IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 13 November 2022, accepted 22 December 2022, date of publication 26 December 2022,
date of current version 30 December 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3232401

RESEARCH ARTICLE

Privacy-Preserving Trust-Aware Group-Based Framework in Mobile Crowdsensing

BAYAN HASHR SAEED ALAMRI, MUHAMMAD MOSTAFA MONOWAR, AND SUHAIR ALSHEHRI
Department of Information Technology, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah 21589, Saudi Arabia
Corresponding author: Bayan Hashr Saeed Alamri (bsaeedalamri@stu.kau.edu.sa)

ABSTRACT In practical mobile crowdsensing (MCS) systems, many cooperative sensing tasks require a group of reliable participants to perform collaboratively. In this article, we address the problem of group formulation in MCS, which aims to recruit highly trusted participants and form a high-reputation group. We propose a novel Privacy-preserving Trust-Aware Group Formation (PTAGF) framework that ensures trust and privacy between the group members. This framework consists of three mechanisms; the member trust assessment mechanism, the group forming mechanism, and the two-layer privacy-preserving mechanism. Furthermore, we prove that the group forming problem is NP-hard, and thus propose a heuristic-based Trust-Aware Group Formulation (TAGF) algorithm. A theoretical analysis is provided, which demonstrates that the proposed framework achieves privacy and security. Finally, we experimentally evaluate the performance of PTAGF on a real-world dataset against two state-of-the-art approaches. The results demonstrate that PTAGF outperforms these approaches in terms of trustworthiness in group selection. Moreover, it achieves reasonable task coverage and running time with different communities size, group sizes, and task scales.

INDEX TERMS Mobile crowdsensing, cooperative task, privacy-preserving, trust-aware, group formation.

I. INTRODUCTION

Mobil Mobile Crowdsensing (MCS) is a new emerging paradigm, taking advantage of sensor-embedded wearable devices and smartphones to collect sensor readings efficiently. MCS applications and services are rapidly increasing, getting more research attention, and moving beyond a single individual participant to community groups and are influenced by group behavior and social networks [1]. In the real world, the number of cooperative sensing tasks increases which requires a group of participants to perform collaboratively for certain applications, which triggers the idea of group recruitment [2]. The quality of sensor readings for these tasks lies in recruiting trusted and reliable participants to accomplish these tasks cooperatively. Moreover, the group member would prefer to cooperate with trustworthy members, especially when the task requires some special skills to be performed. Hence, how to verify the trust and the ability of participants to form effective groups is a key research problem. Most existing group-forming systems achieve good communication and social costs [3], [4], [5], [6]. They either consider the social relationship between group members, the distance between members and the task, or the similarity between the members' interests and the tasks. However, they fail to take the group member's trust and the ability to perform the assigned task into account during the recruiting process, which may form low-reputation groups that are socially connected but not able to perform the sensing tasks. In addition, the privacy of participants during the forming group process and the cooperation between group members should be preserved from platform entities and malicious participants. Grouping-based solution methods protect participants' information during participant recruitment and minimize the information loss and the overhead on the participants' side [7]. Thus recruiting trusted participants to perform cooperative sensing tasks can improve not only the participation willingness of group members but also the quality of mobile crowd sensor readings [5].

To address the shortcoming of existing group-forming mechanisms in MCS, this article presents a novel

The associate editor coordinating the review of this manuscript and approving it for publication was Zheng Yan.

Privacy-preserving Trust-Aware Group Formation (PTAGF) framework (PTAGF comes from Privacy-preserving Trust-Aware Group Formation). In PTAGF the sensing tasks are allocated to groups of participants, not individual participants. A heuristic approach is proposed where a group of trusted members is formed to generate a high-reputation group that performs group sensing activities with high-quality sensor readings while ensuring members' privacy. With this group framework, members in the assigned high-reputation groups will cooperate to perform the complex sensing tasks. The objective of this framework is to form dynamic groups, securely and privately, with highly reputable participants that achieve high sensing data quality.

To address the shortcoming of existing group-forming mechanisms in MCS, we propose a novel group-forming framework that ensures trust and privacy among the group member. In this framework, the participant's recruitment process considers multiple factors to select trusted members not only their social ties. Furthermore, participants' spatial-temporal information is protected during the task allocation process. To the best of our knowledge, this article is the first work that investigates how to form a high-reputation group with all trusted members to perform cooperative tasks along with preserving the privacy of these group members. In summary, the main contribution of this paper is the Privacy-preserving Trust-Aware Group Formation (PTAGF) framework. Overall, our contributions are as follows:

- We introduce a member trust assessment mechanism, which is designed based on the social relationship between the group members, their past behavior, their skills and readiness to perform the current task, and a group reputation that is based on its members' trust.
- We propose a novel group-forming mechanism based on the trust relationship, which is designed to ensure that all selected members are trusted and form a high-reputation group to provide high-quality sensor readings.
- We introduce a Two-Layered Privacy-preserving Mechanism (TLPM), which protects the privacy of members from other group members and malicious entities.
- We formulate the Trust-Aware Group Forming (TAGF) problem in MCS as a bicriteria optimization problem and prove its NP-hardness.
- A heuristic trust-aware group member recruitment algorithm is presented that is based on the concept of trust-relationship between members within the group and can be realized within a limited time complexity.
- We conduct theoretical analysis and extensive simulation using a real-world dataset that shows the proposed framework outperforms the state-of-the-art approaches.

This article is organized into the following sections: Section II reviews related works from two aspects; group-based recruitment systems in MCS, and the privacy-preserving mechanisms in MCS. Section III presents the system model, the motivating example, and the related threat to privacy. Section IV introduces the problem of group formulation in MCS and proves the *NP*-hardness of the problem. Privacy-preserving Trust-Aware Group-based Framework (PTAGF) and its underlying mechanisms are introduced in Section V. Section VI presents a privacy and security analysis of the PTAGF framework. In Section, VII PTAGF is evaluated by conducting various experiments and simulations. Finally, we conclude the article in Section VIII.

II. RELATED WORK

In this section, we review related works from two aspects. The first one is a group-based recruitment system in MCS, and the second one is privacy-preserving mechanisms in MCS.

A. ON THE ASPECT OF A GROUP-BASED RECRUITMENT SYSTEM IN MCS

For the multiple cooperative task (MCT) scenarios, where each sensing task requires more than one participant to complete cooperatively, many studies attempt to recruit groups of participants with reasonable sensing qualities. A variety of grouping algorithms are adopted in different systems to recruit optimal participants who can complete the sensing task with minimum cost.

Group forming schemes that are based on genetic algorithm (GA) are introduced in multiple studies to tackle different problems [2], [8], [9], [10], [11]. In [2], the authors propose a group-based recruiting (GRS) approach that achieves higher collective *QoI* from a selected group. In [10] they employ a genetic algorithm to allocate a group of workers to clusters.

Integrating a greedy algorithm and a genetic algorithm in group-based recruitment mechanisms is presented in [11]. On the other hand, Zhao et al. [4] tackle the reliability-oriented socially-aware crowdsourcing (R-SAC) problem. They propose a greedy maximum reliability user recruitment (G-MRUR) approach to find a near-optimal solution for the R-SAC problem. However, genetic algorithm-based group forming mechanisms lack the optimization for sensing task completion time. Furthermore, it suffers from increased computation overhead with increasing the number of tasks.

In contrast, some schemes adopt a heuristic algorithm to form a reliable group of participants [3], [7], [12], [13]. In [3], the authors propose group-oriented cooperative crowdsensing (GoCC). However, these systems need to balance privacy-preserving with participant recruiting performance. Moreover, there is a trade-off between running time and performance, and also it suffers from high computation overhead. Utilizing the neural network method and the clustering algorithm to learn the similarity between users and group them based on this similarity is introduced by Xu et al. [5]. On the other hand, a greedy heuristic algorithm to solve the time-limited multitask allocation optimization problem is proposed in [14]. A simple sorting algorithm and dynamic programming algorithm are adopted in [15]. K-means clustering for efficient participants grouping and task grouping is introduced in [16] and [17], respectively.

Table 1 summarizes and compares the group recruitment systems and our approach in terms of members' privacy

TABLE 1. Comparison of group recruitment systems.

Ref.	Member privacy inside a group	Member privacy outside a group	Member trust	Group reputation	Social attributes	Energy efficiency	Crowdsensing	Task nature
[3]	✗	✗	✗	✗	✓	✗	✓	One-time
[4]	✗	✗	✗	✗	✓	✗	✗	One-time
[5]	✗	✗	✗	✗	✓	✗	✓	One-time
[15]	✓	✗	✗	✗	✗	✗	✓	Continuous
[8]	✗	✗	✗	✓	✗	✗	✓	One-time
[2]	✗	✗	✗	✓	✓	✓	✓	One-time
[9]	✗	✗	✗	✓	✓	✓	✓	Continuous
[10]	✗	✗	✗	✓	✗	✗	✗	Continuous
[12]	✗	✗	✗	✗	✓	✗	✓	One-time
[13]	✗	✗	✗	✗	✓	✗	✗	One-time
[7]	✗	✓	✗	✗	✗	✗	✓	One-time
[11]	✗	✗	✗	✓	✗	✓	✓	Continuous
[14]	✗	✗	✓	✗	✗	✓	✓	Continuous
[16]	✓	✓	✓	✗	✗	✗	✓	One-time
PTAGF	✓	✓	✓	✓	✓	✓	✓	One-time

inside and outside the group, trustworthiness management, group reputation, energy efficiency, mobile crowdsensing, and task nature. The task nature is divided into a one-time task, where the participant submits a single report; and a continuous task, in which participants' stability needs to be considered because they regularly submit reports.

