IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 15, No. 1, February 2026, pp. 383~398
ISSN: 2252-8938, DOI: 10.11591/ijai.v15.i1.pp383-398

Centrality-optimized coalition formation: a genetic algorithm approach with leadership attributes

Anon Sukstrienwong, Sorapak Pukdesree
School of Information Technology and Innovation, Bangkok University, Pathum Thani, Thailand

Article Info

Article history:

Received Oct 30, 2024
Revised Nov 3, 2025
Accepted Jan 10, 2026

Keywords:

Closeness centrality
Genetic algorithm
Group formation
Leadership attributes
Optimization algorithm

ABSTRACT

In graph theory, centrality is often assessed using traditional methods such as closeness centrality, which measures the average shortest path length between nodes in a network. In this study, we primarily focus on developing the proposed approach and demonstrating its effectiveness through initial experimental results. A novel genetic algorithm (GA)–based method named centrality–optimized leadership coalition formation (COLCF) has been designed. It emphasizes actual agent distances according to closeness centrality and leadership attributes in group formation. We detail the COLCF algorithm, present empirical case studies, and provide efficiency comparisons. In accordance with our simulation results, the proposed algorithm is capable of capitalizing on the ideal coalition structure for achieving high closeness centrality when incorporated with leadership attributes. The experimental results demonstrate the algorithm's robustness and effectiveness in addressing complex coalition formation challenges.

Corresponding Author:

Anon Sukstrienwong
School of Information Technology and Innovation, Bangkok University
9/1 Moo 5, Phahonyothin Road, Khlong Nueng, Khlong Luang District, Pathum Thani, Thailand
Email: anon.su@bu.ac.th

1. INTRODUCTION

The concept of "coalition formation" has been modeled as an essential tool through mathematical theories of coalition behavior, allowing groups to obtain the greatest benefit. Accordingly, it has been applied across diverse contexts and domains. In particular, it is employed in altruistic decision-making in multi-agent systems [1], shared resource management in congestion problems where multiple agents compete for limited resources [2], public allocation guided by information design [3], and capability-based modeling for complex group tasks [4]. In the real-world applications, the term "coalition formation" is alternatively referred to as team formation, strategic alliances, joint venture formation, group formation, partnership formation, and collaborative grouping [5]–[8]. Among these, "team formation" is a term often used in the context of games. It refers to optimizing group outcomes through cooperative efforts based on the attributes of the game's participants [9], [10]. In the business context, it is often referred to as strategic alliances, where alliance companies collaborate to develop innovation and enhance market performance by sharing strengths and mutual benefits. It can also be called "joint venture formation" when referring to a formal business arrangement [11]. Similarly, the term "partnership formation" is used interchangeably with "group formation," but it most commonly describes official agreements between two or more parties for long-term collaboration [12]. The term "group formation" is commonly used in education as well, particularly in collaborative learning and computer-supported collaborative learning (CSCL) environments. For example, some technology companies join together to share their skills and ideas to create new innovations. Therefore, coalition schemes have become essential tools, offering promising solutions to deal with complex challenges, especially in education, where collaboration is employed to achieve optimal learning objectives. This is because enhancing the performance of diverse students through groupwork lies at the core of educational development. Chai *et al.* [13] conducted a systematic review of empirical research focusing on computer-based assessment of collaborative problem-solving (CPS) skills. The authors pointed out that collaborative methods should prioritize real-world skills and effectively capture students' collaborative behaviors and outcomes. Within engineering, the term "coalition formation" describes how autonomous agents self-organize into task-oriented groups to improve efficiency and responsiveness in complex environments [14], [15]. The papers emphasize how task completion is optimized by grouping members based on skills and contextual information.

Consequently, several researchers have investigated various methodologies, rules, policies, domain knowledge, and guidelines for selecting appropriate algorithms for coalition formation to achieve their respective objectives. For example, Bausch [16] proposes rules in the context of politics regarding how leaders are re-selected and how they form the best coalitions of different sizes within large groups. These selected leaders are intended to facilitate effective coordination and enhance overall group cohesion. Ding *et al.* [17] employ coalition formation to optimize a distributed network in a jamming environment on multi-agent behavior. The authors present distributed algorithms that enable users to make decisions resulting in efficient network performance. Rossi *et al.* [18] review several algorithmic strategies for coalition formation and classify multi-agent algorithms for collective behavior, highlighting their role in distributed decision-making. Several studies focus on online coalition formation from theoretical and algorithmic perspectives, as detailed by Bullinger and Romen [19]. This line of study, including a broader overview, is generally supported by the survey on coalition formation in multi-agent systems [6], [20], [21]. Incontestably, achieving their common objectives through joining proper members in tandem and having the right individuals in the groups makes strong collaborative connections [22]. Forming groups of individuals in a virtual environment has been validated as tool for generating positive outcomes by maximizing opportunities for sharing each member's experience and satisfying adequate learning performance [9], [23]–[25].

It is widely acknowledged that, up to the present time, there are various approaches to establishing group formation, including random grouping, self-selection, and criterion-based grouping. Ramírez *et al*. [26] examined these approaches in secondary education, highlighting their characteristics and implications. The random technique is an easy and straightforward method since everyone is assigned to groups randomly. It allows all members to be mixed to achieve heterogeneity in the separated groupings [27]–[29]. This goal, however, is not always satisfied, as randomly assigning students to groups can lead to issues, with some members being reluctant to join the group to which they have been allocated. The second method, self-selected group formation, allows people to establish groups independently. This approach enables everyone to form groups based on their own decisions, preferences, or familiarity with others [27], [29]–[31]. The third method, criterion-based grouping, permits the formation of groups based on specified criteria and algorithms deemed scientifically accurate [29], [31], [32]. Consequently, various grouping criteria have been proposed to create well-structured groups that foster collaboration, enhance task efficiency, and improve overall performance. From another perspective, a multi-target optimization technique was proposed by Miranda *et al*. [33] to facilitate the group formation problem, while accounting for numerous objective functions such as inter-homogeneity, intra-heterogeneity, and empathy. Their experiments were conducted on varying sizes and complexities, aiming to maximize the average similarity of group members. Following that, Boongasame *et al.* [34] presented a fascinating work involving group formation with the characteristic of a coalition leader. The authors focused on a buyer coalition scheme to gain additional discounts for their group purchasers.

However, we are interested by the fact that a group of individuals distributed in different geological locations must form a group based on the results of cost and reward, where each member is placed adjacent to each other. Forming groups of members based on their geographical location, of course, demands a group centrality measurement in order to identify significant individuals who possess leadership attributes and a high score in closeness centrality among a large number of participants. This is because group members can be directly influenced by the group's leader. Therefore, each person's leadership attribute is specified in order to assess closeness centrality. A member can be defined as a node in graph theory. All nodes may have a leader at their center, as a qualified leader can coordinate group activities and comfortably communicate to all members. In addition, group leadership is one of the most essential characteristics of team efficiency. Some authors in [35], [36] have affirmed that leadership correlates to better team performance. Fransen *et al* [35] verified that it is beneficial to designate a leader in a group to empower team members since leadership abilities are more broadly dispersed over time among team members. The relevance of closeness in the network graph [37], is used to determine a node's power, activities, and communication convenience. Moreover, the authors claimed that group activities are always associated with good leadership, strong popularity, or the reputation of a network. Meshcheryakova and Shvydun [38] also emphasized the importance of centrality metrics in evaluating individual roles within a network. The authors highlighted that closeness, betweenness, and eigenvector centrality each offer unique insights into an individual's influence and communication efficiency. Complementing this, Istrate *et al.* [39] proposed a game-theoretic framework for analyzing centrality by modeling individuals and their connections as coalitions, thereby enhancing the measurement of influence and effectiveness within group interactions.

The research in [40], [41] emphasized the importance of closeness centrality in leadership. Mitterlechner highlighted how leaders positioned closer to others in the network can better coordinate activities and influence group dynamics, while Yuan and Knippenberg [41] discussed how a leader's central position in a network positively impacts team performance, with team size acting as a moderating factor. The papers provide us with a perspective on how centrality underscores the importance of incorporating leadership attributes when evaluating network structures. Therefore, including leadership attributes is critical in this work. Consequently, the distance between members and their group leader is, thus, the primary factor in our work for forming optimal groups so that closeness centrality can be achieved. However, creating an optimal group formation based on the closeness centrality of the leader is a demanding problem that requires a sophisticated algorithm. Therefore, the development of efficient algorithms to tackle this problem is crucial in various domains.

This paper is organized into five sections, including this introductory section. Section 2 includes background theory as well as an overview of literature related to our work. Section 3 describes our proposed genetic-like algorithm for generating optimal groups of agents in terms of closeness centrality to a qualified leader. A closeness centrality-based measure that we employed in the fitness function is also detailed. Following that, section 4 presents an empirical case study to validate the proposed algorithm. The conclusion and discussion are provided in section 5.

