the set of workers participating in the task using Algorithm 2 in group G_i be $W_{Gi}(t)$ and the first selected worker with the maximum self-crowdsourcing value be w_{i*}. We use $P(X)$ to denote the probability that the task allocation objective defined in (1) can be achieved. Another set of workers $W'(t)$ in group G_i is then assumed with the same first selected worker w_{i*}. We thus have the following:

$$\left(W'(t) \subseteq G_i \wedge w_{i*} \in W'(t) \wedge \left(\left(\bigcup_{\forall w_{ix} \in W'(t)} S_{ix}\right) \bigcap S_t \neq \phi\right)\right)$$
$$\Rightarrow P\big(W_{G_i}(t)\big) \geq P\big(W'(t)\big).$$

Proof: The proof is similar to the second part of the proof for Lemma 1, so we skip it for saving paper spaces. ■

According to Lemma 2, the group without leaders that is allocated according to its $v_{Gi}(t)$ can effectively approach the task allocation objective.

C. Coordination and Contexts of Groups

1) Coordination Between Groups: Generally, a worker will definitely accept a cooperation request within the same group but will accept a cooperation request outside the group only if certain preconditions can be satisfied. To incentivize non-cooperative groups to help one another, we are inspired by the coadjutant behaviors in society [21] and assume that the groups in the social network are coadjutant. A group G_i

will have certain obligations to provide assistance for another group, G_j, if G_j has provided assistance to G_i in the past. Therefore, G_j can accept the requests of G_i for assistance even when *G_i's* offered monetary reward is less than *G_j's* reservation wage because G_j also expects to obtain the possible assistance from G_i in the future. To advance cooperation between groups, we present a term of credit between groups as the following definition.

Definition 8: Credit between groups. Let there be two groups, G_i and G_j, and $n_{Gj\leftarrow Gi}$ denotes the historical number of G_i's providing real assistance for G_j's executing tasks. The credit of G_i paid by G_j is in proportion to $n_{Gj\leftarrow Gi}$

$$c_{G_i}(\leftarrow G_j) = f(n_{G_j\leftarrow G_i}) \tag{11}$$

where f is a monotonically increasing function. Clearly, when $c_{Gi}(\leftarrow G_j)$ is higher, it is more compulsory that G_j should provide assistance for G_i's request even when G_i cannot provide sufficient monetary reward to G_j, because in the past, G_i has frequently assisted G_j; thus, G_j is now obligatory to compensate G_i.

When G_i requests assistance from G_j to execute the assigned task, G_i will promise two factors to G_j.

1) *The Possible Monetary Reward for Executing Task t:* Such a factor will be estimated according to the possible contribution of G_j to satisfy the skill requirements of task t. Let S_t be the set of necessary skills required by task t, $\overline{S_t}$ be the set of skills for t that are currently lacking, and S_{jx} be the set of skills owned by worker w_{jx}. Let the amount of budget of t that is distributed to G_i be $b_t(G_i)$; then, the possible monetary reward paid by G_i to G_j is as follows:

$$m_{G_i\rightarrow G_j}(t) = \lambda \cdot \left(b_t(G_i) - \sum_{\forall w_{ix}\in W_{G_i}(t)} \gamma_{ix}\right) \times \frac{\left|\bigcup_{\forall w_{jx}\in G_j} S_{jx} \cap \overline{S_t}\right|}{|S_t|} \tag{12}$$

where $0 \leq \lambda \leq 1$, which denotes the percentage of monetary benefit that G_i is willing to distribute to other assistant groups.

2) *The Credit Paid by G_i to G_j for Executing Task t $c_{Gi\rightarrow Gj}(t)$:* If $c_{Gi\rightarrow Gj}(t)$ is high and G_j hopes to request assistance from G_i in the future, G_j can accept the current request from G_i even when G_j cannot receive a satisfactory monetary reward for this request. If G_j accepts the request from G_i, then: $c_{Gi}(\leftarrow G_j) = c_{Gi}(\leftarrow G_j) - c_{Gi\rightarrow Gj}(t)$, $c_{Gj}(\leftarrow G_i) = c_{Gj}(\leftarrow G_i) + c_{Gi\rightarrow Gj}(t)$. Therefore, in real systems, we can have: $|c_{Gi}(\leftarrow G_j)| = |c_{Gj}(\leftarrow G_i)|$.

Then, after receiving the request from G_i on assisting to execute task t, G_j will decide whether to accept the request according to criteria shown in Section V-B.

2) Contextual Crowdsourcing Values of Groups: In the social networked crowd, the communication costs will significantly influence the performance for completing the outsourced task [13].

Definition 9: Communication cost between two groups. Let there be two groups G_i and G_j such that the communication cost between them, $C_{Gi,Gj}$, is defined as follows: 1) if G_i and G_j both have leaders w_{il} and w_{jl}, then $C_{Gi,Gj} = C_{il,jl}$; 2) if G_i has leader w_{il} but G_j has no leaders, then $C_{Gi,Gj} = \min_{\forall w_{jy}\in Gj} C_{il,jy}$; and 3) if G_i and G_j both have no leaders, then $C_{Gi,Gj} = \min_{\forall w_{ix}\in Gi \wedge \forall w_{jy}\in Gj} C_{ix,jy}$.

The context of a group in the social network primarily means other groups that coordinate with this group through the social network. Clearly, every contextual group within the social network will contribute differently to the contextual crowdsourcing value of a group, given G_i, which is determined by the skills owned by the workers in the contextual group, and the communication cost between the contextual group and G_i.

Definition 10: Contextual crowdsourcing value of a group. The contextual crowdsourcing value of group G_i for task t is defined as follows:

$$Cv_{G_i}(t) = \alpha \cdot v_{G_i}(t) + \frac{(1-\alpha)}{|G|} \cdot \sum_{\forall G_j\in(G-\{G_i\})} \left(\frac{\left|\left(\bigcup_{\forall w_{jy}\in G_j} S_{jy}\right) \cap \left(S_t - \bigcup_{\forall w_{ix}\in G_i} S_{ix}\right)\right|}{\left|S_t - \bigcup_{\forall w_{ix}\in G_i} S_{ix}\right|} \times \frac{\sum_{\forall G_j\in(G-\{G_i\})} C_{G_i,G_j}}{C_{G_i,G_j} \cdot |G|}\right) \tag{13}$$

where G is the set of all groups in the social network; α is a parameter to measure the relative importance between a group itself and the contextual groups, $0 < \alpha < 1$.