B. ON THE ASPECT OF PRIVACY-PRESERVING MECHANISMS IN MCS

In recent years multiple methods have been proposed for preserving privacy in MCS. Some of the main methods to protect participants' privacy are differential privacy [18], [19], [20], [21], [22], obfuscation [23], [24], encryption [25], anonymization [26], [27], [28], and data exchange-based strategy [29], [30], [31]. However, some of these techniques either impose a high overhead on the participant's side or cause high information loss. Table 2 shows the comparison between these techniques.

One of the most commonly used techniques for preserving privacy is k-anonymity, which employs indistinguishable features to protect the spatial-temporal privacy of participants. In the k-anonymity protection model [32], the participants' information cannot be distinguished from at least a k-1 individual. In general, anonymity refers to a technique that generalizes real sensor readings with others or dummies [33].

Different schemes address identity and location privacy-preserving in location-based services (LBS) by adopting k-anonymity techniques [28], [30], [34], [35]. In the same way, [26], [36] address identity and location privacy-preserving in crowdsensing-based vehicular social networks (VSNs). Li et al. [15] study bidding privacy preservation incentive mechanisms in a temporally and spatially dynamic MCS system. The authors in [37] and [38] prevent users' location information from leaking and solve the homogeneity attacks problem. In addition, [37], [39] adopt the concept of personal k-anonymity, where the user can freely select his/her privacy level by varying the data anonymization level without knowing the preferences of others. Although [37] has a trade-off between QoS and privacy-preserving, [39] succeeds to achieve high service accuracy while realizing identity anonymity. On the other hand, [33], [40], [41], [42] adopt $k-$anonymity to preserve participants' spatial privacy neglecting temporal privacy. Compared to [41] and [43] schemes, the scheme in [7] minimizes the overall communication overhead.

Multiple studies are adopting k-anonymity with data exchanging strategy to increase the level of participants' privacy [44], [45], [46], [47]. In the data exchange strategy, instead of trusting a third-party (TTP), local and distributed approaches based on collaborating participants are applied. Christin et al. [29], [48], [49] adopt a privacy-preserving exchange strategy that enables participants to disclose their sensor readings without compromising their information privacy in different scenarios. In the same way, [50], [51] study the privacy concern in participatory sensing due to large amounts of collected readings from a large number of participants. Buttner and Huss [52] present two-path hiding strategies to preserve the privacy of sensor readings collected by vehicles. Similar to [29], [49], and [53], the authors in [47] and [54] assess the participant's contribution to the collaborative exchange privacy-preserving mechanism, and identify malicious users in MCS.

Most of the schemes that are based on exchanged strategies exchange the sensor readings with random participants, and these participants' reliability cannot be ensured. Hence, the sensor readings may be disclosed or discarded by malicious participants which will increase the probability of incomplete tasks. Furthermore, the exchange process with peers is performed in an approximately large area which may cause a delay in the reporting process. In addition, there is a trade-off between privacy level and the amount of exchanged readings.

Therefore, we select two techniques to strengthen our approach without imposing overhead on the participants' side or leading to information loss. In this article, we present two privacy-preserving techniques for MCS systems. We research privacy-preserving techniques by combining an exchange-based strategy and the k-anonymity technique. The first one allows the participants to exchange their sensor readings to hide their sensitive information from the service provider (SP) and the malicious entities. The second one allows a group of participants to submit their sensor readings to SP while achieving the k-anonymity privacy level for each group of members.

III. PRELIMINARIES

In this section, we first present the system model, trust and reputation, the motivating example, and the related threats to privacy.

TABLE 2. Comparison between privacy-preserving mechanisms.

Mechanisms	Centralized/ Distributed	Data change	Information loss	Computation resource
Differential privacy	Distributed	Yes	Medium	Medium
Obfuscation	Centralized\ Distributed	Yes	High	Medium
Encryption	Distributed	No	Low	High
Anonymization	Centralized	No	Medium	Low
Exchange-based	Distributed	No	Medium	Low

A. SYSTEM MODEL

We consider a typical participatory sensing model. In general, the MCS system includes three main components, the number of participants in the area of interest (AoI) that are willing to perform sensing tasks, a task publisher who generate sensing tasks and broadcast them to the participant, and the platform which connects the participant and task publisher, which has a vital role in the MCS processes. In our framework, these components work as follow: First, the task publisher generates and publishes a sensing task through the platform that the participants in the AoI voluntarily and actively participate in. The task is published with six dynamic constraints/requirements, which can be changed according to the application and task publisher's needs. Thus, the task is defined as a six-tuple $T = (L^t, d^t, N^t_{Min}, N^t_{Max}, MT^t, SA^t)$, where L^t is the sensing task location in latitude-longitude coordinates, d^t is the task submission deadline, N^t_{Min} and N^t_{Max} are the allowed maximum and the minimum number of members within a group for task t, MT^t is the minimum required trust of the participants for task t, SA^t is a set of sensors required by task t. Second, the platform distributes the sensing task and its requirements/constraints to the participants in the AoI. Next, the participants perform the sensing task without violating its constraints and submit the sensor readings back to the platform. Then the platform aggregates these readings and forwards them to the task publisher. After that, the task publisher gives feedback about these readings as a penalty or reward to the participants.

B. TRUST AND REPUTATION

The concepts of ''Trust'' and ''Reputation'' are used interchangeably in most studies, but we consider them in different terms. We use ''Trust'', to refer to a locally stored value that represents the probability that the participant is socially connected, previously interacted with other participants, and has the skill to perform the sensing task. Hence, this trust represents a distributed value that is updated by observing other participants. Moreover, it is stored locally which enables the participants to examine the opposite party's trust level without the need to contact the platform every time. However, we use ''Reputation'', to refer to a globally stored value that represents the synthesized probability that the participant is trusted, as perceived by all other participants of previous interactions [55]. It reflects the participant's successful accomplishment of the sensing task. The platform stores the reputation values of every participant in the system. Hence, participants should use trust value instead of reputation value to evaluate the trustworthiness of other participants. However, participants can always request other participants' reputations from the platform and use it as an initial trust value.

C. MOTIVATION

The motivations of this framework come from the concern that, groups exist in several MCS scenarios, which supports the lifecycle of real-world group activities (e.g., parties, meetings). Grouping is an important phase of the design space for mobile sensing, especially in social influence phenomena [1], [56], management, economics [57], and social networks [58]. In addition, grouping optimizes MCS systems' operation and design requirements. Furthermore, closer and trusted relationships can help users achieve a good quality task with lower resource consumption [13], [59]. However, the grouping system that only considers the participants' social attributes such as the participant's friends and family circle, proves to be inefficient [2], [9].

To illustrate this, we consider a simple MCS system with 6 participants, who want to form groups to collaboratively perform sensing tasks. We assume that the maximum and the minimum number of members within the group are 4 and 2, respectively. As shown in Table 3, all participants get a trust score. This trust score lies between 0 and 1 for each participant, where the lower the score, the lower trust the participant has from his/her neighbor participants, and the lower chance he/she gets to join a group. For simplicity, we assume all the participants receive the same trust score they give to their neighbor participants. Furthermore, all participants have a reputation value based on the trust score that they get from all the group members they interact with. These participants' reputation value affects the final collective group reputation. We assume that the group reputation will take the lowest reputation of the joined members [60] (later elaborated in Section V-A). Considering a group scenario, there are two approaches to forming a group of participants. The first scenario is based on the close relationship between participants (family, and friends). Table 4 shows the members of each group and the final group's reputation. As evident, some participants (e.g. E and F) cannot join any group although they have a good trust score. In addition, all the groups that are formed have a low group reputation. This is because a member with a lower score might severely affect and decrease the group's reputation score. However, if E and F were to join the groups instead of C and B, the group's reputation would be higher. This can be shown in Scenario 2, where the participants are grouped based on their trust score, as depicted in Table 5. In this scenario, E and F join some of the groups and achieve high-reputation groups such as G4, G7, G8, and G11. Based on this example, it is clear that forming a group of participants that is solely based on their relationship (as in Scenario 1) compromises the group's reputation and leaves out skilled participants from joining sensing groups. On the other hand, grouping based on a trust relationship (as in

TABLE 3. Participants information.

Participants ID (i)	Relationship (j)			Trust Score (i,j)	Reputation value (i)
	Friend	Indirect Friend	Stranger		
A	B, C	D	F, E	0.8	0.78
B	A, C	D	F, E	0.2	0.28
C	A, B, D	Non	F, E	0.1	0.13
D	C	A, B	F, E	0.3	0.39
E	Non	Non	A, B, C, D, F	0.7	0.57
F	Non	Non	A, B, C, D, E	0.9	0.98

* (i, j) means i trust value in j, which means how much i trust j.

TABLE 4. Scenario 1: Grouping when the only social relationship is considered.

Group	G1	G2	G3	G4	G5	G6
Members	{A,B, C,D}	{A,B, C}	{A,C, D}	{A,B, D}	{B,C, D}	{A, B}
Group Reputation	0.13	0.13	0.13	0.28	0.13	0.28
Group	G7	G8	G9	G10	G11	
Members	{A,C}	{A,D}	{B,C}	{B,D}	{C,D}	
Group Reputation	0.13	0.39	0.13	0.28	0.13	

TABLE 5. Scenario 2: Grouping when the trust score is considered.

Group	G1	G2	G3	G4	G5	G6
Members	{A,D, E, F}	{A,D, E}	{A,D, F}	{A, E, F}	{D,E, F}	A, D
Group Reputation	0.39	0.39	0.39	0.57	0.39	0.39
Group	G7	G8	G9	G10	G11	
Members	{A, E}	{A, F}	{D, E}	{D, F}	{E, F}	
Group Reputation	0.57	0.78	0.39	0.39	0.57	

Scenario 2) generates a high reputation group with trusted and skilled members.

