Group Composition for Collaborative Learning With Distributed Leadership in MOOCs Using Particle Swarm Optimization

Matheus R D Ullmann*†, Deller F James‡, Celso G Camilo-Júnior‡ and Tiago do C Nogueira*
*School of Electrical, Mechanical and Computer Engineering
Federal University of Goiás, Goiânia - Goiás - Brazil
Email: matheusullmann@gmail.com
†Federal Institute of Bahia - Campus Barreiras, Barreiras - Bahia - Brazil
‡Institute of Informatics, Federal University of Goiás
Email: dellerjames@gmail.com

***Abstract*—The Massive Open Online Courses (MOOCs) are online courses with open enrollment that involves a huge amount of students from different locations, with different backgrounds and interests. The large number of students implies an enormous and not manageable amount of interactions. This fact, along with the different interests of the students, results in low quality interactions. Due to the large amount of students, it also becomes impossible the composition of learning groups manually. Due to these characteristics present in MOOCs, a method for the formation of groups was developed in this work, as an attempt to meet the dichotomy that exists between the collective. For the formation of groups, an adaptation of the Particle Swarm Optimization algorithm was proposed on the basis of three criteria, level of knowledge, interests and leadership profiles, forming groups with different levels of knowledge, interests similar and distributed leadership, providing a better interaction and construction of knowledge. On a test basis, the algorithm demonstrated that can meet the criteria for grouping in a computation time and is more efficient than the model of random groups. The tests also showed that the algorithm is robust considering the various data sets and variations of iterations.**

***Index Terms*—Massive Open On-line Courses; CSCL; Colaborative Learning; Distance Learning; Particle Swarm Optimization, Group Formation**

I. Introduction

Trend in the area of distance education, the MOOCs (Massive Open Online Courses) are courses linked to the democratization of knowledge, integrating social networks connectivity, access to materials of an expert in a field of study and a collection of free resources for access [1], [2]. Therefore, you can set two essential characteristics for this type of course, should be open and free courses and must allow scalability, providing global coverage.

In the last five years, has seen a great increase in the popularity of Massive Open Online Courses, or MOOCs. This increase in popularity can be reflected both in the number of students enrolled and in the number of universities now offer courses in this format [3]. The MOOCs have the potential to become a new important mechanism for learning. Despite this early promise, however, the MOOCs are still relatively unexplored and little understood [4].

Arguably, the mass adoption of the Internet in recent years, is promoting educational and cultural change never seen before. This digital revolution is contributing to change teaching and learning processes significantly and irreversibly [5]. Increasing connectivity has opened new avenues for learning daily and the influence of connectivity, providing access to digital knowledge diffused in several new tools present on the Web, you can to redefine and take an opinion different about what we understand as education and learning [5]. With the rise of social media, communication, learning and entertainment can now be intrinsically linked, making people increasingly dependent on connecting with others in the virtual world.

This kind of course, by owning as feature classes with a large number of students, eliminates the need for timing in education and the need for students and instructors to be in the same location. However, the very features that enable massive scalability of open online courses, they also bring significant challenges to your teaching, development and management. In particular, the scale makes it difficult for instructors to interact with the many students. The lack of interaction and feelings of isolation has been assigned as reasons for why students give up the MOOCs before your termination [6]. The vast majority of articles that talk about the MOOCs are emphatic in saying that the approval rate is low, spinning around 10% generally [7].

The MOOCs require students depend on each other to have the necessary assistance with regard to the content. The instructors on MOOC have fewer tools with which can improve the travel experiences than they would have in online courses where you don't enroll tens of thousands of students [8]. Critics of the MOOCs argue that a very small percentage of students actually complete most MOOCs because your pedagogical model leaves much to be desired when compared to the educational experiences that enable close interaction between teachers and students, although these courses are

978-1-7281-6929-3/20/$31.00 ©2020 IEEE

largely composed of students who already have a high level of education [8]. However, the interactions between students can be improved by an efficient group, distributing the role of the teacher to the students in each group, allowing the student to have a well-defined role, seeking a greater commitment and productivity of the student [5].

In this sequence, Harris et al. [5] shall ensure that the distribution of the teacher's role in digital media effectively can be a major item for safe and harmonious participation in online groups, so that such distribution is best achieved through a shared form of power within the group or, more specifically, through distributed leadership.

In a MOOC, many students are shy and do not interact much with the others in the forums. It can feature a very large loss for them as they may lose a chance to address any questions that might be of extreme importance and liable to be shared in the forums [9]. So, you have to work in the aspect of student participation in the course at a more personal level, causing them to interact in small groups that have the same interests within the course. In this sense, students must have personal spaces (small groups) to share their interests motivations. A questionnaire to raise the profile of each student allows you to gather data on the knowledge, abilities and interests of each student, and this information can be used to create study groups with the greatest potential for interactions [10]. However, an effective collaborative learning experience is influenced by some factors, among them are the attitude of the student to seek knowledge and learning, interaction with other students and group dynamics [10].