From Definition 10, we have the following hypothesis: the group with closer interacting groups whose skills can further make up the skill shortcomings of the group for the outsourced task will be more effective for satisfying the task allocation objective than will other groups.

V. Context-Aware Task Allocation

When a complex task is outsourced, the requester (or the crowdsourcing system) will first allocate a principal group with the maximum contextual crowdsourcing value to take charge of the task. If the principal group lacks any necessary skills required by the task, other contextual groups should be allocated to assist the principal group to execute the task.

A. Allocation of Principal Group

Now, the nonredundant allocation is very popular in real crowdsourcing of complex tasks, e.g., we find that 79.3% of tasks are allocated nonredundantly while we randomly count 6271 tasks at the Upwork website. Therefore, we assume in this article that each complex task is allocated nonredundantly.

As stated above, we will try to select the group with the maximum contextual crowdsourcing value (defined in

Algorithm 3: Allocation of the Principal Group for Task t. /* Let the Set of All Candidate Groups Be G */

$b = 0$; $n = 1$; $G_{temp} = G$;
While $(b == 0)$ and $(n \leq |G|)$ **do:**
 $G_* = \arg\max_{\forall G_i \in G_{temp}} (Cv_{G_i}(t))$;
 $G_{temp} = G_{temp} - G_*$;
 If $\sum_{\forall w_{ix} \in W_{G_i(t)}} \gamma_{ix} \leq b_t$: $b = 1$;
 $n++$;
If $b == 1$: Output (G_*)
 Else: Output (false);
End.

Definition 10). However, each task has a budget constraint; therefore, the total reservation wages of the workers in the group that actually participate in executing task $(W_{Gi}(t))$ should not exceed the budget. Our task allocation criterion is thus now to assign the task to a group that has the highest contextual crowdsourcing value in the set of groups in which each one's real participating workers' reservation wages do not exceed the task's budget, b_t. The process is shown as Algorithm 3.

The time complexity of Algorithm 3 is $O(|G|)$, where $|G|$ denotes the number of the candidate groups.

B. Allocation of Assistant Groups

Here, we will present two approaches for the allocation of assistant groups: 1) semisupervised approach and 2) fully supervised approach.

First, we apply the semisupervised manner into the group-oriented allocation, i.e., only the allocation of the principal group is supervised by the requester (or the crowdsourcing system), and the allocation of assistant groups is conducted by the principal group autonomously. Moreover, we adopt another approach, fully supervised manner, in which the principal group and assistant groups are all allocated by the requester (or the crowdsourcing system).

Each person has a threshold to decide his attitude to cooperate with others [22]. Therefore, here we also set each group, G_i, to have a predefined threshold τ_{Gi}; the group will accept a request for assistance from another group only if the ongoing task's benefit exceeds the threshold.

Generally, when a group, G_j, decides whether to accept a request from group G_i, it primarily considers the following factors: 1) the possible monetary reward for executing task t, $m_{Gi \to Gj}(t)$; 2) the credit paid by G_i to G_j for executing task kt, $c_{Gi \to Gj}(t)$; 3) the cooperation history that G_i has assisted G_j, $c_{Gi}(\leftarrow G_j)$; and 4) the cooperation history that G_j has assisted G_i, $c_{Gj}(\leftarrow G_i)$. Let the threshold of G_j be τ_{Gj}. G_j will accept the request from G_i if the following condition can be satisfied:

$$\eta_1 \cdot m_{G_i \to G_j}(t) + \eta_2 \cdot c_{G_i \to G_j}(t) + \eta_3 \cdot c_{G_i}\big(\leftarrow G_j\big) - \eta_4 \cdot c_{G_j}(\leftarrow G_i) \geq \tau_{G_j} \quad (14)$$

where η_1, η_2, η_3, and η_4 are four parameters to determine the relative importance of the four factors.

1) Semisupervised Approach: The semisupervised approach is often used in the situation in which the principal group has a leader who can search for other groups autonomously. Certainly, in this approach, the principal group also needs to consult the crowdsourcing system on the information of other groups, such as their skills, their thresholds, their distances, and their credits.

Algorithm 4: Semisupervised Approach for Allocation of Assistant Groups (t, G_i). /* Let G_i Be the Principal Group and the Set of All Candidate Groups Be G */

$\overline{S_t} = S_t - \bigcup_{\forall w_{ix} \in G_i} S_{ix}$;
$G_{ass}(t) = \{\}$; $G = G - \{G_i\}$;
G_i enquires the information of other groups from the crowdsourcing system;
While $((G \neq \phi)$ and $(\overline{S_t} \neq \phi))$:
 $G_* = \arg\min_{\forall G_j \in G} (d_{Gi,Gj})$;
 $G = G - \{G_*\}$;
 If $\bigcup_{\forall w_{*x} \in G_*} S_{*x} \bigcap \overline{S_t} \neq \phi$:
  If $\eta_1 \cdot m_{G_i \to G*}(t) + \eta_2 \cdot c_{G_i \to G*}(t) + \eta_3 \cdot c_{G_i}(\leftarrow G_*) - \eta_4 \cdot c_{G*}(\leftarrow G_i) \geq \tau_{G*}$:
   $\overline{S_t} = \overline{S_t} - \bigcup_{\forall w_{*x} \in G_*} S_{*x}$;
   $G_{ass}(t) = G_{ass}(t) \bigcup \{G_*\}$;
Output $(G_{ass}(t))$;
End.

In real society, each person may cooperate with others initially who are neighbors. He then will cooperate with other people from the near to the distant within the social network, which can reduce communication costs and make it more possible to cooperate with acquaintances [9]. Therefore, we now also make the principal group search for other groups from the near to the distant within the social network until all required skills are satisfied or all groups within the social network are observed. The process is shown in Algorithm 4. The time complexity of Algorithm 4 is $O(|G|)$.

Theorem 2: If all groups for performing task t can be obtained by using Algorithm 4, the total communication costs between the principal group and assistant groups can be minimized.