2. BACKGROUND AND RELATED WORKS

This section provides a brief background of group formation regarding a coalition leader as well as closeness centrality in the network to comprehend the coalition development of our proposed algorithm.

2.1. Group formation

Group formation concepts involve the challenge of grouping members effectively to ensure successful collaboration toward a common goal [29], [31], [42], [43]. Successful group formation depends not only on individual skills but also on the strategic positioning of members within the social network to maximize interaction and influence [44], [45]. Currently, numerous studies endeavor to establish optimized groups with a variety of aims and techniques. For instance, Sethi and Nicholson [46] investigated hypothesis testing using structural equation modeling in order to attain exceptional performance. The authors provided findings regarding team structural characteristics and contextual factors that correlate to antecedents of charged behavior in charged product development teams. Isotani *et al.* [47] focused on technology development that enhances collaboration and class communication. The authors stated that group formation is critical to the acceptability of group activities and the effectiveness of the learning process. In addition, they claimed that several collaborative learning approaches have failed due to insufficient group formation. Leem and Chen [48] proposed a similarity-coefficient-based strategy for machine grouping in order to generate efficient machine cells and part families that optimize similarity values. Guo *et al.* [49] proposed an approach for forming logical groupings. The authors developed a mobile application named GroupMe that enables individuals to organize social activities. Tseng *et al.* [50] suggested a unique method for multi-functional project team formation where there are no clear tasks between customer requirements and project characteristics. Fuzzy set theory is utilized in their work to deal with ambiguous situations, while grey decision theory (uncertainty or incomplete information) is employed to choose desired team members. Indrawan *et al.* [51] provided a framework to facilitate multi-attribute coalition negotiation in the e-marketplace. Yu *et al.* [52] developed a multi-attribute coalition formation-based negotiating protocol aimed at minimizing the workload and time consumption for manufacturer agents. In a related context, Ponce *et al.* [53] provided a comprehensive analysis of how cross-sector organizational structures support effective group formation and coalitions for sustainability outcomes, which can also be adapted in sustainable business models aligned with the SDGs.

When forming an effective group often begins by identifying a capable leader, we require a qualified leader to support the group's development based on the group's objectives [54]. In common understanding, a leader refers to a person who can encourage members to work cooperatively towards their goals, which is supported by some papers; for example in [55], [56] provided empirical evidence showing that coaching and entrepreneurial leadership styles foster collective efficacy within teams. As shown in Pérez *et al.* [57], this is further supported by analyses of esports teams, where key structures and processes emphasize the critical role of leadership in promoting cooperation and team effectiveness. Generally, there are countless opportunities for everyone to develop leadership abilities in preparation for becoming a leader. Some individuals may possess specific attributes that make them more qualified to be the leader of the group. Certain leadership behaviors and attributes, such as supporting others' learning, may contribute to an individual's suitability for a leadership role [58], [59]. As a result, several papers focus on the leader characteristics that influence group formation. Boongasame *et al.* [34] conducted an intriguing study that examined group formation in connection to a coalition leader. The authors aimed to form a buyer coalition utilizing a scenario and related simulation tools in which leadership characteristics are examined. A recent work by Xie *et al.* [60] introduced a coalition formation framework for dynamic task allocation, showing that agent-level leadership behaviors support more effective coalition formation.

In addition, Razin and Piccione [61] measured coalition formation under power relations by determining the characterization of strongly stable social orders using the partition function for a cooperative game associated with individual and group power. Breban and Vassileva [62] presented an inter-agent trust relationship coalition-building method. A coalition-building process was presented and assessed by the authors, taking into account the trust relationships between agents. Nardin and Sichman [63] assessed the trust degree of their leader and investigated how trust influences coalition formation. Tarkowski *et al.* [64] conducted the survey in order to analyze a variety of game-theoretic centrality metrics, detailing how agents can be evaluated and selected based on their coalition-forming potential. Molinero *et al.* [65] further developed this aspect by using influence-game models. The authors also employed the centrality of members to coalition strength and stability.

2.2. Closeness centrality

In a network graph, the closeness centrality of a node indicates how close it is to all other nodes. It is defined as the average value of the shortest path lengths between each node in the network [42], [66]–[70]. It has the advantage of identifying prominent nodes if they are more central and closer to the majority of nodes in the network. As a result, it often determines which node has the most influence among others in a given network [66]. For this present study, closeness centrality makes perfect sense in terms of influence. This is because the most influential person refers to a person who can effortlessly reach out to others. Therefore, a person located in the center of the group may have a high relationship with other persons. Let $G = (V, E)$ be a network modeled as a simple graph with $n = |V|$ vertices, where V is the set of nodes and E is the set of edges. By finding the shortest routes between all pairs of nodes in the graph, the closeness centrality algorithm computes the farness of a vertex, which is defined as the sum of each node's shortest path lengths to all other nodes. The farness of a vertex $x \epsilon V$ is defined as (1).

$$far(x) = \sum_{\substack{v \in V \\ d(v,x) \neq \infty}} d(v, x) \tag{1}$$

Where $d(v, x)$ represents the shortest distance between the nodes x and v.

The farness is then reversed to calculate the closeness centrality score of the node. The equation presented in (2) defines closeness of a node.

$$C(x) = \frac{1}{far(x)} \tag{2}$$

However, $C(x) = 0$ if x cannot reach any vertex in the graph. We can observe that, if the farness is large, the closeness centrality becomes small and vice versa.

3. DETAILED CALCULATION METHODS AND PROPOSED GENETIC ALGORITHM

The proposed centrality–optimized leadership coalition formation (COLCF) algorithm leverages this measure to effectively identify optimal group structures while simultaneously selecting qualified leaders for each group, enhancing overall group performance and coordination. The mathematical model and genetic algorithm (GA), used in this study, including the calculation example of algorithm, are detailed in as follows.

3.1. Mathematical model

Let $G = \{a_1, a_2, a_{3,}\ldots, a_n\}$ denote a set of agents, where n is the total number of agents. In this work, the term "agent" refers to "a person". However, in other domains, it might refer to an individual in a factory or a robot that can join together in a group with a shared goal. Each agent is comprised of a binary-value leader attribute. Also, an agent a_i for $1 \leq i \leq n$, contains attributes represented in the form of $v_i = (longitude, latitude_i, att_{i1}, att_{i2}, \ldots, att_{ik})$, where k is the number of attributes.

Consider a non-empty set G_j, where $1 \leq j \leq p$ representing a divided subgroup of G for a certain coalition formation. Hence, $\bigcup_{j=1}^{p} G_j = G$, where p is the total number of subgroups. Any agent is not permitted to belong to more than one group, which is $\bigcap_{j=1}^{p} G_i = \emptyset$. Typically, to partition a set of n distinct agents into p groups as equal size as possible, the total number of different grouping is calculated by considering all permutation, $\frac{n!}{((q+1)!)^r \cdot (q!)^{p-r} \cdot r! \cdot (p-r)!}$, where $q = \left\lfloor \frac{n}{p} \right\rfloor$ represents the size of the smallest group, $r = n \, mod \, p$ represents the number of groups that have one extra agent. Therefore, there are r groups of $q+1$ and $p-r$ groups of q. For example, if $n = 10$, $p = 3$, then we get $q = \left\lfloor \frac{10}{3} \right\rfloor = 3$, and $r = 10 \, mod \, 3 = 1$. This means that one group will have 4 members. And, there will be two groups containing 3 members. Hence, the number of different ways to separate 3 groups of 10 agents is $\frac{10!}{((3+1)!)^1 \cdot (3!)^{3-1} \cdot 1! \cdot (3-1)!} = \frac{10!}{4! \cdot (3!)^2 \cdot 1! \cdot 2!} = \frac{3628800}{10368} = 350$. This means there are 350 distinct ways to separate 10 agents into 3 groups of nearly equal size. As we can see, an increase in n results in a combinatorial growth in the number of distinct group formations, resulting in exponential complexity. In our specific problem, group formation based on a leader, each agent is represented by a vector of the form ((latitude, longitude), Leader_attribute). The Leader_attribute is a binary variable that indicates whether the agent can serve as the leader of the group. In addition, each divided group must contain a qualified agent leader, who has a leader attribute. Hence, we use the term "complete" to refer to the divided group G_j if it contains at least one agent a_i with the leader attribute $att_i = 1$. Otherwise, the group is labeled with the term "fail" meaning that it has no qualified leader or no one in this group contains leader attribute $att_i = 0$. Let L be a collection of all qualified leaders, where $L \subset G$. The formation of a group will be successful if and only if the number of competent leaders is at least p, which equals |L|. In other words, if $|L| \geq p$. group formation fails.