Thus, it is critical to develop a grouping mechanism that considers the trust relationship between group members, to achieve high task quality, and form a high-reputation group. Hence, this paper proposes a trust-aware grouping mechanism.

D. THREAT MODEL

We assume an honest-but-curious attack model, where platform entities e.g., participants, SP, or task initiator, attempt to breach the participants' privacy. However, they will follow the protocol normally and faithfully, and they are not able to compromise cryptographic mechanisms. They may attempt to infer sensitive information from the reported sensor readings or link them to the identity of particular participants. Furthermore, a curious-but-honest SP may attempt to passively breach the privacy of the participants. Also, it has access to the sensor readings reported by the participants, thus it might infer participants' sensitive information. However, it does not launch an active attack, such as collusion with other participants to obtain other participants' sensitive information. In addition, as an artifact distributed and collaborative nature of the exchange-based strategy, group members can become possible adversaries, who are potentially interested in inferring other members' information. Malicious members can disclose other members' sensitive information by directly accessing the information on exchanged sensor readings. We also assume that the adversary entities can observe reported readings, through eavesdropping or being a malicious SP.

IV. PROBLEM FORMULATION

In this section, we formulate the Trust-Aware Group Forming (TAGF) problem in MCS as a bicriteria optimization problem, and prove its NP-hardness.

A. GROUP FORMING PROBLEM DESCRIPTION

In this subsection, we introduce the problem of group formulation in MCS and prove its *NP*-hardness. The parameters and notations used in this paper are summarized in Table 6.

We investigate the social relationship among participants in a quasistatic scenario, where participants are connected via a social network defined as a directed social graph as illustrated in Fig. 1.

Definition 1: The trust between participants is represented as a directed social graph, which is presented as a 3-tuple $TG = (V, E, \omega)$, where V is the set of participants, E is the edges connected participants associated with ω the weight on the edges, which represent random trust scores. Trust can be an asymmetric relationship. Each $E_{ij} \in E$ indicates that members i and j trust each other. There is no self-loop in the trust graph i.e., there is no edge $(i, i) \in E$. PTAGF should find a group of participants that optimize the following three factors:

- The members' trust value within the group;
- The group reputation score compared to other groups;
- The quality of sensor reading concerning the time constraints.

Let a crowd of participants is organized into n groups, $n \geq 1$. Given a complex task t with constraints, i.e., the number of members, and a deadline for the task submission d^t. Let a set of reliable and skilled participants $R(t_k)$ participate in executing the task t, $R(t_k) \subseteq V$. We use $MT(i, j)$ to denote the trust value between group members (i, j); GR denotes the group reputation and QoI denotes the quality of sensor reading submitted by the group $R(t_k)$.

In summary, the problem of in-aware groups forming in MCS can be described as follows:

- The trust among members within the group, influences the recruiting and performance of a group in MCS tasks. Moreover, MCS needs timely and high-quality sensor readings. Then, how can a group's reputation and its member trust affect the quality of sensor readings *(Group Reputation and Member Trust)*?
- The execution of a complex sensing task needs a group of trusted participants. Then, how is an efficient

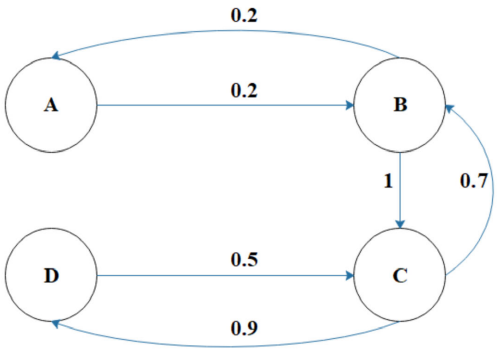


FIGURE 1. Directed social graph.

mechanism to recruit these participants in a group to be designed *(Trust-Aware Group Formulation)*?

B. COMPLEXITY ANALYSIS

Problem 1: With the above definition, we formulate Trust-Aware Group Formation (TAGF) problem and prove its *NP*-hardness. Given a social trust network $TG = (V, E, \omega)$, and a set of complex tasks $T = \{t_1, t_2, \ldots, t_c\}$ $c \geq 1$, and a member i can participate in these complex tasks. We assume every participant in $R(t_k)$ has the required skill to perform t_k. The TAGF problem is to find a group of participants $\acute{V}$ to perform the sensing tasks, such that $\acute{V} \subseteq V$. Furthermore, $N_{Min} \leq |\acute{V} \cap R(t_k)| \leq N_{Max}$ $\forall t_k \in T$. The problem aims to maximize the trust value. Mathematically, TAGF can be formulated as follows:

$$Maximize\ MT(\acute{V}) \tag{1}$$

$$Subject\ to: \acute{V} \subseteq V \tag{2}$$

$$N_{Min} \leq |\acute{V} \cap R(t_k)| \leq N_{Max} \quad \forall t_k \in T \tag{3}$$

$$MT(\acute{V}) = MT(i, j), \quad i, j \in \acute{V} \tag{4}$$

Proposition 1: The TAGF problem is NP-hard.

Proof: We prove the TAGF problem is NP-hard by a reduction from the bicriteria optimization problem, which was proven to be NP-hard [61]. An instance from the bicriteria optimization problem is defined as follows: Given a directed graph $G = (L, E)$ with a set of nodes $L = \{c_1, c_2, \ldots, c_m\}$, sub-graphs $\{S_1 = (\mathcal{F}_1, \mathcal{E}_1), S_2 = (\mathcal{F}_2, \mathcal{E}_2), \ldots, S_n = (\mathcal{F}_n, \mathcal{E}_n)\}$ such that $\mathcal{F}$ is a set of nodes in the federation and $\mathcal{E}$ is a set of trusted edges among nodes and two matrics $\acute{\rho}(\mathcal{F})$ and $\psi_c(\mathcal{F})$ the former represents the average reputation for nodes in the federation and the latter represents the individual payoff of node c in $\mathcal{F}$. The bicriteria optimization problem is to choose a federation with both maximum individual payoff, and average reputation:

$$\begin{cases} max(\mathcal{F}) & \psi_c(\mathcal{F}),\ and \\ max(\mathcal{F}) & \acute{\rho}(\mathcal{F}) \end{cases} \tag{5}$$

TABLE 6. Notation and key parameters.

Notation	Definition	Notation	Definition
T	Set of complex tasks	$TG = (V,$ E,$\omega)$	Directed social graph (trusted graph)
L^t	Task t location	V	Set of vertices representing participants
d^t	Task deadline	E	A set of edges represents a trust connection among participants
N^t_{min}	Minimum number of members within a group for task t	ω	The weight of edges represent the trust value
N^t_{max}	Maximum number of members within a group for task t	E_{ij}	Edge connecting participants i and j
MT^t	Minimum required trust for task t.	$R(t_k)$	Set of participants that have the skill to perform the task t_k.
SA^t	Set of sensors required by task t.	$\acute{V} \subseteq V$	A group composed of chosen members
QoI	Quality of information	MT	Member trust
GR	Group reputation	MT_{old}	Old member trust
I	Set of participants	MR	Member readiness
L_i	Participant's location	DSA_i	Participant's device sensor availability
Rr	Reliability ratio	DRE_i	Participant's device residual energy
SR	Social relationship	P_i	Member's reputation
$\acute{MT}(i, j)$	Normalized member trust	C_i	Number of delivered sensor reading
$H(i)$	Set of i's neighbors within the group.	P_k	Participant' neighbors' reputation
G	Group	τ^t_G	Decreasing function concerning the time
ST^t_G	Time needed by group G to submitthe sensor readings	k	The desired anonymity level
r	Reported sensor readings	$\acute{r}$	Anonymized data
Z	Set of information received by members	$\acute{Z}$	Set of anonymized information received by members
F	Mapping function	B_{cl}	Cloaking box

We transfer an instance of the bicriteria optimization problem to an instance of the TAGF problem. We have $TG = (V, E)$ and a sub-set $\acute{TG} = (\acute{V}, \acute{E})$ such that we consider $\acute{V}$ as a set of participants, the set of chosen group members needs to satisfy the maximum trust value and to join a group with a high average reputation.

$$\begin{cases} max(\acute{V}) & MT_i(\acute{v}),\ and \\ max(\acute{V}) & GR(\acute{V}) \end{cases} \tag{6}$$

V. PRIVACY-PRESERVING TRUST-AWARE GROUP-BASED FRAMEWORK (PTAGF)

This section presents and discusses the PTAGF in detail and its underlying mechanisms as depicted in Fig. 2

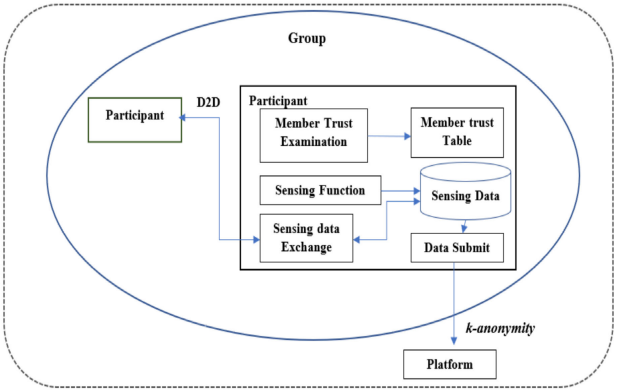


FIGURE 2. General framework of PTAGF.

A. MEMBER-TRUST AND GROUP-REPUTATION (MT-GR) MECHANISM

In this sub-section, we propose the Member-Trust and Group-Reputation (MT-GR) mechanism, which helps in forming a high-reputation group with trusted members and demotivates malicious participants. During the registration with the platform, each participant is defined as a five-tuple, $I = (L_i, DS_i, DRE_i, P_i, C_i)$, where L_i is the participant location in latitude-longitude coordinates. DSA_i and DRE_i denote the available set of sensors and the battery level of the participant's device, respectively. P_i is the participants' reputation computed from their previous behavior, and C_i represents the number of successfully delivered sensor readings by the participants. After the participant is allocated a complex task, he/she starts forming a group with a high reputation.