Proof Sketch: Let G_i be the principal group; then, the set of lacking skills of G_i to implement t is $\overline{S_t}$. If Algorithm 4 is used, the set of assistant groups is $G_{\text{ass}}(t)$, and the total communication costs between G_i and the groups in $G_{\text{ass}}(t)$ is C_*. However, if there is another set of groups, G', $G' \neq G_{\text{ass}}(t)$, that can provide all of the skills in $\overline{S_t}$, and the total communication costs between G_i and G' are C'; if $C' < C_*$, it denotes that there are any further groups that provide the required skills in $\overline{S_t}$, but the nearer groups with required skills do not provide the required skills in $\overline{S_t}$. Clearly, such a situation cannot occur in Algorithm 4. Therefore, we have Theorem 2. □

2) Fully Supervised Approach: In the semisupervised approach, the principal group can only observe other groups according to their distances, which might not achieve the optimal result of crowdsourcing values; moreover, the search process might be too complex for the principal group. Therefore, we now provide a fully supervised approach, in which the assistant groups are all allocated by the requester

Algorithm 5: Fully Supervised Approach for Allocation of Assistant Groups (t, G_i). /* Let G_i Be the Principal Group and the Set of All Candidate Groups Be G */

$\overline{S_t} = S_t - \bigcup_{\forall w_{ix} \in G_i} S_{ix}$;
$G_{ass}(t) = \{\}$; $G = G - \{G_i\}$;
While $((G \neq \{\})$ and $(\overline{S_t} \neq \{\}))$
  $G_* = \arg\max_{\forall G_j \in G}(v_{Gj}(G_i - t))$;
  $G = G - \{G_*\}$;
  If $\bigcup_{\forall w_{*x} \in G_*} S_{*x} \bigcap \overline{S_t} \neq \{\}$:
    If $\eta_1 \cdot m_{G_i \to G_*}(t) + \eta_2 \cdot c_{G_i \to G_*}(t) + \eta_3 \cdot c_{G_i}(\leftarrow G_*) - \eta_4 \cdot c_{G_*}(\leftarrow G_i) \geq \tau_{G_*}$:
      $\{\overline{S_t} = \overline{S_t} - \bigcup_{\forall w_{*x} \in G_*} S_{*x}$; $G_{ass}(t) = G_{ass}(t) \bigcup \{G_*\}$; $\}$
Output $(G_{ass}(t))$;
End.

(or the crowdsourcing system). When the requester (or the crowdsourcing system) wants to assign a group to act as an assistant group, he will measure the assistance value of such group according to the following definition.

Definition 11: Assistance value of a group for another group. Let G_i be the principal group for task t. If $\overline{S_t}$ is the set of skills for t that are currently lacking, the assistance value of G_j for G_i on executing t is defined as follows:

$$v_{G_j}(G_i - t) = \frac{\beta_1 \cdot \left(\left| \bigcup_{\forall w_{jx} \in G_j} S_{jx} \bigcap \overline{S_t} \right| / |\overline{S_t}| \right) + \beta_2 \cdot c_{G_i}(\leftarrow G_j)}{\beta_3 \cdot C_{G_i, G_j} + \beta_4 \cdot \tau_{G_j}} \tag{15}$$

where β_1, β_2, β_3, and β_4 are four parameters to decide the relative importance of the four factors.

Finally, the fully supervised approach for allocation of assistant groups is shown as Algorithm 5.

Theorem 3: Let G_i be the principal group and the set of assistant groups for task t using Algorithm 5 be $G_{\text{ass}}(t)$. It is, then, assumed that there is another set of groups in the crowd, G', and the threshold of each group in G' can be satisfied if the group assists G_i for t. Thus, we have

$$\begin{pmatrix} \forall G' \wedge G' \subseteq G \wedge \\ \left(\forall G_j \in G' \Rightarrow \left(\eta_1 \cdot m_{G_i \to G_j}(t) + \eta_2 \cdot c_{G_i \to G_j}(t) \right.\right. \\ \left.\left. +\ \eta_3 \cdot c_{G_i}(\leftarrow G_j) - \eta_4 \cdot c_{G_j}(\leftarrow G_i) \geq \tau_{G_j}\right)\right) \end{pmatrix}$$
$$\Rightarrow \left(\frac{\sum_{\forall G_j \in G_{\text{ass}}(t)} v_{G_j}(G_i - t)}{|G_{\text{ass}}(t)|} \geq \frac{\sum_{\forall G_j \in G'} v_{G_j}(G_i - t)}{|G'|} \right).$$

Proof Sketch: We can use reductio ad absurdum to prove Theorem 3. Assuming there is a set of groups $G', G' \neq G_{\text{ass}}(t) \wedge G' \subseteq G$, that are allocated for assisting G_i to execute the outsourced task t, and

$$\forall G_j \in G' \Rightarrow \big(\eta_1 \cdot m_{Gi \to Gj}(t) + \eta_2 \cdot c_{Gi \to Gj}(t) + \eta_3 \cdot c_{Gi}(\leftarrow G_j) - \eta_4 \cdot c_{Gj}(\leftarrow G_i) \geq \tau_{Gj}\big).$$

If the following assumption is true: $(\sum_{\forall G_j \in G_{\text{ass}}(t)} v_{Gj}(G_i - t))/|G_{\text{ass}}(t)| < \sum_{\forall G_j \in G'} v_{Gj}(G_i - t)/|G'|$ then there exists at least one group with higher assistance value perceived by G_i for t and whose threshold is satisfied by G_i but that cannot be selected by Algorithm 5, and another group with lower assistance value perceived by G_i for t will be allocated. However, in each round for selecting the assistant group, the group with the highest assistance value for G_i for t and whose threshold is satisfied by G_i will definitely be the first to be assigned. Thus, the above assumption cannot occur in reality when Algorithm 5 is used. Therefore, we can have Theorem 3. □

Theorems 2 and 3 show that the advantage of the semisupervised approach is that the communication costs between the principal group and the assistant groups can be optimized, which can significantly improve the performance for completing the task in social networks. In comparison, the advantage of the fully supervised approach is that the maximum crowdsourcing values of the assigned groups can be achieved theoretically because all assistant groups are selected by the requester (or the crowdsourcing system) globally.

VI. Experimental Validation and Analyses

A. Experiment Setting

The experiments are conducted on a real-world dataset extracted from GitHub, https://github.com/, from which data on 3733 groups and 4661 workers (a worker can be affiliated with more than one group) are collected. The dataset including workers (registered users at GitHub) and their skills (users set up tags at GitHub to show their capabilities, e.g., Pathon, Java, etc.), groups and their members (GitHub provides team management functions so that users can work together), and social network of workers containing following context (the people the worker follows) and follower context (the people who follow the worker). The tasks involved in this article are complex tasks that cannot be completed by a single person and require multiple workers to work together to meet the multiple skill needs of the tasks, which are collected from hybrid categories of tasks at Github, such as Web developing, rescue mission, and ground mapping.

We extract the top 40 skills that are most frequently possessed by workers; then, we select the workers who can provide at least one of these top 40 skills and select the groups that each contains at least two of these workers; finally, we obtain 602 groups and 1494 workers. The social network on these workers is constructed according to their following relationship, in which the vertices denote the workers and the edge between two vertices denotes that there is following or followed relationship between the two corresponding workers. In fact, the "following" or "being followed by" relationships are not cooperation relationships, which are only used to denote that two workers are connected in the social network.