The problem of group formation based on the closeness centrality of the group leader and attributes can be a challenging and time-consuming task, especially when an optimal solution is required. As the number of agents and subgroups grows bigger, the search space of the problem becomes larger. Hence, this problem becomes more complex, having the difficulty of finding the optimal solution. Consequently, it has become a more challenging task for us to form the optimal groups that obtain qualified leaders within the divided smaller group, where the mean distance between group members and a group leader is as low as possible.

In our work, a group formation is successful if and only if it satisfies the following conditions:

i) Each established group contains a qualified leader.
ii) The number of qualified leaders is at least p, i.e., $|L| \geq p$.
iii) All divided groups are balanced in terms of the distance between their members and their group leaders. This ensures that the split groups are more central, meaning that all agents are near their agent leader.
iv) Finally, we assume that all agents can communicate directly with each other, without any limitations.

3.2. Genetic algorithm design

GA, a well-known optimization technique, are employed for establishing optimal groupings of agents that fulfill the specified criteria and task requirements for effective coalition formation [22], [71]–[73]. The evolution of a GA begins with a fully random initial set of potential solutions, which are referred to as the population. The role of the initial population with different initialization techniques can have a great impact on the efficiency and effectiveness of GAs [74], [75]. GAs leverage evolutionary principles by using selection, crossover, and mutation to iteratively improve candidate solutions over successive generations. Hence, our GA is designed using only two processes: selection and mutation, as shown in Figure 1.

As the purpose of algorithm is to have all agents located as close as possible to their leader in group they belong to, it is therefore necessary to calculate the value of $C(x)$ for all divided groups. Keep in mind that an agent leader (a_i) of a group contains a leader attribute $att_i = 1$. The initial population is randomly selected, and its fitness value will be calculated. In the selection process, the algorithm selects candidates based on fitness value, and only a certain percentage of individuals with the best fitness value are picked. By doing this, more fit individuals are stochastically selected from current population to be in next generation. The next operator is the mutation process, which is used to introduce diversity into the population. Permutation mutation is adopted in this design to ensure that all agents are selected. The new generation of candidate solutions is then used in next generation. By evaluating the fitness function of each chromosome and iteratively evolving the population through selection and mutation, the algorithm can converge on a set of chromosomes that represent good solutions to problem of group formation. This capability, particularly in GA-like algorithms, has been supported across various works in the literature [71]–[80]. The algorithm terminates when either a maximum number of generations (Max_Gen <Gen?) has been produced or a satisfactory fitness level has been reached for the population. Importantly, the chromosome representation

used in this paper is designed as a fixed-size array, as presented in Figure 2, containing all agents. The value of T_j for $1 \leq j \leq p$ indicates that the agent located in this gene is the leader of group j.

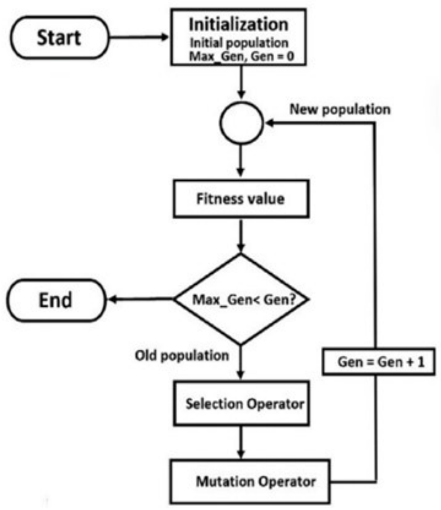

Figure 1. Flowchart for COLCF

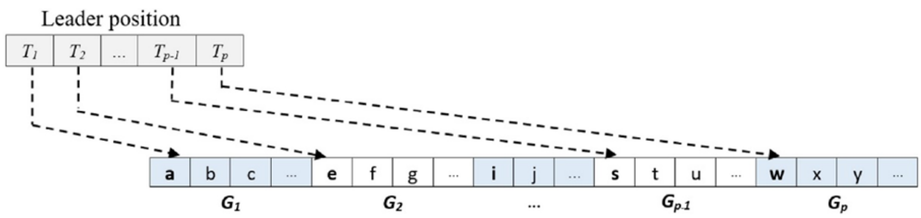

Figure 2. Chromosome representation associated with the array of leader positions, where T_j for $1 \leq j \leq p$

The algorithm initializes with Gen=0. And it runs repeatedly until the generation hits the Max_Gen. For this investigation, the fitness is associated with the closeness centrality as presented in (1) and (2). To calculate the distance between two geographical locations in kilometers, we then use the haversine formula given by (3).

$$d = 2 \times 6371 \times arcsin\left(\sqrt{sin^2\left(\frac{\Delta\phi}{2}\right) + \cos(\phi_1) \times \cos(\phi_2) \times sin^2\left(\frac{\Delta\lambda}{2}\right)}\right) \quad (3)$$

Where d is distance between two points in kilometers, $\Delta\phi = \phi_2 - \phi_1$ is difference in latitude, $\Delta\lambda = \lambda_2 - \lambda_1$ is difference in longitude, ϕ_1 and ϕ_2 are the latitudes of the two points in radians, λ_1 and λ_2 are the longitudes of the two points in radians.

Example: let two geographical points in (latitude, longitude) format be defined as follows. The distance between them is calculated as follows:

– Point 1 (13.658592620562027,100.6486061283200).
– Point 2 (13.61961628801031,100.73134604937815).

First, convert the latitude and longitude values from degrees to radians using the conversion factor $(\pi/180)$. Each point in radian is then represented in the form (ϕ, λ). Values rounded to 5 decimal places for readability.

Point 1 in radians: $(\phi_1, \lambda_1) = (0.23839, 1.75665)$
Point 2 in radians: $(\phi_2, \lambda_2) = (0.23771, 1.75809)$
$\Delta\phi = \phi_2 - \phi_1 \approx 0.23771 - 0.23839 = -0.00068$ radians
$\Delta\lambda = \lambda_2 - \lambda_1 \approx 1.75809 - 1.75665 = 0.00144$ radians

$$d = 2 \times 6371 \times arcsin\left(\sqrt{sin^2\left(\frac{-0.00068}{2}\right) + cos(0.23839) \times cos(0.23771) \times sin^2\left(\frac{0.00144}{2}\right)}\right)$$
$$= 2 \times 6371 \times arcsin\left(\sqrt{0.0000001156 + 0.97197 \times 0.97203 \times 0.0000005184}\right)$$

$= 2 \times 6371 \times \arcsin(\sqrt{0.0000006057})$
$= 2 \times 6371 \times \arcsin(0.0007781)$
$\approx 2 \times 6371 \times 0.0007781$
$\approx 9.91421\ (km)$

Therefore, the distance between point 1 and point 2, written as $d(\text{Point 1}, \text{Point 2})$, is approximately 9.91421 km.

The COLCF algorithm calculates the closeness centrality between a leader and the other members within the divided group. Subsequently, the chromosome that yields a higher score of closeness centrality will be considered the best candidate for forming optimal groupings of agents. It should be noted that in (4) and (5), the variable x is assigned to the leader agent of the divided group. As can be seen, $far(x)$ is the sum of the distances between the agent leader and the others in the group. The value of $far(x)$ can be large. Therefore, $C(x)$ is defined as 1 divided by $far(x)$, resulting in a value is less than 1. The fitness value of the chromosome is determined by taking the mean value of all closeness centralities of the divided groups, which is also less than 1. The closeness centrality between the agent leader and the other agents is normalized by dividing by the maximum possible distance (n–1), where n is the number of agents in the divided group. This means that the maximum possible value for the closeness centrality of each group is 1.

$$fitness(chromsom) = \frac{1}{p}\left(\sum_{V=G_1}^{G_p} C(x)\right) \quad (4)$$

Then, substituting (1) and (2) into (4), we obtain:

$$fitness(chromsom) = \frac{1}{p}\left(\sum_{V=G_1}^{G_p}\left(\frac{1}{\sum_{\substack{v\in V\\ d(v,x)\neq\infty}} d(v,x)}\right)\right) \quad (5)$$

where x is an agent leader of the divided group and p is the total number of subgroups.

3.4. Step-by-step example and calculation of the algorithm

To illustrate the operation of the proposed algorithm, consider the following example. Suppose we divide three equal groups (p=3) from nine agents in G={a, b, c, d, e, f, g, h, i}, where only *a, b, c,* and d possess the leader attribute, and they are scattered across the map. The location of each agent is presented, as illustrated in Figure 3. We assume that each agent's location is represented in the (x, y) coordinates, and their vector representations are in the form of (Leader_attribute, (x, y)). Therefore, it is straightforward to calculate the distance between any two agents. For better understanding and simplicity, we then employ the Euclidean distance formula instead of the Haversine formula to calculate distances between two members on the 2D map. Hence, the vectors of the agents are defined as follows: a=(1, (2,5)), b=(1, (1,1)), c=(1, (5,4)), d=(1, (2,2)), e=(0, (3,4)), f=(0, (6,3)), g=(0, (3,3)), h=(0, (4,2)), and i=(0, (4,5)).