1) MEMBER TRUST (MT)

Using definition 1, we update the MT which is represented as the weight on the edge between vertices in Fig. 1. The MT is based on multiple factors, which are:

Old Member-Trust (MT_{old}): The past behavior of member i. In the case that two members did not have any interaction in the past, they can fetch each other reputation P_i from the global database on the platform.

Member Readiness (MR): Represents the readiness of the member device to perform the sensing task. This factor is device-related, and it is computed by the platform based on two parameters [2], [9], [11].

Device Residual Energy (DRE)*:* is a parameter that measures the battery levels, and updates dynamically during the sensing task.

Device Sensor Availability (DSA)*:* We assume that the platform is aware of each device's sensors when the participant registers to the system. We define the following decision variables:

$$MR_{(DRE,DSA)} = \begin{cases} 1 & \text{if } DRE > \text{predefined battery level required for the task and } DSA \text{ available} \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

Reliability ratio (Rr): We use Rr to compute the ratio of the successfully delivered sensor readings to the total number of exchanged ones [54]. The set of successfully delivered sensor readings C_i is updated regularly and stored on the global database. The rate of member reliability is given by:

$$Rr_i = \frac{(|\text{set of successfully delivered sensor readings } C_i)|)}{(|\text{set of total exchanged sensor readings}|)} \quad (8)$$

Social Relationship (SR): This factor indicates how closely the members of the group are socially connected. The social relationship can take a value between 0 and 1, (i.e., $SR \in [0, 1]$), where 1 is the highest value.

Finally, participants rate each other based on their MT, by the following formula:

$$MT = (w_1 \cdot MT_{old}) + (w_2 \cdot MR) + (w_3 \cdot Rr) + (w_4 \cdot SR) \quad (9)$$

In this formula, these weights $[w_1, w_2, w_3, w_4]$, are determined by the task publisher. They specify the importance of each parameter, such that $w_i \geq 0 \ \forall i$. The value of MT is normalized between 0 and 1, as follows:

$$\acute{MT}_{(i,j)} = \frac{MT(i,j)}{\sum_{k \in H(i)} MT(i,k)} \quad (10)$$

Were $H(i) = \{j| \ \exists (i,j) \in E\}$ is the set of $i's$ neighbors within the group. Finally, the member trust must be greater or equal to the MT^t. After that, the platform calculates the reputation score of member i, denote as $P(i)$ which reflects the summation of all assigned MT to i by its neighbors, a following:

$$P_i = \sum_{\forall_{k \in H(i)}} MT_{(k,i)} \cdot P_k \quad (11)$$

The reputation score for each registered member will be updated regularly each time his/her group submits a sensor reading.

2) GROUP REPUTATION (GR)

The group's reputation is determined by the lowest reputation score among its members [60] since a member with a lower reputation might severely affect and decrease the group reputation score.

$$GR = min_i\{Pi\} \quad i \in G \quad (12)$$

Furthermore, this prevents groups from affecting the final sensing results by allowing bad reputation members to join the group just to gain benefits from this group.

3) QUALITY OF INFORMATION OF GROUP

In our framework, we are based on collective quality rather than individual quality. Therefore, the QoI of the group-submitted sensor readings is evaluated based on the group's reputation, and the time required for the group to complete

and submit the sensor readings. QoI of the highest reputation group is calculated using the following equation:

$$QoI = GR \cdot (\tau_G^t)^2 \tag{13}$$

where the τ_G^t is a decreasing function concerning time, and is calculated as:

$$\tau_G^t = min[1 - max(0, min[\log_{d^t}(ST_G^t), 1])] \tag{14}$$

where ST_G^t is the time in seconds needed for group G to submit the sensor readings. It is calculated as the difference between the time when the task is allocated to group G until the group submits the sensor reading. On the other hand, d^t is the time constraint (deadline) of task t. Thus, the objective is to maximize the QoI of the group-reported sensor reading and it can be formulated as:

$$QoI_G = Max_{G \in V}(QoI_G^t) \tag{15}$$

B. TRUST-AWARE GROUP FORMING (TAGF)

In this sub-section, a trust-aware group formation mechanism as a recruitment and task allocation approach is proposed. Since the TAGF problem is NP-hard, we propose a heuristic algorithm to solve the TAGF problem.

The proposed Trust-Aware Group Forming (TAGF) is given in Algorithm 1. TAGF has three input parameters: the set of tasks T along with their requirements, the social relationship graph, *socialRelMatrix*, and the trust values between participants, *trustMatrix*. The TAGF uses G to keep a set of group members and it is initialized with an empty set. Then, it initializes V with a set of populations. The group reputation GR and the QoI are initialized with size zero. First, TAGF chooses to start the group with a member who has the highest reputation in the area of the published task (line 4). After that, it iteratively selects members who satisfy the member trust threshold of the task MT^t till the maximum group size is reached N_{Max}^t, (lines 5-19). In the case that no candidate satisfies the MT^t threshold and at the same time the group size is below the N_{Min}^t, the group will not be formed and the task will be aborted, (lines 20-22). In the second case, if TAGF has enough members that satisfy N_{Min}^t, but it does not find the next member that satisfies MT^t (lines 23-25), then it will skip to (line 30). In the third case, it finds the next member with a high or equal member trust to MT^t, then it adds this member to the group (lines 26-28). The iterations continue until TAGF finds a set of groups that reach the N_{Max}^t requirements with all trusted members. Then, it calculates the group's reputation which tasks the lowest reputation of the joined members (line 30). In addition, it calculates the QoI of the group (lines 31-33). Finally, TAGF returns a set of formed groups G, their group reputation GR, and their QoI. For the group selection part, Algorithm 2 receives the set of groups that were formed by Algorithm 1, their GR, and their QoI. Then it selects the highest group reputation among these groups (line 1). After that, it recruits this group to complete the sensing task and returns its QoI.

Algorithm 1 Trust-Aware Group Forming (TAGF)

Input: task T requirements $T = (L^t, d^t, N_{Min}^t, N_{Max}^t, MT^t, SA^t)$, *socialRelMatrix*, *trustMatrix*
Output: G, GR, QoI

1: **Initialization:**
2: $G \leftarrow \phi$, $V \leftarrow$ initial set of population, $GR = 0$, $QoI = 0$
3: **End of Initialization:**
4: $G =$ {highest Reputation Member in V} // *Chose the first member i to start a group*
5: **while** $(G < N_{Max}^t)$ **do**
6: Calculate MT based on (9)
7: Normalized MT based on (10)
8: Recalculating member's reputation based on (11)
9: $nextMemberMT = 0$, $nextMember = 0$;
10: **for** $i = 1$ to number of members **do** // *Select highest trust candidate*
11: member=members(i)
12: **for** $j = 1$ to candidates **do**
13: candidate = candidates(j);
14: **if** $normMT$ (member,candidate) $>$ $nextMemberMT$ && $normMT$ (member, candidate) $\geq$ MT^t **then**
15: $nextMemberMT = normMT$(member, candidate);
16: $nextMember = candidate$;
17: **end if**
18: **end for**
19: **end for**
20: **if** $nextMember == 0$ **then** //*No candidate was found*
21: **if** $G < N_{Min}^t$ **then**
22: **Error:** Not enough members with trust value $> MT^t$;
23: **else**
24: break while (skip to line 30)
25: **end if**
26: **else** // *A candidate was found*
27: $G = GU\ nextMember$; //*add candidate to G in position i*
28: **end if**
29: **end while**
30: $GR = min(reputation(G(member)))$;
31: $logD = log(ST)/log(deadline)$;
32: $tG = 1 - max(0, min(logD, 1))$;
33: $QoI = GR * (tG)^2$;
34: **return:** G, GR, QoI

Algorithm 2 Group Selection

Input: G^*, GR^*, QoI^* //*Sets of Groups and their respective GR and QoI*
Output: Chosen high reputation group G, QoI

1: $indexBestGroup =$index of max(GR^*)
2: **return:** $G^*[indexBestGroup]$, $QoI^*[indexBestGroup]$

C. TWO-LAYERED PRIVACY-PRESERVING MECHANISM (TLPM)

To ensure the participants' privacy we propose a two-layer privacy-preserving mechanism (TLPM). TLPM starts two layers of defense to protect the participants' privacy. The first layer is the exchange-based strategy, where each member within a group exchanges their sensor reading with other encounter group members. The purpose of this exchange is to jumble and hide the source information that is tagged with the sensor reading. Making it challenging to guess which sensor reading belong to which member, the second layer is based on the anonymization approach, by adopting the k -anonymity technique. Thus, to realize k -anonymity after each group formed with at least $k \geq N_{Min}$ members, each member within the group sends his/her sensor data to the server as anonymized data. The k-anonymity technique is good at balancing service quality and privacy preservation [37], but it is difficult to achieve optimal anonymity. Moreover, it is insufficient to consider only k-anonymity as privacy protection, because some groups may contain or form with the minimum number of participants, N_{Min}, so there is a need for another layer of defense without extra overhead on the participants' side.

D. EXCHANGE-BASED PRIVACY LAYER

As the first layer of privacy protection, we adopt exchanging strategy [48], [54]. After forming a group, the members within the same group start exchanging their sensor readings with their encounter peers. This exchange process unlinks the sensitive information tagged with sensor readings from the source identity. Before exchanging, each member encrypts his/her readings with the server's public key to prevent other members from disclosing its content. The exchange is performed using a realistic exchange/ complete exchange approach [29], [48], [49], where all the sensor readings that are collected by one member are exchanged entirely at encounters. Adopting this approach is more suitable for our system because it appears as if the readings are captured by a single real member. Furthermore, the exchange performs asymmetrically [48], where the amounts of exchanged readings differ between both encounter members. Additionally, the reporting process to the platform adopts a time-based approach [48] according to the task deadline.