Our two group-oriented task allocation approaches, semisupervised approach (*Group_semi supervised*) and fully supervised approach (*Group_fully supervised*), are compared with the following benchmark approaches.

1) Team formation algorithm based on Minimal Cost Contribution (*Team formation*) [13]. Several initial teams

are formed; then, the new members are selected to add to the initial teams by considering their communication costs with all of the current members of the team in addition to their personal costs. Finally, a best team is selected. The details are described in [13].

2) Individual-oriented algorithm (*Individual*) [24]. The algorithm initially assigns the task to a principal worker who has the best contextual factor based on the skills the worker has, the communication cost with other workers, and the reservation wage of the worker. The principal worker then searches for other assistant workers based on the skills the worker has and the reservation wage of the worker from near to far within the social network until all of the skills required by the task are satisfied.
3) Group-oriented algorithm based on skill local search (*Group_skill local search*). The algorithm initially assigns the task to a principal group that can best meet the skill requirement of the task. Then, the principal group will search from near to far within the social network for cooperation until all of the skills required by the task are satisfied.

We implement these algorithms by Java and test them on a PC with an Intel Dual Core 3.40 GHz CPU. All of the results are recorded by averaging over 100 instances. We run the experiment in five situations: the occupancy rate of groups with leaders is 0% (that is, all of the groups are without leaders), 30%, 50%, 80%, and 100% (that is, all of the groups have leaders). There are two series of parameters to determine the relative importance between different factors in the definitions of crowdsourcing values: we use α and $1-\alpha$ to denote the relative importance between a group and its contexts, respectively; moreover, we use β_1, β_2, and β_3 to denote the relative importance among the three factors for determining crowdsourcing value. In the experiments in Sections VI-B and VI-C and 6.4 in the Appendix, these parameters are: $\alpha = 0.75$; $\beta_1 = 1$, $\beta_2 = 0.5$, and $\beta_3 = 0.5$.

B. Experiments on the Performance

We test the following six types of performances in experiments: 1) synergy performance among the assigned workers; 2) consistency performance of the assigned workers; 3) conflict performance of the assigned workers; 4) adaptability for varying tasks; 5) average pairwise communication costs among the assigned workers; and 6) reservation wages of the assigned workers. Metrics 1, 2, and 3 can test the task quality performance of our approach; metric 4 can test the effects of our approach on addressing the limitation of team formation; and metrics 5 and 6 can measure the performance of our approach on reducing costs.

The reasons that we use the above metrics to test the performance are shown as follows.

1) It is generally accepted that the synergy performance of a group can significantly determine the cooperation performance of such group to complete tasks [25], [37].
2) The consistency performance of a group can reflect the cooperation structure of the group that can significantly influence the cooperation performance and reliability of such group to complete tasks [30], [38].
3) Conflict performance can also influence the group's performance to complete tasks in previous studies, which often use the harmonic mean of average path length (HMAPL) of the group in social networks to indicate the group's potential of conflict [28], [29].
4) The adaptability for varying tasks can test the effect of our approach on addressing the limitation of team formation that can be only used for special tasks.
5) The communication costs and reservation wages are two main costs in completing the complex tasks.

1) Synergy Performance Among Assigned Workers: In this section, we test the synergy performance between the workers selected by the group-oriented task allocation and that of selected by other benchmark approaches.

We use the weighted synergy graph to model the task-based relationship among workers, in which the vertices represent workers and the edges represent the task-based relationship between workers [37]. If two workers are originally in a group or are assigned to cooperate for the first time to complete a task, the corresponding two vertices are connected by an edge with an initial weight of 100; then, the weight will be reduced by one for each cooperation. The distance between two vertices is the weight between them. The shorter the distance is, the more cooperative they are. Suppose that there are two workers x and y, let $d(x, y)$ be the shortest distance between these two workers in the graph. Let $s(x, y)$ denote the pairwise synergy value of x and y

$$s(x, y) = 5 - \ln(d(x, y)). \tag{16}$$

Let $\mathcal{S}(A)$ denote the synergy value of a set of workers A, which is the average of the pairwise synergy values of all worker pairs in A

$$S(A) = \frac{1}{\binom{A}{2}} \sum_{\{x,y\}\in A} s(x, y). \tag{17}$$

The larger the synergy value of a set is, the more harmonious the cooperation among workers in the set is [37].

We now compare our two group-oriented approaches with the other three benchmark approaches on the synergy performance of the assigned workers. The results on the average synergy values of 100 tasks in five situations (with varying occupancy rates of groups with leaders) are shown in Fig. 2(a). The results show that the three group-oriented approaches can achieve higher synergy values than can the other two approaches; thus, the group-oriented crowdsourcing paradigm can achieve better cooperation performance among workers. In addition, the workers selected by the *Group_skill local search* approach achieve the highest synergy value; the potential reason is that the approach only considers the skills of the workers but ignores the wages of the workers and the communication costs between them; therefore, the workers selected by the approach are within fewer groups and then can achieve higher synergy value.

2) Consistency Performance of the Assigned Workers: It is well known that decision-making is crucial in teamwork, thus whether members can reach agreements is an important aspect of judging the quality of the group [18]. In social networks, closer relationship helps groups achieve consensus