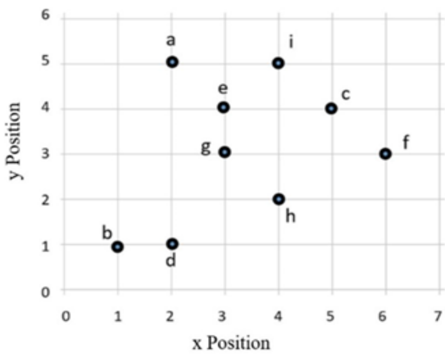

Figure 3. Location of each agent on the 2D map

To form three groups of equal size, our genetic-based algorithm explores different possible configurations while satisfying the given constraints. The algorithm initializes the population with a set of randomly generated chromosomes, each of which represents a possible grouping of the agents. By using (3), the fitness function of the chromosome is evaluated based on the closeness centrality of each group's leader, reflecting their importance in the overall network. The chromosome shown in Figure 4 encodes the chromosome that G_1={b, d, g}, G_2={a, e, i}, and G_3={c, f, h}. Additionally, the leader of G_1 is b, the leader of G_2 is a, and the leader of G_3 is c, and the network can be displayed in Figure 5.

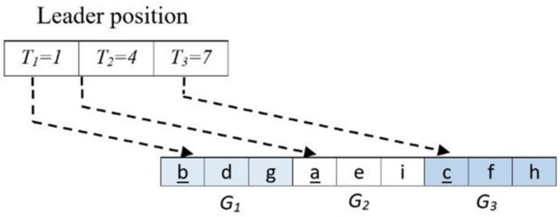

Figure 4. The chromosome representing G_1={b, d, g}, G_2={a, e, i}, and G_3={c, f, h}, where b, a, and c are the leaders of G_1, G_2, and G_3, respectively

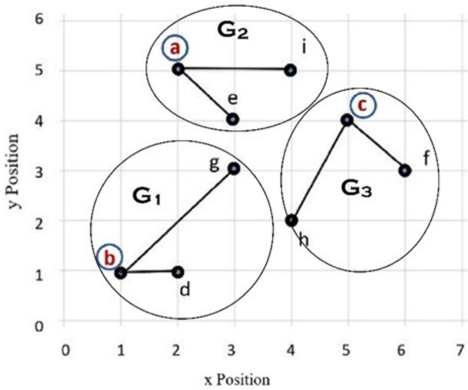

Figure 5. All groupings decoded from chromosome in Figure 4 on the 2D map, showing each group's centrality

For each of the three groups, one representative agent, who is the leader of the group, will be used to compute the shortest path distances to the other nodes within the same group. The sums of these distances are then inverted, and the average of the three results is taken as the fitness value. By using (7), the fitness is calculated as $\frac{1}{3}\left(\frac{1}{1+\sqrt{8}}+\frac{1}{\sqrt{2}+2}+\frac{1}{\sqrt{2}+\sqrt{5}}\right)$, which evaluates to approximately 0.27602. We can observe that the offspring yields a higher fitness value, indicating that the centrality of divided groups associated with the group leader is better than that of its parent. Generally, the offspring with a higher fitness value has a greater chance of being selected for the next generation compared to their parents.

$$
\begin{aligned}
Fitness(parent\ chromosome) &= \frac{1}{3}\left(\sum_{V=G_1}^{G_3}\left(\frac{1}{\sum_{\substack{v\in V \\ d(v,x)\neq\infty}} d(v,x)}\right)\right) \\
&= \frac{1}{3}\left(\frac{1}{d(b,d)+d(b,g)}+\frac{1}{d(a,e)+d(a,i)}+\frac{1}{d(c,f)+d(c,h)}\right) \\
&= \frac{1}{3}\left(\frac{1}{1+\sqrt{8}}+\frac{1}{\sqrt{2}+2}+\frac{1}{\sqrt{2}+\sqrt{5}}\right) \\
&\approx 0.27602
\end{aligned}
$$

We then apply the permutation mutation to generate an offspring. In general, the offspring is similar to its parent except for the mutated gene. Suppose that the permutation mutation modifies only the first group. An example of offspring resulting from the mutation is shown in Figure 6, where G_1 consists of {d, b, g}, while G_2 and G_3 remain unchanged. Then, we can demonstrate all group members connected to their leader, as presented in Figure 7. It is important to note that, for a group to be considered qualified, the leader must possess the leadership attribute. For example, if the group was represented as G_1={g, b, d}, it would be unqualified because g's leadership attribute is equal to 0. To ensure the quality of the offspring generated by the algorithm, the chromosomes containing no unqualified subgroups like this are eliminated from the candidate chromosomes. As a result, only the offspring with the qualified groups based on this are considered for further evolution.

As mentioned previously, the COLCF algorithm considers the chromosome that produces the greatest closeness centrality score between an agent leader and the others within the divided groups. Therefore, the chromosome that produces the greatest score for closeness centrality is deemed the most suitable option for creating ideal agent groups. This can be verified through the following calculation. Let us examine the map illustrated in Figure 7. We can see that d in the group of {d, b, g} has the lowest total distance to all others. The fitness is calculated as $\frac{1}{3}\left(\frac{1}{1+\sqrt{5}}+\frac{1}{\sqrt{2}+2}+\frac{1}{\sqrt{2}+\sqrt{5}}\right)$, which evaluates to approximately 0.29195. As we can see, d can be a leader of the group as its closeness centrality is the highest. Details of the chromosome's fitness value calculation are presented as follows.

$$\begin{aligned} Fitness(offspring) &= \frac{1}{3}\left(\frac{1}{d(d,b)+d(d,g)}+\frac{1}{d(a,e)+d(a,i)}+\frac{1}{d(c,f)+d(c,h)}\right) \\ &= \frac{1}{3}\left(\frac{1}{1+\sqrt{5}}+\frac{1}{\sqrt{2}+2}+\frac{1}{\sqrt{2}+\sqrt{5}}\right) \\ &\approx 0.29195 \end{aligned}$$

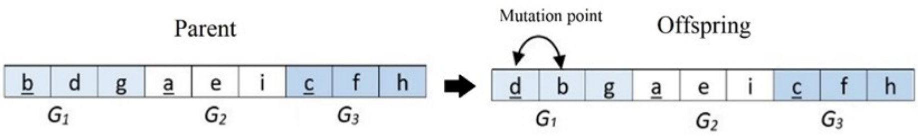

Figure 6. Permutation mutation

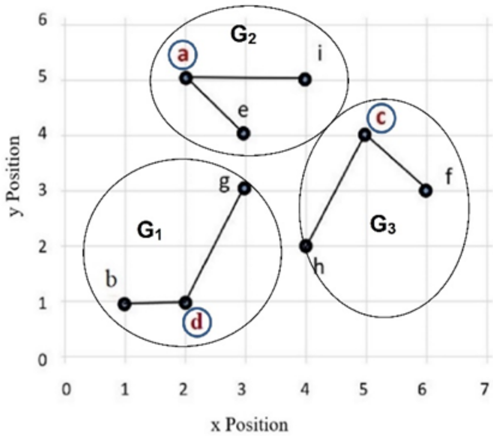

Figure 7. Representation of each group on the 2D map based on the chromosome shown in Figure 6

The fitness value of the offspring is approximately 0.29195, which is higher than that of its parent (0.27602). This improvement showed that the offspring had a better grouping solution compared to its parent. During each generation, the fitness function of each chromosome is evaluated, and the population is iteratively evolving through selection and mutation. The algorithm continued to evolve the population until it reached the criterion; maximum number of generations. At this point, the COLCF algorithm had found a set of chromosomes that represented good solutions to the problem of group formation, and the optimal solution could be found. Its ability to explore and exploit the search space, while introducing diversity through mutation, made it an effective tool for finding good solutions to the problem.

4. RESULTS AND DISCUSSION

To demonstrate the scalability of the proposed algorithm, we conducted a series of experiments guided by insights and methodologies reported in recent comprehensive surveys and studies [81]–[87]. Our empirical experiments varied the number of agents, groups, initial population size, and maximum generations to evaluate the algorithm's performance under different scalable conditions. The detailed experimental setup is summarized in Table 1.