E. k-ANONYMITY PRIVACY LAYER

Exploiting personalized k-anonymity [37], we apply a group-based personalized privacy-preserving method to protect group members' information. As we mentioned in Section V-B, each sensing task has N^t_{Min} and N^t_{Max} requirements, thus the group will realize k-anonymity with at least N_{Min}-anonymity. Hence, each group can determine its privacy level. The anonymity server within the platform transforms the reported sensor readings r from each member i into anonymized data $\acute{r}$. The set of information received from each group member denotes Z. Thus, the sensor readings of member i in set Z are shown as:

$$r_i \in Z : \langle i, \{t_i, x_i, y_i\}, k\rangle$$

where t_i represents the timestamp at which the sensor reading is captured, x_i and y_i represent the longitude and latitude coordination. The k value indicates the group member number which represents the desired anonymity level. Hence, k must satisfy two conditions; $k > 1$, and $N_{Min} \leq k \leq N_{Max}$. Thus, a large group size indicates a large k value, which indicates a high privacy degree. Each member's information is assigned a cloaking box to be undistinguished from at least $k-1$ other group members' information. Accordingly, to preserve the group member's information we define $\acute{r}$ as the anonymized information. The set of anonymized information for each group member is denoted as $\acute{Z}$:

$$\acute{r} \in \acute{Z} : \langle i, \{[t_i^{minV}, t_i^{maxV}], [x_i^{minV}, x_i^{maxV}], [y_i^{minV}, y_i^{maxV}]\}\rangle$$

Therefore, there exists a mapping between r and $\acute{r}$, such that $\acute{r} = F(r_i)$. In addition, there exists a cloaking box such that B_{cl} $(\acute{r})$ is a spatial-temporal cloaking box for i member's information.

$$B_{cl}(\acute{r_i}) = ([t_i^{minV}, t_i^{maxV}], [x_i^{minV}, x_i^{maxV}], [y_i^{minV}, y_i^{maxV}])$$

By using this method, r in Z and $\acute{r}$ in $\acute{Z}$ satisfy spatial and temporal containment. In particular, the spatial-temporal box $B_{cl}(\acute{r_i})$ contains the real position of the member within the group. Furthermore, this mechanism is spatial-temporal k-anonymity.

VI. SECURITY ANALYSIS

Our security analysis focuses on how the TLPM achieves participants' identities and spatial-temporal privacy preservation and provides strong protection against attacks.

1) PRIVACY AGAINST SP AND MALICIOUS ENTITIES

The service provider and other platform entities have access to all the reported sensor readings. In addition, they can perform as honest-but curies attackers to disclose the information of group members.

Lemma 1: The participant's identity privacy is preserved.

Proof: The group members' privacy is evaluated as the level of anonymity of their sensor readings. The number of members in a group is at least k, thus the SP and the adversary guessing the identity of the member is with probability $(1/k)$. In addition, each group will change its k-anonymity dynamically according to the task requirements. Furthermore, the exchange strategy makes it harder to disclose the participant's identity as group members constantly exchange their sensor readings with encounter peers. The probability of disclosing the participant's identity is far less than $(1/k)$. Thus, identity privacy is realized, because the best knowledge other entities can know is that the contributor of the sensor reading is from one of the k group members.

Lemma 2: The participant's spatial-temporal privacy is preserved.

Proof: Adopting the k-anonymity mechanism leads to spatial-temporal k-anonymity. The k-anonymity demands that for the anonymized information of a member there exists at least $k-1$ other anonymized information within the same group. Furthermore, the exchange strategy hides the source information of sensor readings. Thus, the member's precious spatial-temporal information gets mixed with the exchanged/reported member. In addition, as the exchange is done with multiple peers before reporting to the server, the privacy level will increase.

2) TLPM MAINTAINS HIGH ACCURACY OF THE SENSOR READINGS

The exchange strategy may cause dropping or tampering with the sensor readings before reporting them. Moreover, the k-anonymity may degrade the spatial-temporal accuracy of the submitted readings. This can happen if the submitted sensor readings have a high anonymized tolerance, i.e., the anonymized spatial-temporal is far from the original one.

Lemma 3: The proposed TLPM is truthful.

Proof: During the exchange-based layer of the TLPM, the group members are already verifying each other trustworthiness, hence the exchange of sensor readings will be performed with reliable members. Accordingly, the member within the group will behave truthfully to join a high-reputation group. If they behave untrusted or unreliable, their member trust (MT) degree will decrease, which leads to being removed from the group.

Lemma 4: The TLPM is spatial-temporal containment.

Proof: It signifies that in the k-anonymity privacy layer, the spatial-temporal box $B_cl(r'_i)$ contains the real position of the member within the group. For that, if the group member's real location and timestamp of his/her sensor readings are $i : \{t_i, x_i, y_i\}$, the spatial containment satisfies; $x_i \in [x_i^{minV}, x_i^{maxV}]$, $y_i \in [y_i^{minV}, y_i^{maxV}]$ and temporal containment satisfies; $t_i \in [t_i^{minV}, t_i^{maxV}]$.

3) RESISTANCE TO INTERNAL AND EXTERNAL ATTACKS

An external attack eavesdrop on the communication process between the group member and the SP. Thus, it can extract sensitive information and disclose the content of the sensor readings. Furthermore, during the exchange process, the encounter members have full access to the exchange readings. Hence, they may acquire sensitive information about other members.

Lemma 5: TLPM is protected against external attacks.

Proof: An example of an external attack is the eavesdropping attack. It is impossible to determine the real contributor of the sensor readings that correspond to the eavesdropped information, due to the use of an exchange strategy and encrypting them with the SP public key.

Lemma 6: TLPM is protected against internal attacks.

Proof: The internal attacks can be launched by both SP or participants, such as inference attacks. The group member that exchanges the sensor readings with others cannot decrypt these sensor readings, because they are encrypted with SP public key before exchanging them. Moreover, even if the SP has full access to these readings reported by participants it cannot acquire accurate spatial-temporal information, due to the k-anonymity and exchange strategy. The cloaking group region has at least k members, who already exchanged their reading multiple times before reporting them to the SP.

VII. PERFORMANCE EVALUATION

In this section, PTAGF is evaluated by conducting various experiments and simulations. All mechanisms are implemented in MATLAB R2021b, and experiments are conducted on a machine equipped with Intel(R) Core (TM) i7-8565U CPU and 16.0 GB RAM, running on Windows 10 (64-bit).

A. DATASET

To evaluate the performance of PTAGF, a real-world location-based social network dataset Gowalla [62] is used. The dataset collects 6.4 million locations (i.e., where users share their locations by checking in) and a friendship network. It consists of 196,591 nodes (i.e., participants) and 950,327 edges (i.e., friendship), throughout Feb. 2009 - Oct. 2010.

B. COMPARISON APPROACHES

Our simulation compares PTAGF performance with two representative group recruitment approaches, GoCC [3] and G-MRUR [4]. Both approaches address the MCTs in MCS, hence finding reliable groups to complete these tasks. Our proposed framework also finds a reliable group with a high reputation with Algorithm 1 described in Section V.C to perform these tasks. Moreover, both approaches consider the social ties to recruit group members, which is an important factor in our group-forming mechanism.

- **GoCC-G**: This approach [3] is a heuristic-based participant selection. It is mining the social network to recruit participants with closer relationships, Algorithm (4).
- **G-MRUR**: This approach [4] is a greedy-based user selection, which finds an approximation solution Algorithm (1). It recruits reliable workers, by considering both matching between task type and recruited worker's interest and reliability feedback.

C. PERFORMANCE METRICS

To evaluate PTAGF comprehensively, the following metrics are used:

- **Trustworthiness in group selection**: Measured by the total trust scores of all the group members, the higher the better.
- **Task coverage percentage**: Measured by the ratio between the number of completed tasks to the total number of tasks, the higher the better.
- **Running Time** : Measured by the time taken by an approach to find the solution, the lower the better.

D. SIMULATION SETUP

In the simulation, we use part of the original Gowalla dataset and select the participants and their social ties randomly to

TABLE 7. Simulation setting.

	n	Nmax	Nmin	S
Set #1	300,. . . ,900	[5,30]	[3,10]	10
Set #2	600	[5,10],...,[5,50]	[3,10]	10
Set #3	600	[5,30]	[3,5],. . . ,[3,25]	10
Set #4	600	[5,30]	[3,10]	5,. . . ,25

construct different sub-network similar to [5] and [63]. The trust value, i.e. *trustMatrix* in Algorithm 1, is represented by the weight of the corresponding edges among participants on the social graph. It is uniformly distributed in the interval $\in$ [0, 1]. However, similar to [64] and [65] if two participants do not have a social relationship originally, in this case, we set an initial connection weight between the two unconnected users as 1. After that, their trust in each other will decrease or increase for each cooperation according to eq.(9). By doing so, we examine if our approach can select trusted participants within the AoI even if they are not socially connected. Furthermore, we generate the factors: number of exchange sensor readings, residual energy, and sensor availability, following the distribution within the ranges [0,50], [0,100], and [0,1], respectively. To comprehensively evaluate the performance of PTAGF four parameters are varied in the simulation for different scenarios. These settings have been proven to be reasonable setting in many practical scenarios [5], [6].