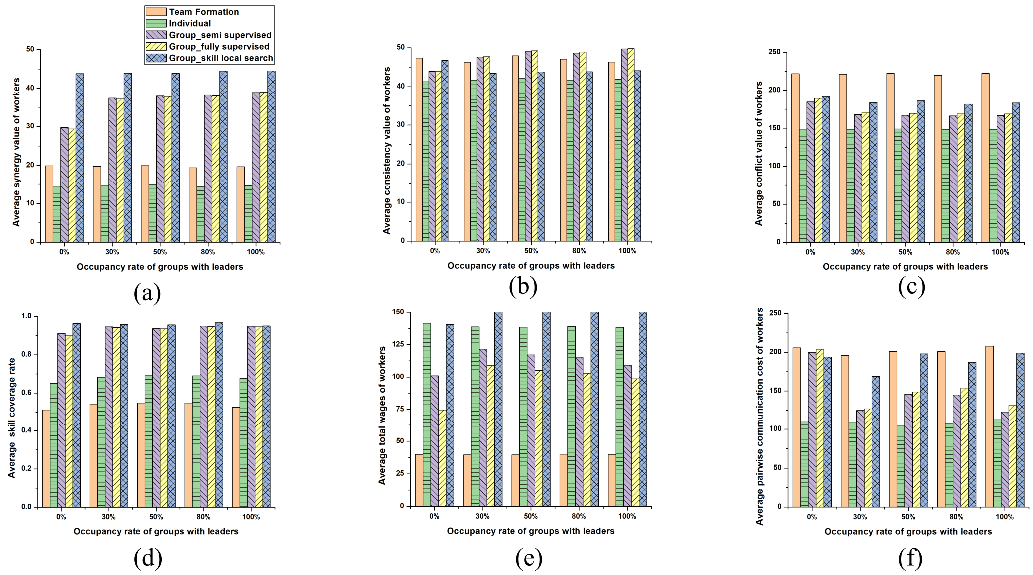


Fig. 2. Experiments on the performance. (a) Synergy performance. (b) Consistency performance. (c) Conflict performance. (d) Adaptability for varying tasks. (e) Effectiveness-total reservation wages. (f) Effectiveness-average pairwise communication cost.

more easily [38]. Clustering Coefficient is often used to measure the degree of the nodes tend to cluster in graph theory; generally, the higher the Clustering Coefficient is, the closer the nodes are in the graph [28]. Therefore, we use the average Clustering Coefficient $C(A)$ to measure the consistency of the selected workers set A [38].

In graph $G(V, E)$, $v_i \in V$ indicates a worker in set A. Let N_i be the set of neighbors of node v_i in the social network contexts, and $e_{j,k} \in E$ indicates that there is a direct connection between nodes v_j and v_k. Thus, the consistency value of v_i can be defined as

$$C_i = \frac{2 \cdot \left|\left\{e_{j,k} : v_j, v_k \in N_i, e_{j,k} \in E\right\}\right|}{|N_i| \cdot (|N_i| - 1)}. \tag{18}$$

The average consistency value of set A is defined as

$$C(A)=\frac{1}{|V|}\sum_{i=1}^{|V|} C_i. \tag{19}$$

Intuitively, higher average Clustering Coefficient means that the group is closer and members can reach agreements more easily [30].

The results on the average consistency value of 100 tasks in the five situations are shown in Fig. 2(b). The results show that when the occupancy rate of groups with leaders is 0%, our two group-oriented approaches have relatively poor performance than the *Team formation* approach and *Group_skill local search* approach; while the rate increases to 30% and more, our two approaches achieve higher consistency value than other approaches. The reason is that the group with a leader will select workers having social relationship with the leader more probably; therefore, the group members will be closer.

3) Conflict Performance of the Assigned Workers: The social relationship of the group members is one of the causes of conflict. The distance between two workers in a social network can reflect their familiarity, and two unfamiliar workers may result in conflicts in teamwork [29]. The HMAPL of the group in social networks could indicate the group's potential of conflict [28], [29].

In graph $G(V, E)$, $v_i \in V$ indicates a worker in set A. Let $d(x, y)$ be the shortest distance between workers x and y in the graph. The HMAPL of A is defined as

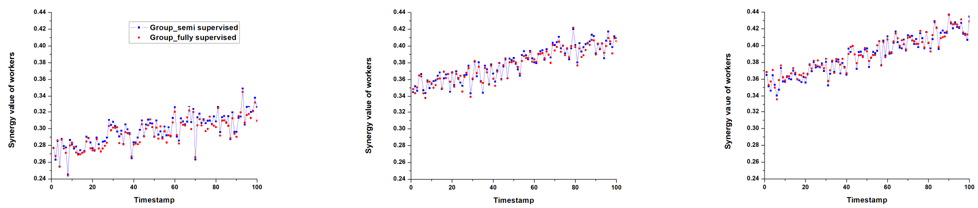
$$APL(A) = \frac{|V| \cdot (|V| - 1)}{\sum_{x \neq y} \frac{1}{d(x,y)}}. \tag{20}$$

Intuitively, smaller HMAPL of a group means that the members are more familiar and the group does not have conflicts easily.

The results show that our two group-oriented approaches have relatively better performance than *Team formation* approach and *Group_skill local search* approach, but not to *Individual* approach. The potential reason is that the principal worker searches for other assistant workers from the near to the distant within the social network in *Individual* approach; thus, the assigned workers are more familiar. In particularly, when the occupancy rate of groups with leaders increases to 30% and more, the gap between our two approaches and the *Individual* approach gradually narrows, as the group leader could select workers who have social relationship with the leader more probably.

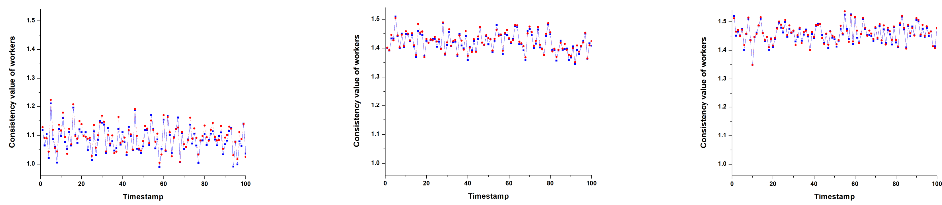
4) Adaptability of the Assigned Workers: In this section, we test the adaptability of the workers selected by our approaches for varying tasks.

We select 100 tasks randomly and then use the five approaches to select five sets of workers for the first task. Then, we use the remaining tasks to examine the adaptability of the five sets. First, for each of the remaining tasks, we apply the individual-oriented method to the five worker sets to select five subsets of workers, respectively. We then count the average skill coverage rate of the skills of the five selected worker



(a) (b) (c)

Fig. 3. Dynamics experiments on the synergy of assigned workers. (a) Occupancy rate of groups with leaders is 0%. (b) Occupancy rate of groups with leaders is 50%. (c) Occupancy rate of groups with leaders is 100%.



(a) (b) (c)

Fig. 4. Dynamics experiments on the consistency of assigned workers. (a) Occupancy rate of groups with leaders is 0%. (b) Occupancy rate of groups with leaders is 50%. (c) Occupancy rate of groups with leaders is 100%.

subsets. The higher the average skill coverage of the selected workers is, the better the approach can adapt to varying tasks.

The experimental results are shown in Fig. 2(d). The results show that the workers selected by the group-oriented approaches can obtain a higher skill coverage rate than can the workers selected by the other two approaches; thus, groups can be assigned varying tasks, and the group-oriented approaches are more adaptive than the *Team formation* approach and *Individual* approach in crowdsourcing markets.