Table 1. Initial parameter setup for the algorithm based on the experimental results

Experiment no.	Number of agents	Number of leader–type agents	Number of groups	Initial chromosome population (M)	Max_Gen
1	100	84 (84%)	10	300	200
2	200	152 (76%)	20	300	400
3	400	330 (75%)	40	300	400

In our experiments, we considered a fully connected network of multiple agents. The goal was to evaluate the proposed algorithm's ability to partition a varying number of agents into optimal subgroups, aiming to achieve high closeness centrality by incorporating leadership attributes based on each agent's location. By varying the number of agents, we assessed how effectively the algorithm scaled and adapted to different network sizes.

i) Experiment 1: we started with 100 agents and divided them equally into 10 subgroups, where the number of leader-type agents is 84. In order to achieve optimal groupings of agents that fulfill certain requirements, the COLCF algorithm was employed. We initialized the algorithm with specific parameters as presented in Table 1. The maximum number of generations (Max_Gen) was 200, while the initial chromosome population (M) was 300. The experimental results are shown in Figure 8, indicating that the algorithm quickly converged to the optimal solution after approximately 150 generations. Additionally, the execution time for this experiment was about 4 seconds. The best fitness value found by the algorithm was 0.624895. Moreover, the agent leader of each divided group was detailed under the graph result. For instance, 10–48–1–19–8–20–6–7–78–62 implies that the first group was agent #10 and the second group was agent #48.

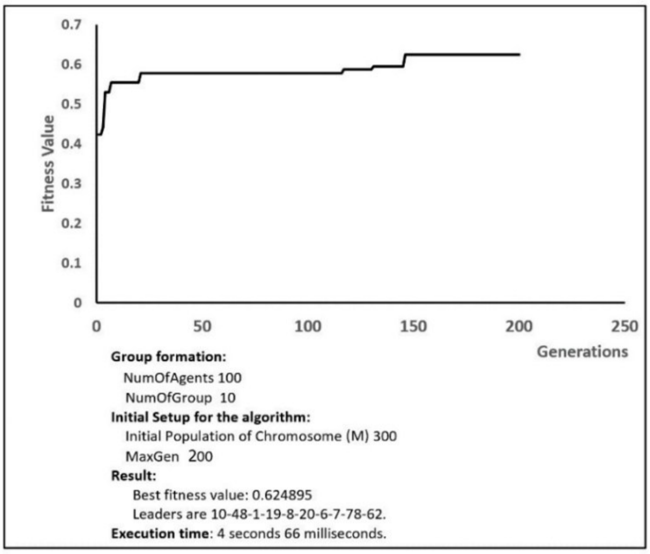

Figure 8. Group formation of 100 agents using the COLCF algorithm (example 1)

ii) Experiment 2: as we aimed to evaluate the scalability of our proposed algorithm by increasing the number of agents and subgroups, the number of agents was increased to 200, which were divided into 20 subgroups. In this experiment, a total of 152 agents (76% of all agents) possessed a leadership attribute. The initial population of chromosomes (M) was kept at 300, while the maximum number of generations (Max_Gen) was set to 400. The experimental result, presented in Figure 9, demonstrates the algorithm's ability to handle larger problems with more agents and subgroups while still converging to the optimal solution. In addition, it shows that our algorithm successfully converged to the optimal solution after 200 generations. This experiment took approximately 15 seconds to complete, and the best fitness value achieved was 0.774177.

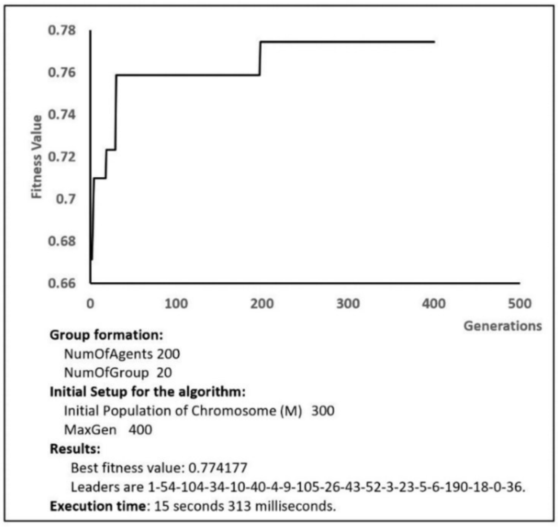

Figure 9. Group formation result for 200 agents using the COLCF algorithm (example 2)

iii) Experiment 3: we increased the problem size to 400 agents, dividing them into 40 subgroups of equal size. There were 330 leader-type agents in total, which was about 75% of all agents. To accommodate the larger problem size, the maximum number of generations (Max_Gen) was up to 400, and the maximum chromosome population (M) was limited to 300. Figure 10 shows that the algorithm successfully converged to the optimal solution after 300 generations. The execution time needed for this experiment was approximately 57 seconds. The COLCF yielded a fitness value of approximately 0.818259. This experimental result indicates that the algorithm was able to search for an optimal solution in forming groups for a larger problem.

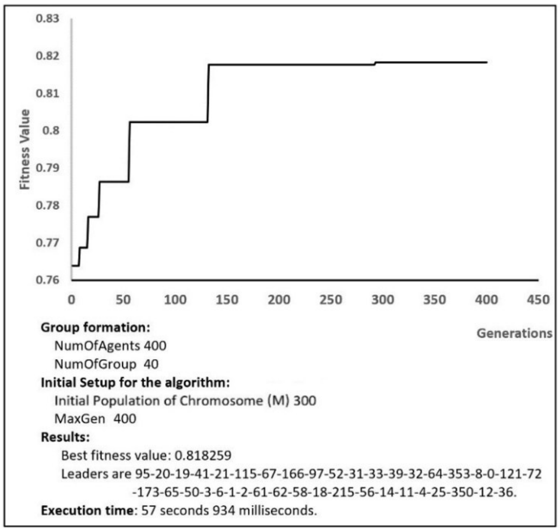

Figure 10. Group formation for 400 agents using the COLCF algorithm (example 3)

4.1. Performance and baseline comparison

The proposed COLCF algorithm's performance was evaluated by adjusting the varying number of agents and groups across three experiments. As shown in Table 2, the algorithm can handle the problem as the number of agents increases. The best fitness values of three experiments are 0.624895 (100 agents), 0.774177 (200 agents), and 0.818259 (400 agents), respectively. Based on Figures 8 to 10, the algorithm demonstrated reliable convergence behavior. For the first experiment with the small problem (100 agents), it was shown that the algorithm typically reaches optimal or near-optimal solutions by the 150th generation, while it reaches the optimal result at around 200 generations in the second experiment. However, in the third experiment with the largest problem (400 agents), convergence extended toward the 300th generation, showing that increased scale can lead to a higher number of generations required for convergence. Execution time also scaled with problem size, rising from 4 seconds to 57 seconds. While this increase is expected, it remains acceptable considering the improved fitness values.

In this paper, we use the fitness of the initial random population (generation 0) as a baseline to evaluate the effectiveness of the proposed algorithm. The fitness of the best chromosome in each generation was recorded. By comparing the fitness values from generation 0 to the final generation, the average improvement gained through the COLCF was demonstrated in Table 3. In all experiments, the fitness value of the last generation achieved a higher value than that of the initial generation, demonstrating the improvement gained through the evolutionary process. Specifically, COLCF increased the fitness by 47.63, 15.37, and 7.10% across the three experiments, respectively. Although the percentage improvement decreases as the number of agents increases, this trend is expected due to the rising complexity of the grouping problem. The decreasing percentage is therefore not indicative of poor performance but rather a reflection of the growing difficulty of the problem.

Table 2. Performance and scalability summary

Experiment no.	Number of agents	Number of groups	Best fitness value	Convergence generation	Execution time (s)
1	100	10	0.624895	~150	4
2	200	20	0.774177	~200	15
3	400	40	0.818259	~300	57

Table 3. Average fitness improvement from generation 0 to final generation

Experiment no.	Gen 0 fitness value	Final fitness value	Avg improvement (%)
1	0.423255	0.624895	47.63
2	0.671264	0.774417	15.37
3	0.76395	0.818259	7.10

4.2. Leader proportion on group formation

To optimize group formation, the COLCF algorithm identifies a qualified leader for each subgroup based on centrality and leadership attributes. In each generation, agents and leader agents are encoded within chromosomes and selected from the previous generation according to their fitness values. This process gradually refines the group structure, enhancing connectivity and identifying the most influential leaders within the network. The results indicate that the algorithm can effectively determine suitable leader agents. For example, in experiment 1, agents such as #10 and #48 were identified as leaders of their respective groups. These agents demonstrated high centrality within the graph. Moreover, as the problem scale increased, the algorithm maintained its resilience even in larger, more complex configurations. One factor not examined in this study is the proportion of leader-type agents. As demonstrated previously, the total number of distinct groupings is calculated by the formula: $\frac{n!}{((q+1)!)^r \cdot (q!)^{p-r} \cdot r! \cdot (p-r)!}$, where n is distinct agents, p is the number of groups, $q = \left\lfloor \frac{n}{p} \right\rfloor$, and $r = n \bmod p$. In addition, if n is large, it may lead to more feasible solutions, thereby increasing the complexity of the problem as the search space becomes larger. As the best group leaders need to be identified for each divided group, a larger number of candidate leaders may increase the algorithm's computational complexity. On the other hand, fewer candidate leaders may allow the algorithm to identify the best leader of the divided groups more quickly. This is supported by [82] and [83], affirming that the size of the solution space and candidate selection significantly impact the efficiency of optimization algorithms. Although the experiments are not designed around this factor, the observed proportions (84 at 84%, 152 at 76%, and 330 at 75% in each respective experiment as presented in Table 1) suggest that this factor could affect convergence speed. There has been no exploration or specific analysis conducted on this factor. Therefore, it warrants further investigation, as it has the potential to influence group formation.