- **Community size** (n): The number of users within the AoI varied from 300 to 900 in steps of 100.
- **Group maximum members** (N_{Max}): Uniformly distributed in a specific interval, which varies in [5, 10], [5, 20], [5, 30], [5, 40], and [5, 50].
- **Group minimum members** (N_{Min}): Indicates that a task requires at least N_{Min} to perform, uniformly distributed in [3, 5], [3, 10], [3, 15], [3, 20], and [3, 25].
- **Task scale** (S): This is the number of published sensing tasks. S varies from 5 to 25 in the step of 5.

Finally, Table 7 summarizes the parameter settings by varying each parameter while fixing the other to simulate different scenarios. In the simulation, the average is taken over 100 runs.

E. PERFORMANCE COMPARISON

In this subsection, we vary the values of key parameters to explore their impacts on the group-forming systems. The simulation results of comparing PTAGF with the two approaches, GoCC-G and G-MRUR, are reported.

1) IMPACT OF COMMUNITY SIZE (n)

To evaluate the scalability of the designed group-forming mechanism, we vary the community size (n) from 300 to 900 in steps of 100. Fig. 3 shows the performance of different approaches under various n.

As expected in Fig. 3a, the trustworthiness of selected members achieved by all three approaches increases with increasing n. However, PTAGF outperforms GoCC-G and G-MRUR by 51.61%. This demonstrates that the PTAGF can form a group of trusted and reliable members, and that increases the group's reputation. This is because more participants are available to choose from. As a result, the group formed by PTAGF more likely contains more trusted and reliable members that can perform the sensing task with higher quality.

From Fig. 3b we can observe that with increasing n, the performance of PTAGF in terms of task coverage may not be the best. However, it performs better than GoCC-G, with 83.33%. PTAGF verifies the trust of each participant before allowing them to join the formed groups, by considering multiple factors as demonstrated in (9). Hence, the total number of members in the selected group may be lower than G-MRUR. Consequently, the task coverage of PTAGF when increasing n is lower than G-MRUR. However, PTAGF has reasonable task coverage that can be performed with more trusted members and highly reputable groups to give high-quality sensor readings.

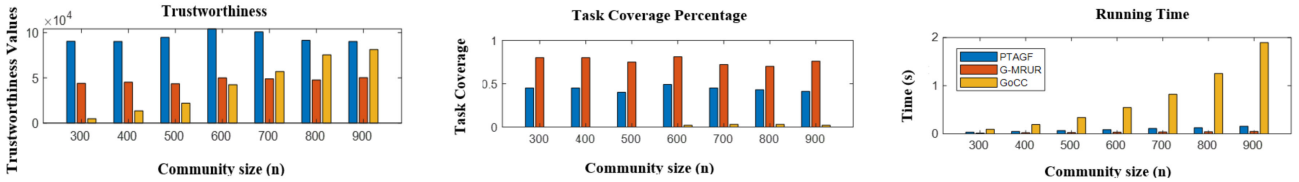
Fig. 3c illustrates that as the community size increase, the running time of all three approaches increases. This is because, with an increase of n more participants are available in the AoI to choose from, resulting in increasing the running time in the member recruiting process. As illustrated in Fig.3c, PTAGF is much more efficient than GoCC-G, but it takes slightly more time than G-MRUR. It should be noted that PTAGF verifies each member's trust, as elaborated in Section V-A. As a consequence, PTAGF has a slightly higher time cost to recruit more trusted participants. Indeed this considers a small price PTAGF pays for recruiting trustworthy group members.

2) IMPACT OF GROUP MAXIMUM MEMBERS(N_{Max})

Fig. 4, depicts the impact of the group size on the performance of the three group-forming approaches. From the experimental results, it can be seen in Fig.4a that all three approaches' trust members' values increase with increasing the N_{Max}. However, PTAGF has the best performance in terms of the trustworthiness of the group members when the N_{Max} increases above the [5, 30] scale. On average, PTAGF outperforms GoCC-G, and G-MRUR by 42.85% and 33.3%, respectively. These results demonstrate that even if the sensing task requires a maximum group size to complete the task, PTAGF always recruits trusted members to form these groups.

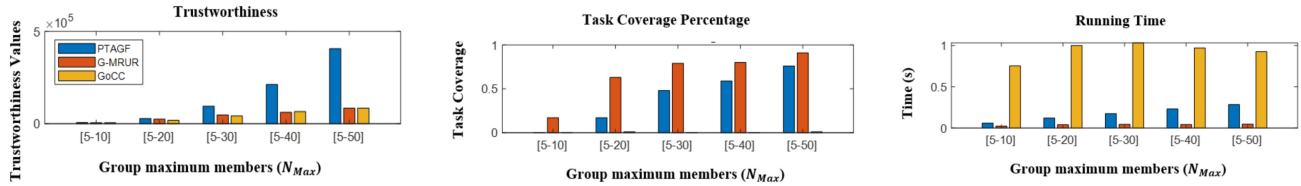
As the group size is maximized, PTAGF and G-MRUR have a high task coverage than GoCC-G as shown in Fig. 4b. Although the PTAGF task coverage level is lower than GMRUR, its approaches the GMRUR when increasing N_{Max}, and it increases proportionally. This signifies that PTAGF increases the number of trusted members and that leads to an increase in the tasks coverage percentage.

As observed in Fig. 4c, the running time increases with the group size. It clearly shows that PTAGF has a slightly higher time cost than G-MRUR. The reason is that as N_{Max} increases there are many members to verify their trust level before allowing them to join the formed group. This emphasizes the



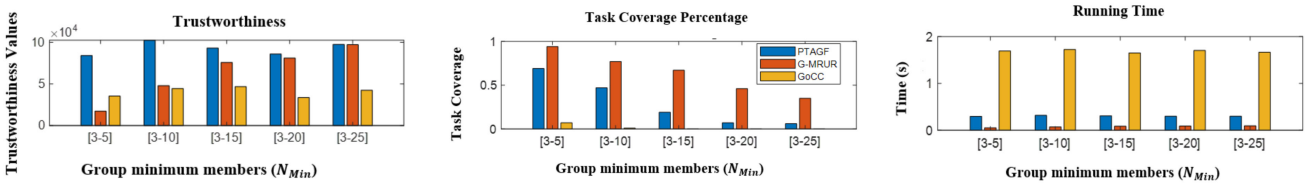
(a) (b) (c)

FIGURE 3. Impact of community size (*n*).



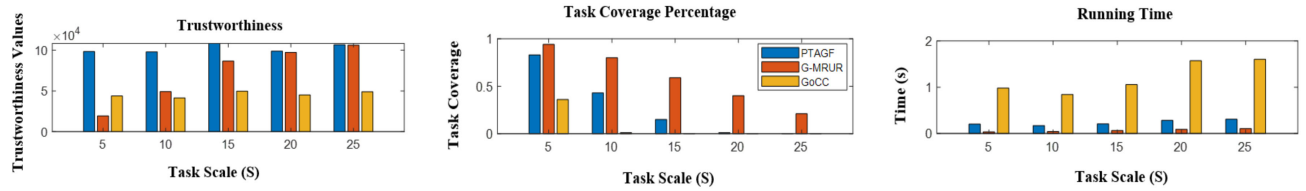
(a) (b) (c)

FIGURE 4. Impact of group maximum members (N_{Max}).



(a) (b) (c)

FIGURE 5. Impact of group minimum members (N_{Min}).



(a) (b) (c)

FIGURE 6. Impact of task scale (*S*).

small impact of maximizing the group size on the running time of PTAGF. The PTAGF can still find the most trusted and reputable groups to complete the sensing tasks at a lower running time, of less than 0.3 s.

3) IMPACT OF GROUP MINIMUM MEMBERS (N_{Min})

Fig. 5 illustrates the impact of N_{Min} on the performance metrics for all three group-forming mechanisms. Fig. 5a shows the average trustworthiness level for each group-forming algorithm when N_{Min} increases. As evident, PTAGF still outperforms GoCC-G and G-MRUR in terms of grouping more trusted members even if we have a small group size requirement. The results of PTAGF are 59.09% higher than GoCC-G and 29.54% than G-MRUR, respectively. As expected, PTAGF recruits more trusted members to form these small groups to complete the sensing tasks with high quality.

Fig. 5b shows the results on average task coverage percentage varying N_{Min} values. Here, PTAGF does not perform well in task coverage compared to G-MRUR when fixing the community size. The potential reason is that as the range of N_{Min} increases, it is difficult for existing groups to meet participants who satisfy the member trust threshold with fixed n. However, PTAGF performs better than GoCC-G in task coverage even with small group sizes.

Fig. 5c illustrates a running time comparison of the three algorithms under different sizes of N_{Min}. Indeed, the running time of the three approaches increases with increasing of

N_{Min}. As shown in Fig. 5c, the running time of PTAGF is slightly higher than G-MRUR. Since we need to find trusted candidate participants according to several factors, as elaborated in Section V-A, this may induce a longer running time. However, the running time of PTAGF is still acceptable in a real-world scenario, and this is a price PTAGF pays to choose the best and most trusted members.

4) IMPACT OF TASK SCALE (*S*)

The task scale can depict the workload of the group-forming approaches. Fig. 6 shows the comparison between the three approaches under various S. From Fig.6a we observe that PTAGF outperforms GoCC-G and G-MRUR in recruiting more trusted members when the number of published tasks increases. As S increases PTAGF needs more reliable participants to complete the sensing tasks. Although, when more than 15 tasks are published, PTAGF trustworthiness in members selection decreases slightly because of a fixed number of n, which leads to a lack in the number of trusted members to recruit from. Accordingly, it may increase the probability of selecting lower trusted members, thus the trustworthiness value of PTAGF decreases with increasing S. However, on average, PTAGF outperforms GoCC-G and G-MRUR in terms of trustworthiness in member selection by 66.67% and 33%, respectively.