5) Effectiveness on Reducing Costs: To evaluate the effectiveness of our approaches on reducing costs, we test the following performance metrics: the total reservation wages of all selected workers, and the average pairwise communication costs among all selected workers.

Fig. 2(e) shows the average of the selected workers' total wages for 100 tasks. The results denote that both the average values of the total wages of the workers selected by the *Group_semi supervised* approach and *Group_fully supervised* approach are higher than that of the workers selected by *Team formation* approach, but lower than that of the workers selected by *Individual* approach and *Group_skill local search* approach. When the occupancy rate of groups with leaders is 0%, the average of total reservation wages of the workers selected by our two approaches is relatively lower than other situations in which the rate increases to 30% and more.

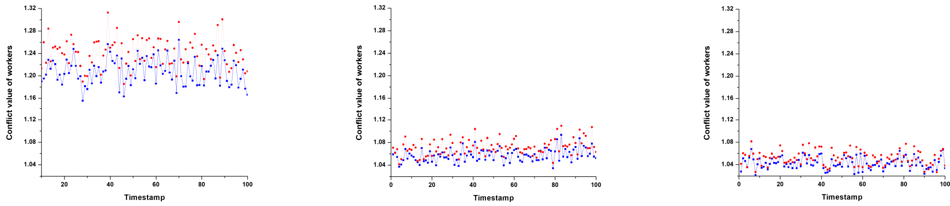
The results on the average pairwise communication cost of the assigned workers for 100 tasks are shown in Fig. 2(f). The results show that the performance of the *Group_semi supervised* approach and *Group_fully supervised* approach are superior to that of *Team formation* approach and are inferior to *Individual* approach. Moreover, when the occupancy rate of groups with leaders is 0%, our two group-oriented approaches have relatively worse performance than *Group_skill local search* approach; while in other four situations in which the occupancy rate of groups with leaders is 30% and more, our two group-oriented approaches have an obvious advantage over *Group_skill local search* approach and narrow the gap with the *Individual* approach. The potential reason is that groups with leaders will select workers who have social relationship with leaders more probably; therefore, the average pairwise communication cost becomes lower.

C. Experiments on the Dynamics

1) Dynamics of the Synergy Performance: From the definition on synergy performance in Section VI-B1, the synergy values are dynamically changed with the cooperation number between workers. Now, we set the number of tasks as the timestamp, which is set from 1 to 100; then, we test the dynamics of synergy performance of our two approaches. The results on three types of situations with different occupancy rates of groups with leaders are shown in Fig. 3(a)–(c).

We can see that the synergy value of the workers selected by the *Group_semi supervised* approach shows a slight advantage compared to the *Group_fully supervised* approach. The potential reason is that when using the *Group_semi supervised* approach, the principal group searches other groups for help from the near to the distant within the social network; therefore, it is more possible that adjacent groups had cooperation in the past and that the workers selected by *Group_semi supervised* approach can achieve a higher synergy value.

2) Dynamics of the Consistency Performance: According to Section VI-B2, the consistency performance is related to the social relationship between workers. If two workers in the same group do not have social relationships originally, they will be connected after completing task. Now, we set the number of tasks as the timestamp, which is set from 1 to 100; then, we test the dynamics of the consistency values of the assigned worker groups of our two approaches, shown in Fig. 4(a)–(c). We can see that the dynamics of the consistency values of the workers selected by our approaches are very close.



(a) (b) (c)

Fig. 5. Dynamics experiments on the conflict of assigned workers. (a) Occupancy rate of groups with leaders is 0%. (b) Occupancy rate of groups with leaders is 50%. (c) Occupancy rate of groups with leaders is 100%.

3) Dynamics of the Conflict Performance: Now, we test the dynamics of the APL values of the selected workers, shown in Fig. 5(a)–(c). We can see that the HMAPL of group of workers selected by the *Group_semi supervised* approach is significantly smaller than the *Group_fully supervised* approach, which indicates that the workers selected by the *Group_semi supervised* approach are less likely to have conflict. The potential reason is that when using *group_semi supervised* approach, the principal group searches other groups for help from the near to the distant within the social network, thus the selected workers are more familiar.

Moreover, we compare the *Group_semi supervised* approach with the *Group_fully supervised* approach on other three performance metrics: 1) adaptability performance; 2) the selected workers' total wages; and 3) the selected workers' pairwise communication costs; we make experiments on the uncertainty of our two approaches resulted from the following parameters: 1) relative importance factor between a group and its contextual groups; 2) number of skills of task; and 3) number of tasks. The experimental results are shown in the Appendix.

VII. Conclusion

To address the common issue that workers are often naturally organized into groups through social networks in real crowdsourcing systems, this article presents a novel group-oriented crowdsourcing paradigm which is different from previous generally used individual-oriented and team formation approaches. With the new approach, the tasks are allocated to the worker groups. This article initially proves the problem is NP-hard; then, this article presents a heuristic approach that can be realized within a limited time complexity. In the heuristic approach, a function of crowdsourcing value is defined to measure the priority of a group being selected to participate in a task. This article theoretically proves that the heuristic approach can ensure that the optimization objective is approached.

Finally, this article conducts extensive experiments on a real-world crowdsourcing dataset. The experimental results show that our presented group-oriented approaches can nearly always achieve better synergy performance, consistency performance, conflict performance, adaptability, and effectiveness on reducing costs, as compared with previous benchmark individual-oriented and team formation approaches.

Regarding the future work, we will address the following issues.

1) This article assumes that the groups are fixed during task allocation and execution. In reality, the social networks and groups may sometimes be dynamic [35], in which workers may depart or join the groups dynamically. Therefore, the adaptive mechanism and self-organization mechanism [31] will be introduced for addressing dynamic groups in the future.
2) This article assumes that the groups are reliable during task allocation and execution. However, due to the uncertainty and openness of social networks [36], the groups may sometimes behave unreliably and the credits of groups may be not trusted. Thus the trust and reputation mechanisms [35], [36] will be introduced for addressing unreliable groups in the future.
3) This article only models two typical groups, the ones with leaders and the ones without leaders. In fact, there are varying complex and dynamic groups [39]. Therefore, in the future we will introduce more group models into the modeling of worker groups and group crowdsourcing behaviors.

Acknowledgment

The authors would like to thank Zhanpeng He and Peng Shi for their work in the experiments.