5. CONCLUSION

In this paper, the optimized group formation algorithm based on the GA is proposed to form the best possible groupings of agents, considering the closeness centrality of the group leader. The algorithm demonstrates that it works effectively in solving problems with different levels of difficulty and size. Additionally, its scalability was demonstrated by its ability to handle larger problems with more agents and subgroups. The unique features of our approach, including the incorporation of closeness centrality and leadership attributes, distinguish it from other existing methods. Further research can focus on improving the algorithm's performance and efficiency, such as exploring different selection and crossover methods. The proportion of agent leaders will be further analyzed, as it may affect overall group formation. Another potential direction of research could be implemented in the mobile platforms to allow the algorithm's flexibility to be used anywhere and to adapt to different problem domains. Finally, it would be worthwhile to apply the algorithm to real-world problems and measure its performance in comparison to other existing algorithms.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Anon Sukstrienwong	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓		
Sorapak Pukdesree				✓		✓				✓				

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

Data availability is not applicable to this paper as all data were generated through random simulations and were not stored. However, the simulation process is fully described in the methods section, and results are reproducible using the described algorithm.

REFERENCES

[1] A. M. Kerkmann, S. Cramer, and J. Rothe, "Altruism in coalition formation games," *Annals of Mathematics and Artificial Intelligence*, vol. 92, no. 3, pp. 601–629, Jul. 2023, doi: 10.1007/s10472-023-09881-y.

[2] R. Sultana and V. Kavitha, "Cooperate or compete: coalition formation in congestion games," in *2024 IEEE 63rd Conference on Decision and Control (CDC),* Milan, Italy, 2024, pp. 1731-1738, doi: 10.1109/CDC56724.2024.10885950.

[3] S. Vosooghi, "Coalition formation in public-good games and the power of information design," *SSRN*, 2024.

[4] J. Chen, M. Guo, B. Xin, Q. Wang, S. Lu, and Y. Wang, "Coalition formation problem: a capability-centric analysis and general model," *Science China Information Sciences*, vol. 67, Oct. 2024, doi: 10.1007/s11432-023-4169-2.

[5] N. Maqtary, A. Mohsen, and K. Bechkoum, "Group formation techniques in computer-supported collaborative learning: a systematic literature review," *Technology, Knowledge and Learning*, vol. 24, no. 2, pp. 169–190, 2019, doi: 10.1007/s10758-017-9332-1.

[6] S. Sarkar, M. C. Malta, and A. Dutta, "A survey on applications of coalition formation in multi-agent systems," *Concurrency and Computation: Practice and Experience*, vol. 34, no. 11, Feb. 2022, doi: 10.1002/cpe.6876.

[7] A. Andrejczuk, R. Berger, J. A. R.-Aguilar, C. Sierra, and V. M.-Puchades, "From agents to teams: a survey on task allocation via coalition formation in multi-agent systems," in *Autonomous Agents and Multiagent Systems*, vol. 12403, Switzerland: Springer, 2021, pp. 1–25, doi: 10.1007/978-3-030-69122-1_1.

[8] T. Rahwan, T. P. Michalak, M. Wooldridge, and N. R. Jennings, "Coalition structure generation: a survey," *Artificial Intelligence*, vol. 229, 2015, doi: 10.1016/j.artint.2015.08.004.

[9] O. Granstrand, "A cooperative game theory systems approach to the value analysis of (innovation) ecosystems," *Technovation*, vol. 130, Feb. 2024, doi: 10.1016/j.technovation.2023.102926.

[10] M. Zumkley, F. Sage, S. D. Panal, and J. Prinz, "Does diversity matter? Explaining the effects of team diversity on team performance using data from a student's business simulation game," *Team Performance Management: An International Journal*, vol. 31, no. 1–2, pp. 1–12, 2025, doi:10.1108/TPM-01-2023-0005.

[11] L. E.-Ugartemendia, M. M. Chávez, and N. G.-Eguía, "Impact of joint venture on its partner company's performance: a systematic literature review," *Managerial Finance*, vol. 51, no. 10, pp: 1514–1530, 2025, doi: 10.1108/MF-07-2024-0517.

[12] M. Bouteraa, A. Boujelbene, and H. Boujelbene, "International joint ventures' knowledge acquisition: critical literature review," *Sustainability*, vol. 15, no. 8, 2023, doi: 10.3390/su15086364.

[13] H. Chai, T. Hu, and L. Wu, "Computer-based assessment of collaborative problem solving skills: a systematic review of empirical research," *Educational Research Review*, vol. 43, 2023, doi: 10.1016/j.edurev.2023.100591.

[14] Y. Chen, Y. Wang, Y. Jin, T. Wang, X. Nie, and T. Yan, "A survey of intelligent multi-agent formation control," *Applied Sciences*, vol. 13, no. 10, May 2023, doi: 10.3390/app13105934.

[15] S. Cohen, "Coalition formation in sequential decision-making under uncertainty," in *22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, London, United Kingdom, May 2023, pp. 2955–2964, doi: 10.5555/3545946.3599136.

[16] A. W. Bausch, "Coalition formation and selectorate theory: an experiment," *Political Science Research and Methods*, vol. 5, no. 2, pp. 261–275, 2017, doi: 10.1017/psrm.2015.63.

[17] H. Ding, Y. Niu, C. Han, and P. Xiang, "A coalition formation game-based multi-user grouping approach in the jamming environment," *Electronics*, vol. 11, no. 14, 2022, doi: 10.3390/electronics11142241.

[18] F. Rossi, S. Bandyopadhyay, M. Wolf, and M. Pavone, "Review of multi-agent algorithms for collective behavior: a structural taxonomy," *IFAC-PapersOnLine*, vol. 51, no. 12, pp. 112–117, 2018, doi: 10.1016/j.ifacol.2018.08.066.

[19] M. Bullinger and R. Romen, "Online coalition formation under random arrival or coalition dissolution," in *31st Annual European Symposium on Algorithms (ESA)*, Germany: Dagstuhl publishing, 2023, vol. 274, pp. 1-18, doi: 10.4230/LIPIcs.ESA.2023.27.

[20] C. Zhu, M. Dastani, and S. Wang, "A survey of multi-agent deep reinforcement learning with communication," *Autonomous Agents and Multi-Agent Systems*, vol. 38, no. 1, Jan. 2024, doi: 10.1007/s10458-023-09633-6.

[21] M. Bullinger and R. Romen, "Stability in online coalition formation," in *Proceedings of the 38th AAAI Conference on Artificial Intelligence*, Vancouver, British Columbia, Canada: AAAI Press, 2024, vol. 38, no. 9, pp. 9537–9545, doi: 10.1609/aaai.v38i9.28809.

[22] N. R. Jennings, T. Rahwan, and M. Wooldridge, "Team formation and collaboration: a comprehensive survey on group formation methods and applications," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 49, no. 1, pp. 1–18, Jan. 2019, doi: 10.1109/TSMC.2018.2839847.

[23] N. van der Meer, V. van der Werf, W.-P. Brinkman, and M. Specht, "Virtual reality and collaborative learning: a systematic literature review," *Frontiers in Virtual Reality*, vol. 4, 2023, doi: 10.3389/frvir.2023.1159905.

[24] I. Ignacio and H. C. Chen, "Digital collaborative learning in higher education: a systematic review," *International Journal of Academic Research in Progressive Education and Development*, vol. 13, no. 1, pp. 1111–1120, 2024, doi: 10.6007/IJARPED/v13-i1/20697.

[25] M. H. Lee, C. H. Choi, and J. H. Lee, "A systematic review of collaborative learning in virtual environments: effects and design principles," *IEEE Transactions on Learning Technologies*, vol. 15, no. 1, pp. 34–48, 2022, doi: 10.1109/TLT.2021.3092023.

[26] D. Ramírez, E. J. G.-Lavín, J. Pulgar, and C. Candia, "Self-select groups increase stability but reduce collaboration in high school physics," *Physical Review Physics Education Research*, vol. 21, Jun. 2025, doi: 10.1103/PhysRevPhysEducRes.21.010157.

[27] R. J. Tumpa, S. Skaik, M. Ham, and G. Chaudhry, "A holistic overview of studies to improve group-based assessments in higher education: a systematic literature review," *Sustainability*, vol. 14, no. 15, 2022, doi: 10.3390/su14159638.