From Fig. 6b it should be noted that the task coverage of all three approaches decreases with increasing task scale. The reason is that with increasing S while fixing n and N_{Max} there will be a lack in the number of members to cover all the published tasks. In addition, the experiment results show that PTAGF outperforms GoCC-G in terms of task coverage with increasing S. Also, as expected the task coverage percentage for PTAGF is decreased, because there will be not enough trusted members in the AoI who can join a reputable group. This decrease in the task coverage demonstrates that PTAGF prefers to complete the publishing tasks with trusted members over completing all tasks with untrusted members, to get a high-quality sensor reading.

From Fig. 6c we observe that the running time also increases with increasing the task scale. In all cases, PTAGF takes slightly more time than G-MRUR but is much more efficient than GoCC-G. Moreover, the average running time of PTAGF across all cases is lower than 0.2, when there are 600 participants and 25 sensing tasks, respectively. As discussed above, this is the performance price PTAGF pays for recruiting trustworthy group members.

VIII. CONCLUSION

In this article, we studied the problem of group formulation in MCS, which aimed to recruit highly trusted participants and form a high-reputation group. We proposed a novel group forming framework (PTAGF) that ensures trust and privacy between the group members. In particular, to improve the quality of the senores readings we verified the trust of each participant before they become a group member by considering multiple factors. A theoretical analysis was provided, which proves that the TLPM achieves participants' identities and spatial-temporal privacy preservation and provides strong protection against attacks. Finally, we experimentally evaluated the performance of PTAGF on a real-world dataset against two state-of-the-art approaches. The results show that PTAGF outperforms these approaches in terms of trustworthiness in group selection, and it achieves reasonable task coverage and running time.

In the future, we would like to study the participants' mobility and continuous sensing tasks during task allocation and participant recruitment processes. Furthermore, we plan to consider cooperation between multiple groups to optimize the task coverage percentage. Moreover, developing incentive mechanisms to reward group members according to their contributions, while satisfying payment rationality and budget feasibility is still a challenge to address. We further plan to utilize fog and edge nodes to optimize task allocation and increase the efficiency of our framework.

REFERENCES

[1] N. D. Lane, "Community-aware smartphone sensing systems," *IEEE Internet Comput.*, vol. 16, no. 3, pp. 60–64, May 2012.
[2] R. Azzam, R. Mizouni, H. Otrok, A. Ouali, and S. Singh, "GRS: A group-based recruitment system for mobile crowd sensing," *J. Netw. Comput. Appl.*, vol. 72, pp. 38–50, Sep. 2016.
[3] W. Tan, L. Zhao, B. Li, L. Xu, and Y. Yang, "Multiple cooperative task allocation in group-oriented social mobile crowdsensing," *IEEE Trans. Services Comput.*, vol. 15, no. 6, pp. 3387–3401, Nov. 2022.
[4] L. Zhao, W. Tan, B. Li, L. Xu, and Y. Yang, "Multiple cooperative task assignment on reliability-oriented social crowdsourcing," *IEEE Trans. Services Comput.*, vol. 15, no. 6, pp. 3402–3416, Nov. 2022.
[5] J. Xu, Z. Rao, L. Xu, D. Yang, and T. Li, "Incentive mechanism for multiple cooperative tasks with compatible users in mobile crowd sensing via online communities," *IEEE Trans. Mobile Comput.*, vol. 19, no. 7, pp. 1618–1633, Jul. 2020.
[6] S. Luo, Y. Sun, Z. Wen, and Y. Ji, "c2: Truthful incentive mechanism for multiple cooperative tasks in mobile cloud," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2016, pp. 1–6.
[7] T. Li, Z. Qiu, L. Cao, D. Cheng, W. Wang, X. Shi, and Y. Wang, "Privacy-preserving participant grouping for mobile social sensing over edge clouds," *IEEE Trans. Netw. Sci. Eng.*, vol. 8, no. 2, pp. 865–880, Apr. 2021.
[8] A. Suliman, H. Otrok, R. Mizouni, S. Singh, and A. Ouali, "A greedy-proof incentive-compatible mechanism for group recruitment in mobile crowd sensing," *Future Gener. Comput. Syst.*, vol. 101, pp. 1158–1167, Dec. 2019.
[9] R. Azzam, R. Mizouni, H. Otrok, S. Singh, and A. Ouali, "A stability-based group recruitment system for continuous mobile crowd sensing," *Comput. Commun. J.*, vol. 119, pp. 1–14, Apr. 2018.
[10] M. Abououf, R. Mizouni, S. Singh, H. Otrok, and A. Ouali, "Multi-worker multi-task selection framework in mobile crowd sourcing," *J. Netw. Comput. Appl.*, vol. 130, pp. 52–62, Mar. 2019.
[11] A. Alagha, R. Mizouni, S. Singh, H. Otrok, and A. Ouali, "SDRS: A stable data-based recruitment system in IoT crowdsensing for localization tasks," *J. Netw. Comput. Appl.*, vol. 177, Mar. 2021, Art. no. 102968.
[12] X. Yin, C. Qu, Q. Wang, F. Wu, B. Liu, F. Chen, X. Chen, and D. Fang, "Social connection aware team formation for participatory tasks," *IEEE Access*, vol. 6, pp. 20309–20319, 2018.
[13] J. Jiang, B. An, Y. Jiang, C. Zhang, Z. Bu, and J. Cao, "Group-oriented task allocation for crowdsourcing in social networks," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 51, no. 7, pp. 4417–4432, Jul. 2021.
[14] Z. Zheng, Z. Qin, K. Li, and T. Qiu, "A team-based multitask data acquisition scheme under time constraints in mobile crowd sensing," *Connection Sci.*, vol. 34, no. 1, pp. 1119–1145, Dec. 2022.

[15] T. Li, T. Jung, Z. Qiu, H. Li, L. Cao, and Y. Wang, "Scalable privacy-preserving participant selection for mobile crowdsensing systems: Participant grouping and secure group bidding," *IEEE Trans. Netw. Sci. Eng.*, vol. 7, no. 2, pp. 855–868, Apr. 2020.

[16] Y. Zhang, Z. Ying, and C. L. P. Chen, "Achieving privacy-preserving multitask allocation for mobile crowdsensing," *IEEE Internet Things J.*, vol. 9, no. 18, pp. 16795–16806, Sep. 2022.

[17] R. Estrada, I. Valeriano, and D. Torres, "Multi-task versus consecutive task allocation with tasks clustering for mobile crowd sensing systems," *Proc. Comput. Sci.*, vol. 198, pp. 67–76, Jan. 2022.

[18] Z. Gao, Y. Huang, L. Zheng, H. Lu, B. Wu, and J. Zhang, "Protecting location privacy of users based on trajectory obfuscation in mobile crowdsensing," *IEEE Trans. Ind. Informat.*, vol. 18, no. 9, pp. 6290–6299, Sep. 2022.

[19] H. Cao, S. Liu, L. Wu, and Z. Guan, "SCRAPPOR: An efficient privacy-preserving algorithm base on sparse coding for information-centric IoT," *IEEE Access*, vol. 6, pp. 63143–63154, 2018.

[20] X. He, G. Cormode, A. Machanavajjhala, C. M. Procopiuc, and D. Srivastava, "DPT: Differentially private trajectory synthesis using hierarchical reference systems," *Proc. VLDB Endowment*, vol. 8, no. 11, pp. 1154–1165, Jul. 2015.

[21] H. Jin, L. Su, H. Xiao, and K. Nahrstedt, "INCEPTION: Incentivizing privacy-preserving data aggregation for mobile crowd sensing systems," in *Proc. 17th ACM Int. Symp. Mobile Ad Hoc Netw. Comput.*, Jul. 2016, pp. 341–350.

[22] X. Jin and Y. Zhang, "Privacy-preserving crowdsourced spectrum sensing," *IEEE/ACM Trans. Netw.*, vol. 26, no. 3, pp. 1236–1249, Jun. 2018.

[23] B. Agir, T. G. Papaioannou, R. Narendula, K. Aberer, and J.-P. Hubaux, "User-side adaptive protection of location privacy in participatory sensing," *GeoInformatica*, vol. 18, no. 1, pp. 165–191, Jan. 2014.

[24] R.-H. Hwang, Y.-L. Hsueh, and H.-W. Chung, "A novel time-obfuscated algorithm for trajectory privacy protection," *IEEE Trans. Serv. Comput.*, vol. 7, no. 2, pp. 126–139, Apr./Jun. 2014.

[25] V. A. Memos, K. E. Psannis, Y. Ishibashi, B.-G. Kim, and B. B. Gupta, "An efficient algorithm for media-based surveillance system (EAMSuS) in IoT smart city framework," *Future Gener. Comput. Syst.*, vol. 83, pp. 619–628, Jun. 2018.

[26] H. Li, D. Liao, G. Sun, M. Zhang, D. Xu, and Z. Han, "Two-stage privacy-preserving mechanism for a crowdsensing-based VSN," *IEEE Access*, vol. 6, pp. 40682–40695, 2018.

[27] S. Ni, M. Xie, and Q. Qian, "Clustering based K-anonymity algorithm for privacy preservation," *Int. J. Netw. Secur.*, vol. 19, no. 6, pp. 1062–1071, 2017.

[28] K. Vu, R. Zheng, and J. Gao, "Efficient algorithms for K-anonymous location privacy in participatory sensing," in *Proc. IEEE INFOCOM*, Mar. 2012, pp. 2399–2407.

[29] D. Christin, D. R. Pons-Sorolla, M. Hollick, and S. S. Kanhere, "TrustMeter: A trust assessment scheme for collaborative privacy mechanisms in participatory sensing applications," in *Proc. IEEE 9th Int. Conf. Intell. Sensors, Sensor Netw. Inf. Process. (ISSNIP)*, Apr. 2014, pp. 1–6.