References

[1] H. Jiang and S. Matsubara, "Efficient task decomposition in crowdsourcing," in *Proc. 17th Int. Conf. Principles Pract. Multi Agent Syst. (PRIMA)*, Dec. 2014, pp. 65–73.
[2] L. Tran-Thanh, T. D. Huynh, A. Rosenfeld, S. Ramchurn, and N. R. Jennings, "BudgetFix: Budget limited crowdsourcing for interdependent task allocation with quality guarantees," in *Proc. 13th Int. Conf. Auton. Agents Multiagent Syst. (AAMAS)*, May 2014, pp. 477–484.
[3] C. Wang, L. Cao, and C.-H. Chi, "Formalization and verification of group behavior interactions," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 45, no. 8, pp. 1109–1124, Aug. 2015.
[4] Q. Liu, T. Luo, R. Tang, and S. Bressan, "An efficient and truthful pricing mechanism for team formation in crowdsourcing markets," in *Proc. IEEE Int. Conf. Commun. (ICC)*, London, U.K., Jun. 2015, pp. 567–572.
[5] I. Lykourentzou, R. E. Kraut, S. Wang, and S. P. Dow, "Team dating: A self-organized team formation strategy for collaborative crowdsourcing," in *Proc. ACM CHI Conf. Human Factors Comput. Syst. (CHI)*, San Jose, CA, USA, May 2016, pp. 1243–1249.
[6] P. Liu and Z. Li, "Task complexity: A review and conceptualization framework," *Int. J. Ind. Ergon.*, vol. 42, no. 6, pp. 553–568, 2012.
[7] M. Yin and M. L. Gray, "The communication network within the crowd," in *Proc. 25th Int. World Wide Web Conf. (WWW)*, Montreal, QC, Canada, Apr. 2016, pp. 1293–1303.
[8] M. L. Gray, S. Suri, S. S. Ali, and D. Kulkarni, "The crowd is a collaborative network," in *Proc. 19th ACM Conf. Comput. Supported Cooperative Work Soc. Comput. (CSCW)*, San Francisco, CA, USA, Feb./Mar. 2016, pp. 134–147.

[9] Y. Jiang, Y. Zhou, and Y. Li, "Reliable task allocation with load balancing in multiplex networks," *ACM Trans. Auton. Adapt. Syst.*, vol. 10, no. 1, 2015, Art. no. 3.

[10] C. Gao and J. Liu, "Network-based modeling for characterizing human collective behaviors during extreme events," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 47, no. 1, pp. 171–183, Jan. 2017.

[11] M. C. Yuen, I. King, and K. S. Leung, "A survey of crowdsourcing systems," in *Proc. IEEE 3rd Int. Conf. Soc. Comput. (SocialCom)*, Boston, MA, USA, Oct. 2011, pp. 766–773.

[12] L. Tran-Thanh, S. Stein, A. Rogers, and N. R. Jennings, "Efficient crowdsourcing of unknown experts using multi-armed bandits," *Artif. Intell.*, vol. 214, pp. 89–111, Sep. 2014.

[13] M. Kargar, A. An, and M. Zihayat, "Efficient bi-objective team formation in social networks," in *Proc. Eur. Conf. Mach. Learn. Principles Pract. Knowl. Disc. Databases (ECML PKDD)*, Bristol, U.K., Sep. 2012, pp. 483–498.

[14] M. Rokicki, S. Zerr, and S. Siersdorfer, "Groupsourcing: Team competition designs for crowdsourcing," in *Proc. 24th Int. Conf. World Wide Web (WWW)*, Florence, Italy, May 2015, pp. 906–915.

[15] J. Chamberlain, "Groupsourcing: Distributed problem solving using social networks," in *Proc. 2nd AAAI Conf. Human Comput. Crowdsourcing (HCOMP)*, Pittsburgh, PA, USA, Nov. 2014, pp. 22–29.

[16] V. V. Vazirani, *Approximation Algorithms*. Heidelberg, Germany: Springer, 2003.

[17] H. Oh, M. H. Chung, and G. Labianca, "Group social capital and group effectiveness: The role of informal socializing ties," *Acad. Manag. J.*, vol. 47, no. 6, pp. 860–875, 2004.

[18] N. L. Kerr and R. S. Tindale, "Group performance and decision making," *Annu. Rev. Psychol.*, vol. 55, pp. 623–655, Feb. 2003.

[19] M. Goldman, J. W. Stockbauer, and T. G. McAuliffe, "Intergroup and intragroup competition and cooperation," *J. Exp. Soc. Psychol.*, vol. 13, no. 1, pp. 81–88, 1977.

[20] E. Franchi, A. Poggi, and M. Tomaiuolo, "Social media for online collaboration in firms and organizations," *Int. J. Inf. Syst. Model. Design*, vol. 7, no. 1, pp. 18–31, 2016.

[21] J. N. Constantino and R. D. Todd, "Genetic structure of reciprocal social behavior," *Amer. J. Psychiatry*, vol. 157, no. 12, pp. 2043–2045, 2000.

[22] J. Wang, F. Fu, T. Wu, and L. Wang, "Emergence of social cooperation in threshold public goods games with collective risk," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 80, Jul. 2009, Art. no. 016101.

[23] S. He, D. H. Shin, J. Zhang, and J. Chen, "Toward optimal allocation of location dependent tasks in crowdsensing," in *Proc. 33rd Annu. IEEE Int. Conf. Comput. Commun. (INFOCOM)*, Toronto, ON, Canada, Apr./May 2014, pp. 745–753.

[24] J. Jiang, B. An, Y. Jiang, and D. Lin, "Context-aware reliable crowdsourcing in social networks," *IEEE Trans. Syst., Man, Cybern., Syst.*, to be published. doi: 10.1109/TSMC.2017.2777447.

[25] S. Liemhetcharat and M. Veloso, "Modeling and learning synergy for team formation with heterogeneous agents," in *Proc. 11th Int. Conf. Auton. Agents Multiagent Syst. (AAMAS)*, Valencia, Spain, Jun. 2012, pp. 365–374.

[26] A. Kittur, E. H. Chi, and B. Suh, "Crowdsourcing user studies with mechanical turk," in *Proc. SIGCHI Conf. Human Factors Comput. Syst. (CHI)*, Florence, Italy, Apr. 2008, pp. 453–456.

[27] C. J. Ho and J. W. Vaughan, "Online task assignment in crowdsourcing markets," in *Proc. 26th AAAI Conf. Artif. Intell. (AAAI)*, Toronto, ON, Canada, Jul. 2012, pp. 45–51.

[28] M. E. J. Newman, "The structure and function of complex networks," *SIAM Rev.*, vol. 45, no. 2, pp. 167–256, 2003.

[29] J. Liu, H. Qiu, N. Zhong, and C. Gao, "A dynamic trust network for autonomy-oriented partner finding," *J. Intell. Inf. Syst.*, vol. 37, no. 1, pp. 89–118, 2011.