[28] A. Krouska, C. Troussas, and M. Virvou, "Applying genetic algorithms for student grouping in collaborative learning: a synthetic literature review," *Intelligent Decision Technologies*, vol. 13, no. 4, pp. 395–406, Feb. 2020, doi: 10.3233/IDT-190184.

[29] D. A. Donovan, G. L. Connell, and D. Z. Grunspan, "Student learning outcomes and attitudes using three methods of group formation in a nonmajors biology class," *CBE--Life Sciences Education*, vol. 17, no. 4, 2018, doi: 10.1187/cbe.17-12-0283.

[30] J. S. Bergtold, "Assessment of group formation methods on performance in group-based learning activities," *Frontiers in Education*, vol. 9, 2024, doi: 10.3389/feduc.2024.1362211.

[31] N. J. Francis, C. Pritchard, Z. Prytherch, and S. Rutherford, "Making teamwork work: enhancing teamwork and assessment in higher education," *FEBS Open Bio*, vol. 15, no. 1, pp. 35–47, Jan. 2025, doi: 10.1002/2211-5463.13936.

[32] I. M. M. Ramos, D. B. Ramos, B. Gadelha, and E. H. T. Oliveira, "Group formation in programming teaching: a systematic review of literature," *Revista de Educación a Distancia*, vol. 24, no. 74, pp. 1–20, 2024,

[33] P. B. Miranda, R. F. Mello, and A. C. Nascimento, "A multi-objective optimization approach for the group formation problem," *Expert Systems with Applications*, vol. 162, 2020, doi: 10.1016/j.eswa.2020.113828.

[34] L. Boongasame, P. Temdee, and F. Daneshgar, "Forming buyer coalition scheme with connection of a coalition leader," *Journal of Theoretical and Applied Electronic Commerce Research*, vol. 7, no. 1, pp. 111–122, 2012, doi: 10.4067/S0718-18762012000100009.

[35] K. Fransen, E. Delvaux, B. Mesquita, and S. V. Puyenbroeck, "The emergence of shared leadership in newly formed teams with an initial structure of vertical leadership: a longitudinal analysis," *Journal of Applied Behavioral Science*, vol. 54, no. 2, pp. 140–170, 2018, doi: 10.1177/0021886318756359.

[36] M. S. Rodgers, T. Y. Kim, T. Chen, and E. David, "Effects of leader group prototypicality on leadership outcomes through leader identity threat: the moderating effects of leader perceptions of organizational support," *Journal of Management Studies*, vol. 61, no. 4, pp. 1427–1456, 2024.

[37] J. Zhang and Y. Luo, "Degree centrality, betweenness centrality, and closeness centrality in social network," in *Proceedings of the 2nd International Conference on Modelling, Simulation and Applied Mathematics (MSAM)*, Bangkok, Thailand: Atlantis Press Mar. 2017, doi: 10.2991/msam-17.2017.68.

[38] N. Meshcheryakova and S. Shvydun, "A comparative analysis of centrality measures in complex networks," *Automation and Remote Control*, vol. 85, no. 8, pp. 773–784, Aug. 2024, doi: 10.31857/S0005117924080031.

[39] G. Istrate, C. Bonchiș, and C. Gatină, "It's not whom you know, it's what you, or your friends, can do: coalitional frameworks for network centralities," in *19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, Auckland, New Zealand, 2020, pp. 566–574, doi: 10.5555/3398761.3398830.

[40] M. Mitterlechner, "Leadership in integrated care networks: a literature review and opportunities for future research," *International Journal of Integrated Care*, vol. 20, no. 3, pp. 1–14, 2020, doi: 10.5334/ijic.5420.

[41] Y. Yuan and D. van Knippenberg, "Leader network centrality and team performance: team size as moderator and collaboration as mediator," *Journal of Business and Psychology*, vol. 37, no. 2, pp. 283–296, 2022, doi: 10.1007/s10869-021-09745-4.

[42] M. C. Camur and C. Vogiatzis, "A survey on optimization studies of group centrality metrics," *Networks*, vol. 84, no. 4, pp. 491–508, Dec. 2024, doi: 10.1002/net.22248.

[43] X. Wang, G. Song, and R. Ghannam, "Enhancing teamwork and collaboration: a systematic review of algorithm-supported pedagogical methods," *Education Sciences*, vol. 14, no. 6, 2024, doi: 10.3390/educsci14060675.

[44] I. Belik, "Measuring group leadership in networks based on Shapley value," *Social Network Analysis and Mining*, vol. 13, no. 1, 2023, doi: 10.1007/s13278-023-01032-9.

[45] J. Park, K. M. Campbell, and S. Walker, "Understanding work teams from a network perspective: a review and future research directions," *Journal of Engineering and Technology Management*, vol. 63, 2022, doi: 10.1016/j.jengtecman.2021.101695.

[46] R. Sethi and C. Y. Nicholson, "Structural and contextual correlates of charged behavior in product development teams," *Journal of Product Innovation Management*, vol. 18, no. 3, pp. 154–168, 2001, doi: 10.1111/1540-5885.1830154.

[47] S. Isotani, A. Inaba, M. Ikeda, and R. Mizoguchi, "An ontology engineering approach to the realization of theory-driven group formation," *International Journal of Computer-Supported Collaborative Learning*, vol. 4, no. 4, pp. 445–478, 2009, doi: 10.1007/s11412-009-9072-x.

[48] C. W. Leem and J. J. G. Chen, "Fuzzy-set-based machine-cell formation in cellular manufacturing," *Journal of Intelligent Manufacturing*, vol. 7, no. 5, pp. 355–364, 1996, doi: 10.1007/bf00123911.

[49] B. Guo, H. He, Z. Yu, D. Zhang, and X. Zhou, "GroupMe: supporting group formation with mobile sensing and social graph mining," in *International Conference on Mobile and Ubiquitous Systems: Computing, Networking, and Services*, Hagenberg, Austria: Springer, Dec. 2012, pp. 200–211, doi: 10.1007/978-3-642-40238-8_17.

[50] T. L. B. Tseng, C. C. Huang, H. W. Chu, and R. R. Gung, "Novel approach to multi-functional project team formation," *International Journal of Project Management*, vol. 22, no. 2, pp. 147–159, 2004, doi: 10.1016/S0263-7863(03)00058-9.

[51] M. Indrawan, T. Kijthaweesinpoon, B. Srinivasan, and A. S. M. Sajeev, "Coalition formation protocol for e-commerce," in *International Conference on Intelligent Sensing and Information Processing*, Chennai, India: IEEE, Jan. 2004, pp. 403–408, doi: 10.1109/ICISIP.2004.1287691.

[52] F. Yu, T. Kaihara, and N. Fujii, "Coalition formation based multi-item multi-attribute negotiation of supply chain networks," *Procedia CIRP*, vol. 7, pp. 85–90, 2013, doi: 10.1016/j.procir.2013.05.015.

[53] E. O.-Ponce, A. Clarke, and B. A. Colbert, "Sustainability cross-sector partnerships: the strategic role of organizational structures," *Corporate Social Responsibility and Environmental Management*, vol. 27, pp. 2122–2134, 2021, doi: 10.1002/csr.1952.

[54] M. Chamberlin, J. D. Nahrgang, H. Sessions, and B. de Jong, "An examination of shared leadership configurations and their effectiveness in teams," *Journal of Organizational Behavior*, vol. 45, pp. 595–619, 2024, doi: 10.1002/job.2774.

[55] G. Hoang, T. T. Luu, and M. T. Babalola, "Entrepreneurial leadership: a systematic literature review and research agenda," *Leadership & Organization Development Journal*, vol. 46, no. 2, pp. 285–313, Dec. 2024, doi: 10.1108/LODJ-05-2024-0350.

[56] A. Mastrorilli, F. P. Santarpia, and L. Borgogni, "Team collective efficacy as a mediator of coaching leadership effects on exhaustion: a multilevel investigation," *Team Performance Management*, vol. 31, no. 1/2, pp. 38–62, 2025, doi: 10.1108/TPM-03-2024-0029.

[57] J. G.-Pérez, A. G.-Naveira, M. M.-Vilar, and J. A.-Sánchez, "Key structure and processes in esports teams: a systematic review," *Current Psychology*, vol. 43, no. 6, pp. 20355–20374, 2024, doi: 10.1007/s12144-023-04650-6.

[58] D. Lundqvist, A. Wallo, A. Coetzer, and H. Kock, "Leadership and learning at work: a systematic literature review of learning-oriented leadership," *Journal of Leadership & Organizational Studies*, vol. 30, no. 2, pp. 205–238, 2022, doi: 10.1177/15480518221133970.

[59] C. Tuschner, J. Krath, J. Bings, M. Schwenkmezger, M. Etzkorn, and H. von Korflesch, "Leading in the digital age: a systematic review on leader traits in the context of e-leadership," in *30th European Conference on Information Systems (ECIS)*, Timisoara, Romania: AISeL, 2022.