[30] T. Peng, Q. Liu, D. Meng, and G. Wang, "Collaborative trajectory privacy preserving scheme in location-based services," *Inf. Sci.*, vol. 387, pp. 165–179, May 2017.

[31] Y. Ren, W. Liu, T. Wang, X. Li, N. N. Xiong, and A. Liu, "A collaboration platform for effective task and data reporter selection in crowdsourcing network," *IEEE Access*, vol. 7, pp. 19238–19257, 2019.

[32] L. Sweeney, "K-anonymity: A model for protecting privacy," *Int. J. Uncertainty, Fuzziness Knowl.-Based Syst.*, vol. 10, no. 5, pp. 557–570, 2002.

[33] D. Wu, L. Fan, C. Zhang, H. Wang, and R. Wang, "Dynamical credibility assessment of privacy-preserving strategy for opportunistic mobile crowd sensing," *IEEE Access*, vol. 6, pp. 37430–37443, 2018.

[34] P. Zhao, J. Li, F. Zeng, F. Xiao, C. Wang, and H. Jiang, "ILLIA: Enabling k-anonymity-based privacy preserving against location injection attacks in continuous LBS queries," *IEEE Internet Things J.*, vol. 5, no. 2, pp. 1033–1042, Apr. 2018.

[35] S. Zhang, G. Wang, and Q. Liu, "A dual privacy preserving scheme in continuous location-based services," *IEEE Internet Things J.*, vol. 5, no. 5, pp. 4191–4200, Oct. 2018.

[36] L. Xing, X. Jia, J. Gao, and H. Wu, "A location privacy protection algorithm based on double K-anonymity in the social internet of vehicles," *IEEE Commun. Lett.*, vol. 25, no. 10, pp. 3199–3203, Oct. 2021.

[37] Y. Wang, Z. Cai, X. Tong, Y. Gao, and G. Yin, "Truthful incentive mechanism with location privacy-preserving for mobile crowdsourcing systems," *Comput. Netw.*, vol. 135, pp. 32–43, Apr. 2018.

[38] K. L. Huang, S. S. Kanhere, and W. Hu, "Preserving privacy in participatory sensing systems," *Comput. Commun.*, vol. 33, no. 11, pp. 1266–1280, 2010.

[39] M. A. Alsheikh, Y. Jiao, D. Niyato, P. Wang, D. Leong, and Z. Han, "The accuracy-privacy trade-off of mobile crowdsensing," *IEEE Commun. Mag.*, vol. 55, no. 6, pp. 132–139, Jun. 2017.

[40] Y. Cheng, J. Ma, and Z. Liu, "A lightweight privacy-preserving participant selection scheme for mobile crowdsensing," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2022, pp. 1509–1514.

[41] X. Wang, Z. Liu, X. Tian, X. Gan, Y. Guan, and X. Wang, "Incentivizing crowdsensing with location-privacy preserving," *IEEE Trans. Wireless Commun.*, vol. 16, no. 10, pp. 6940–6952, Oct. 2017.

[42] Y. Zhang, M. Li, D. Yang, J. Tang, G. Xue, and J. Xu, "Tradeoff between location quality and privacy in crowdsensing: An optimization perspective," *IEEE Internet Things J.*, vol. 7, no. 4, pp. 3535–3544, Apr. 2020.

[43] W. Shen, B. Yin, Y. Cheng, X. Cao, and Q. Li, "Privacy-preserving mobile crowd sensing for big data applications," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2017, pp. 1–6.

[44] F. Qiu, F. Wu, and G. Chen, "SLICER: A slicing-based K-anonymous privacy preserving scheme for participatory sensing," in *Proc. IEEE 10th Int. Conf. Mobile Ad-Hoc Sensor Syst.*, Oct. 2013, pp. 113–121.

[45] F. Qiu, F. Wu, and G. Chen, "Privacy and quality preserving multimedia data aggregation for participatory sensing systems," *IEEE Trans. Mobile Comput.*, vol. 14, no. 6, pp. 1287–1300, Jun. 2015.

[46] D. Reinhardt and I. Manyugin, "OP4: An OPPortunistic privacy-preserving scheme for crowdsensing applications," in *Proc. IEEE 41st Conf. Local Comput. Netw. (LCN)*, Nov. 2016, pp. 460–468.

[47] D. Tang and J. Ren, "A novel delay-aware and privacy-preserving data-forwarding scheme for urban sensing network," *IEEE Trans. Veh. Technol.*, vol. 65, no. 4, pp. 2578–2588, Apr. 2016.

[48] D. Christin, J. Guillemet, A. Reinhardt, M. Hollick, and S. S. Kanhere, "Privacy-preserving collaborative path hiding for participatory sensing applications," in *Proc. IEEE 8th Int. Conf. Mobile Ad-Hoc Sensor Syst.*, Oct. 2011, pp. 341–350.

[49] D. Christin, A. Reinhardt, and M. Hollick, "On the efficiency of privacy-preserving path hiding for mobile sensing applications," in *Proc. 38th Annu. IEEE Conf. Local Comput. Netw.*, Oct. 2013, pp. 818–826.

[50] I. Boutsis and V. Kalogeraki, "Privacy preservation for participatory sensing data," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun. (PerCom)*, Mar. 2013, pp. 103–113.

[51] C. Bassem, "On cooperative obfuscation for privacy-preserving task recommendation in mobile CrowdSensing," in *Proc. 17th Int. Conf. Wireless Mobile Comput., Netw. Commun. (WiMob)*, Oct. 2021, pp. 90–95.

[52] C. Buttner and S. A. Huss, "Path hiding for privacy enhancement in vehicular ad-hoc networks," in *Proc. IEEE 82nd Veh. Technol. Conf. (VTC-Fall)*, Sep. 2015, pp. 1–5.

[53] Z. Xu, W. Yang, and J. Wang, "MSPP: A trajectory privacy-preserving framework for participatory sensing based on multi-strategy," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2019, pp. 1–6.

[54] B. H. Alamri, M. M. Monowar, and S. Alshehri, "A privacy-preserving collaborative reputation system for mobile crowdsensing," *Int. J. Distrib. Sensor Netw.*, vol. 14, no. 9, Sep. 2018, Art. no. 155014771880218.

[55] X. Wang, K. Govindan, and P. Mohapatra, "Collusion-resilient quality of information evaluation based on information provenance," in *Proc. 8th Annu. IEEE Commun. Soc. Conf. Sensor, Mesh Ad Hoc Commun. Netw.*, Jun. 2011, pp. 395–403.

[56] W. A. Mason, F. R. Conrey, and E. R. Smith, "Situating social influence processes: Dynamic, multidirectional flows of influence within social networks," *Personality Social Psychol. Rev.*, vol. 11, no. 3, pp. 279–300, Aug. 2007.

[57] V. 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.

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

[59] 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.

[60] E. K. Asl, J. Bentahar, H. Otrok, and R. Mizouni, "Efficient community formation for web services," *IEEE Trans. Services Comput.*, vol. 8, no. 4, pp. 586–600, Jul. 2015.

[61] L. Mashayekhy, M. M. Nejad, and D. Grosu, "A trust-aware mechanism for cloud federation formation," *IEEE Trans. Cloud Comput.*, vol. 9, no. 4, pp. 1278–1292, Oct. 2021.
[62] E. Cho, S. A. Myers, and J. Leskovec, "Friendship and mobility: User movement in location-based social networks," in *Proc. 17th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining (KDD)*, 2011, pp. 1082–1090.
[63] J. Wang, F. Wang, Y. Wang, D. Zhang, L. Wang, and Z. Qiu, "Social-network-assisted worker recruitment in mobile crowd sensing," *IEEE Trans. Mobile Comput.*, vol. 18, no. 7, pp. 1661–1673, Jul. 2019.
[64] J. Jiang, B. An, Y. Jiang, C. Zhang, Z. Bu, and J. Cao, "Group-oriented task allocation for crowdsourcing in social networks," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 51, no. 7, pp. 4417–4432, Jul. 2021.
[65] M. Kargar, A. An, and M. Zihayat, "Efficient bi-objective team formation in social networks," in *Proc. Joint Eur. Conf. Mach. Learn. Knowl. Discovery Databases*. Cham, Switzerland: Springer, 2012, pp. 483–498.

MUHAMMAD MOSTAFA MONOWAR received the B.Sc. degree in computer science and information technology from the Islamic University of Technology (IUT), Bangladesh, in 2003, and the Ph.D. degree in computer engineering from Kyung Hee University, South Korea, in 2011. He worked as a Faculty Member with the Department of Computer Science and Engineering, University of Chittagong, Bangladesh. He is currently working as an Associate Professor with the Department of Information Technology, King Abdulaziz University, Saudi Arabia. His research interests include wireless networks, mostly ad-hoc, sensor, and mesh networks, including routing protocols, MAC mechanisms, IP and transport layer issues, cross-layer design, QoS provisioning, security and privacy issues, and natural language processing. He has served as a program committee member for several international conferences/workshops. He served as an editor for a couple of books published by CRC Press and the Taylor & Francis Group. He also served as a guest editor for several journals.

BAYAN HASHR SAEED ALAMRI received the B.Sc. degree from Taibah University, Medina, Saudi Arabia, in 2013, and the M.Sc. degree in information technology from King Abdulaziz University, in 2018, where she is currently pursuing the Ph.D. degree. Her main research interests include networks and wireless networks, network security, cloud computing security, wireless sensor networks protocols, the Internet of Things, and mobile crowdsensing.

SUHAIR ALSHEHRI received the Ph.D. degree in computing and information sciences from the Golisano College of Computing and Information Sciences, Rochester Institute of Technology, in 2014. She is currently an Assistant Professor with the Department of Information Technology, Faculty of Computing and Information Technology, King Abdulaziz University. Her main research interests include security and privacy in computer and information systems, and applied cryptography.