[30] A. O. Sousa, "Consensus formation on a triad scale-free network," *Physica A Stat. Mech. Appl.*, vol. 348, pp. 701–710, Mar. 2005.

[31] D. Ye, M. Zhang, and A. V. Vasilakos, "A survey of self-organization mechanisms in multiagent systems," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 47, no. 3, pp. 441–461, Mar. 2017.

[32] J. Ren, Y. Zhang, K. Zhang, and X. Shen, "Exploiting mobile crowdsourcing for pervasive cloud services: Challenges and solutions," *IEEE Commun. Mag.*, vol. 53, no. 3, pp. 98–105, Mar. 2015.

[33] T. Lappas, K. Liu, and E. Terzi, "Finding a team of experts in social networks," in *Proc. 15th ACM SIGKDD Int. Conf. Knowl. Disc. Data Min. (KDD)*, Paris, France, Jun./Jul. 2009, pp. 467–476.

[34] C. T. Li and M. K. Shan, "Composing activity groups in social networks," in *Proc. 21st ACM Int. Conf. Inf. Knowl. Manag. (CIKM)*, Oct./Nov. 2012, pp. 2375–2378.

[35] L. Gao, Y. Ding, and H. Ying, "An adaptive social network-inspired approach to resource discovery for the complex grid systems," *Int. J. Gen. Syst.*, vol. 35, no. 3, pp. 347–360, 2006.

[36] F. Liu, L. Wang, L. Gao, H. Li, H. Zhao, and S. K. Men, "A Web service trust evaluation model based on small-world networks," *Knowl. Based Syst.*, vol. 57, pp. 161–167, Feb. 2014.

[37] S. Liemhetcharat and M. Veloso, "Weighted synergy graphs for effective team formation with heterogeneous ad hoc agents," *Artif. Intell.*, vol. 208, pp. 41–65, Mar. 2014.

[38] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.

[39] V. N. Sayankar, "Effect of group behavior and group dynamics in work culture of organization," *Int. J. Marketing Financ. Services Manag. Res.*, vol. 3, no. 10, pp. 69–75, 2015.

[40] B. Guo *et al.*, "Mobile crowd sensing and computing: The review of an emerging human-powered sensing paradigm," *ACM Comput. Surveys*, vol. 48, no. 1, 2015, Art. no. 7.

[41] B. Guo, Y. Liu, L. Wang, V. O. K. Li, J. C. K. Lam, and Z. Yu, "Task allocation in spatial crowdsourcing: Current state and future directions," *IEEE Internet Things J.*, vol. 5, no. 3, pp. 1749–1764, Jun. 2018.

Jiuchuan Jiang received the Ph.D. degree in computer science from Nanyang Technological University, Singapore, in 2019.

He is currently with the College of Information Engineering, Nanjing University of Finance and Economics, Nanjing, China. He has published several scientific articles in refereed journals and conference proceedings, such as the IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS, IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS: SYSTEMS, the *ACM Transactions on Autonomous and Adaptive Systems*, and the *International Conference on Autonomous Agents and Multiagent Systems*. His current research interests include crowdsourcing, multiagent systems, and social networks.

Bo An received the Ph.D. degree in computer science from the University of Massachusetts, Amherst, MA, USA, in 2010.

He is an Associate Professor with the School of Computer Science and Engineering, Nanyang Technological University, Singapore. He has published over 90 referred papers at AAMAS, IJCAI, AAAI, ICAPS, KDD, JAAMAS, AIJ, and ACM/IEEE TRANSACTIONS. His current research interests include artificial intelligence, multiagent systems, game theory, and optimization. His research results have been successfully applied to many domains, including infrastructure security and e-commerce.

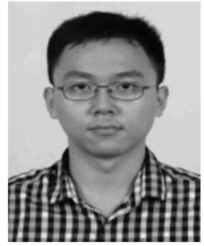
Dr. An was a recipient of the 2010 IFAAMAS Victor Lesser Distinguished Dissertation Award, an Operational Excellence Award from the Commander, First Coast Guard District of the United States, the 2012 INFORMS Daniel H. Wagner Prize for Excellence in Operations Research Practice, and the 2018 Nanyang Research Award (Young Investigator). His publications won the Best Innovative Application Paper Award at AAMAS'12 and the Innovative Application Award at IAAI'16. He was invited to give Early Career Spotlight talk at IJCAI'17. He led the team HogRider which won the 2017 Microsoft Collaborative AI Challenge. He was named to IEEE Intelligent Systems' "AI's 10 to Watch" list for 2018. He was invited to be an Advisory Committee Member of IJCAI'18. He is an Editorial Board Member of JAIR and an Associate Editor of *Autonomous Agents and Multi-Agent Systems* and IEEE INTELLIGENT SYSTEMS. He was elected to the board of directors of IFAAMAS.

Yichuan Jiang (SM'13) received the Ph.D. degree in computer science from Fudan University, Shanghai, China, in 2005.

He is currently a Full Professor with the School of Computer Science and Engineering, Southeast University, Nanjing, China. His current research interests include multiagent systems, crowdsourcing, and social networks. He has published over 100 scientific articles in refereed journals and conference proceedings, such as the IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS, IEEE TRANSACTIONS ON MOBILE COMPUTING, IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS: SYSTEMS, IEEE TRANSACTIONS ON CYBERNETICS, the *ACM Transactions on Autonomous and Adaptive Systems*, the *Journal of Autonomous Agents and Multi-Agent Systems*, IJCAI, AAMAS, and AAAI.

Prof. Jiang was a recipient of the best paper award from PRIMA06 and best student paper awards twice from ICTAI13 and ICTAI14.

Chenyan Zhang received the bachelor's degree in computer science from Southeast University, Nanjing, China, in 2017. She is currently pursuing the M.E. degree in computer science with the School of Computer Science and Engineering, Southeast University, and also with the Co-Innovation Center of Shandong Colleges and Universities for Future Intelligent Computing, Shandong Technology and Business University, Yantai, China.



Zhan Bu received the Ph.D. degree in computer science from the Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 2014.

He is currently an Associate Professor with the College of Information Engineering, Nanjing University of Finance and Economics, Nanjing. His current research interests include social networks and multiagent systems.

Jie Cao received the Ph.D. degree in manufacturing automation from Southeast University, Nanjing, China, in 2002.

He is currently a Professor and the Dean of the College of Information Engineering, Nanjing University of Finance and Economics, Nanjing. His current research interests include business intelligence and social networks.