[60] B. Xie, X. Gu, J. Chen, and L. Shen, "A multi-responsibility-oriented coalition formation framework for dynamic task allocation in mobile-distributed multi-agent systems," *International Journal of Advanced Robotic Systems*, vol. 15, no. 6, 2018, doi: 10.1177/1729881418813037.

[61] M. Piccione and R. Razin, "Coalition formation under power relations," *Theoretical Economics*, vol. 4, no. 1, pp. 1–15, 2009.

[62] S. Breban and J. Vassileva, "A coalition formation mechanism based on inter-agent trust relationships," in *1st International Joint Conference on Autonomous Agents and Multiagent Systems*, Bologna, Italy: Association for Computing Machinery, Jul. 2002, pp. 306–307, doi: 10.1145/544741.544812.

[63] L. G. Nardin and J. S. Sichman, "Trust-based coalition formation: a multiagent-based simulation," in *Proceedings of the 4th World Congress on Social Simulation*, Santa Monica, CA, USA: Computational Social Science Society, 2012.

[64] M. K. Tarkowski, P. L. Szczepański, T. P. Michalak, P. Harrenstein, and M. J. Wooldridge, "Efficient computation of semivalues for game-theoretic network centrality," *Journal of Artificial Intelligence Research*, vol. 63, pp. 145–189, Oct. 2018, doi: 10.1613/jair.1.11239.

[65] X. Molinero, F. Riquelme, and M. Serna, "Power indices of influence games and new centrality measures for agent societies and social networks," in *Advances in Intelligent Systems and Computing*, vol. 291, Cham: Springer, 2014, pp. 23–30, doi: 10.1007/978-3-319-07596-9_3.

[66] S. M. Arul, G. Senthil, S. Jayasudha, A. Alkhayyat, K. Azam, and R. Elangovan, "Graph theory and algorithms for network analysis," in *E3S Web of Conferences*, Les Ulis, France: EDP Sciences, 2023, vol. 399, doi: 10.1051/e3sconf/202339908002.

[67] C. Perez and R. Germon, "Graph creation and analysis for linking actors: application to social data," in *Automating Open Source Intelligence: Algorithms for OSINT*, 2016, pp. 103–129, doi: 10.1016/B978-0-12-802916-9.00007-5.

[68] Y. Xiao, Y. Chen, H. Zhang, X. Zhu, Y. Yang, and X. Zhu, "A new semi-local centrality for identifying influential nodes based on local average shortest path with extended neighborhood," *Artificial Intelligence Review*, vol. 57, Apr. 2024, doi: 10.1007/s10462-024-10725-2.

[69] Y. V Nandini, T. J. Lakshmi, M. K. Enduri, and H. Sharma, "Link prediction in complex networks using average centrality-based similarity score," *Entropy*, vol. 26, no. 6, 2024, doi: 10.3390/e26060433.

[70] N. Bendahman and D. Lotfi, "Unveiling influence in networks: a novel centrality metric and comparative analysis through graph-based models," *Entropy*, vol. 26, no. 6, 2024, doi: 10.3390/e26060486.

[71] M. Guo, B. Xin, J. Chen, and Y. Wang, "Multi-agent coalition formation by an efficient genetic algorithm with heuristic initialization and repair strategy," *Swarm and Evolutionary Computation*, vol. 55, Jun. 2020, doi: 10.1016/j.swevo.2020.100686.

[72] M. A. Londe, L. S. Pessoa, C. E. Andrade, and M. G. C. Resende, "Biased random-key genetic algorithms: a review," *European Journal of Operational Research*, vol. 321, pp. 1–22, 2025, doi: 10.1016/j.ejor.2024.03.030.

[73] O. R. Sánchez, C. A. Collazos, and M. A. Redondo, "A strategy based on genetic algorithms for forming optimal collaborative learning groups: an empirical study," *Electronics*, vol. 10, no. 4, 2021, doi: 10.3390/electronics10040463.

[74] S. Katoch, S. S. Chauhan, and V. Kumar, "A review on genetic algorithm: past, present, and future," *Multimedia Tools and Applications*, vol. 80, no. 5, pp. 8091–8126, 2021, doi: 10.1007/s11042-020-10139-6.

[75] J. Thompson, "Genetic algorithms and applications," in *Handbook of Formal Optimization*, Springer, 2024, pp. 981–1006, doi: 10.1007/978-981-97-3820-5_30.

[76] S. Parvaze *et al.*, "Optimization of water distribution systems using genetic algorithms: a review," *Archives of Computational Methods in Engineering*, vol. 30, pp. 4209–4244, 2023, doi: 10.1007/s11831-023-09944-7.

[77] I. Rahimi, A. H. Gandomi, F. Chen, and E. M.-Montes, “A review on constraint handling techniques for population-based algorithms: from single-objective to multi-objective optimization,” *Archives of Computational Methods in Engineering*, vol. 30, pp. 2181–2209, 2023, doi: 10.1007/s11831-022-09859-9.
[78] T. Zhang, S.-C. Chu, S. Weng, T.-K. Dao, and T.-T. Nguyen, “Analysis of schema formation in genetic algorithms: a review,” in *Genetic and Evolutionary Computing*, Singapore: Springer, 2025, pp. 265–276, doi: 10.1007/978-981-96-1535-3_27.
[79] C. Liang, Y. Toyokawa, and H. Ogata, “Optimizing group formation with a mixed genetic algorithm: an empirical study in active reading using marker data,” *International Journal of Computer-Supported Collaborative Learning*, vol. 20, pp. 519–548, 2025, doi: 10.1007/s11412-025-09452-9.
[80] R. Kumar, A. Singh, and S. K. Singh, “Genetic algorithms in optimization: a systematic review and future directions,” *Archives of Computational Methods in Engineering*, vol. 29, no. 4, pp. 1831–1856, 2022, doi: 10.1007/s11831-021-09622-6
[81] T. Harada and E. Alba, “Parallel genetic algorithms: a useful survey,” *ACM Computing Surveys (CSUR)*, vol. 53, no. 4, pp. 1-39, 2020, doi: 10.1145/3400031.
[82] J. Liu, R. Sarker, S. Elsayed, D. Essam, and N. Siswanto, “Large-scale evolutionary optimization: a review and comparative study,” *Swarm and Evolutionary Computation*, vol. 78, 2023, doi: 10.1016/j.swevo.2023.100983.
[83] M. Xu, Y. Mei, F. Zhang, and M. Zhang, “Learn to optimise for job shop scheduling: a survey with comparison between genetic programming and reinforcement learning,” *Artificial Intelligence Review*, vol. 58, 2025, doi: 10.1007/s10462-024-11059-9.
[84] V. A. Cicirello, “Evolutionary computation: theories, techniques, and applications,” *Applied Sciences*, vol. 14, no. 6, Mar. 2024, doi: 10.3390/app14062542.
[85] X. Wang, D. Yang, and J. Lyu, “Genetic algorithm-based secure cooperative control for high-order nonlinear multi-agent systems with unknown dynamics,” *Journal of Cloud Computing*, vol. 13, no. 1, Jan. 2024, doi: 10.1186/s13677-023-00532-5.
[86] D. Farinati and L. Vanneschi, “A survey on dynamic populations in bio-inspired algorithms,” *Genetic Programming and Evolvable Machines*, vol. 26, 2024, doi: 10.1007/s10710-024-09492-4.
[87] G. Papazoglou and P. Biskas, “Review and comparison of genetic algorithm and particle swarm optimization in the optimal power flow problem,” *Energies*, vol. 16, no. 3, 2023, doi: 10.3390/en16031152.

BIOGRAPHIES OF AUTHORS

Anon Sukstrienwong currently serves as an associate professor of Computer Science in the Department of Information Technology and Innovation at Bangkok University, Thailand. He graduated with a Bachelor of Science in Applied Mathematics from King Mongkut's Institute of Technology Ladkrabang, Thailand, in 1994. He later earned a master’s degree from the University of Colorado at Denver, USA, in 1999. His research interests include optimization algorithms, collaborative systems, and information technology. His work focuses on developing advanced systems for problem-solving and managing complex information technologies. Additionally, his research contributes to bridging theoretical approaches with practical applications in real-world environments. He has published in the fields of optimization and computer science, contributing to both theoretical and applied aspects of these areas. He can be contacted at email: anon.su@bu.ac.th.

Sorapak Pukdesree is currently affiliated with Bangkok University. He holds a bachelor’s degree in Computer Science from Bangkok University, earned in 1994. He obtained his master degree in Information Technology from the University of Tennessee at Chattanooga, USA, in 1999, and completed his Ph.D. in Computer Science at Rangsit University in 2022. His research interests include data management strategies, biometric authentication, cloud computing, and IoT applications. He has published numerous papers on integrating technology in education and developing efficient data systems. He can be contacted at email: sorapuck.p@bu.ac.th.