The online collaboration is a potentially powerful form of learning. For this reason the greater access to digital tools is fast becoming a usual medium where collaborative learning happens. Some authors even speculate that virtual learning is a natural and fundamental part of the future of education [5]. Thus, we can conclude that prior knowledge and skills are necessary but not sufficient conditions for students in MOOCs complete successfully the course [3].

The formation of group proposed in this paper takes advantage of the diversity among the students and, at the same time, makes it manageable the different interests among students, in addition to incorporating the concept of distributed leadership. Principles of formation of groups are proposed in this paper as an attempt to meet the dichotomy that exists between the collective, which involves the formation of an online community of learning on a massive scale, and the individual, with different interests, previous knowledge and expectations. The groups are heterogeneous with respect to the level of knowledge and homogeneous regarding interests, namely, the groups have students with different levels of knowledge and interests, and each group has distributed leadership. The groups are formed automatically by using Particle Swarm Optimization algorithm (PSO). The use of such an algorithm is necessary due to the complexity of the problem. Given the large volume of students, the groups are formed by hand, but by a computational technique.

Over the years, the meta-heuristics for Particle Swarm Optimization (PSO) has been used to solve various problems of complexity NP-Hard [11]–[21], as is the case of this formation of groups. The results of these studies indicate that it is possible to solve NP-hard problems effectively and efficiently using PSO. Thus, for this study, were the tests in order to verify and validate your efficiency and applicability to the problem in question. In experiments performed by the algorithm, the number of students and the number of criteria for the formation of groups, for example, were varied and compared with the formation of random groups, getting better results, not just in the final result, but also in the Runtime.

The present work is divided as follows, II session, are presented the theoretical assumptions for collaborative learning, the criteria for the formation of knowledge level and interest groups but also distributed leadership and Introduction to PSO algorithm. In session III, the method used for the formation of groups in MOOCs. In session IV, calibration is displayed and the experiments performed by the PSO algorithm for the formation of groups. V session, presented the conclusions and further work.

II. Related Work

The related works are divided into two aspects, the first is the criteria for the formation of groups in collaborative learning and the second concerns the Particle Swarm Optimization algorithm.

A. Criteria For Forming Groups in Collaborative Learning

The criteria for the formation of groups used in this work, the level of knowledge and interests of students and, also, the concept of distributed leadership, discussed below.

1) Level of Knowledge and Interests: For many years, theories and teaching structures have indicated prior knowledge of students as a major factor for effective learning, in which the development of understanding through the process of assimilation, which is sustained by prior knowledge [3]. From these fundamental concepts, one can infer that the understanding of a student is developed with the construction and modification of their existing knowledge structures. On the other hand, the students' knowledge structures affect their interests.

The effectiveness of any educational situation, particular or environment, is dependent on the Association of different perspectives, experiences and prior knowledge [3]. As in a MOOC don't have the presence of a tutor or teacher, a shed can be merge people with skills and previous knowledge levels higher with people with lower knowledge levels, so that the knowledge can be spread with more ease. It has a very strong association with the performance of students in the course, compared to homogeneous groups with respect to the level of knowledge [3], [22].

Webb [23] states that students with lower levels of knowledge in a particular subject have the studies yield improved when inserted into heterogeneous groups, this is that these students receive more elaborate explanations of his colleagues that have a greater level of knowledge on the subject. Likewise, those students who have a higher level of knowledge also leave benefit, because to explain the contents to other students, they are taken to reorganise the mind making the information more clearly in different aspects [24].

In this context, another aspect that can be taken into account in the formation of groups in collaborative learning is the interest of the members of the group, since the interest has the potential to change the individuals ' involvement in learning [25]. In this sense, many services have been implemented in order to bring together people with common interests [26].

In this way, the project must be carried out of appropriate structures that exploit the points in common between the interests of students to create virtual interactions that promote effective socialization, so that the members of a group can share maximum common interests as possible, that is, one should maximize the degree of common interest within each group [26].

The level of understanding or understanding and the interests of students are two criteria that are often used by teachers in the classrooms of the real world. Studies indicate that formation of groups based on different levels of understanding and similar interests can encourage better discussions during the learning period [27]. However, for the shared knowledge creation process, two or more individuals must possess complementary skills for interacting, creating a shared understanding [5]. In addition, teachers can take into account the interests of students in the formation of groups, making these groups achieve a higher level of interaction [16].

In this way, it is laid down that the prior knowledge and skills, are these content knowledge or generic learning skills, such as problem-solving, and the interests of students about the contents and objectives in a course, can influence considerably on the success of students ' learning. However, there are other criteria that can be used in a collaborative learning groups, given the course facilitator deems necessary, for example, the distributed leadership.

2) Distributed Leadership: According to Lourenço [28], leadership is a neologism corresponding to the leader, and thus, the term leadership, have meanings like, leader or governance. In the literature, there are several different definitions of leadership, however, a simple definition can be characterized in the fact of a specific person in a group set-up the leadership role and have, as a rule, responsibilities and functions that cannot be shared with other people in the group.

However, in more recent studies the leadership is being seen as a process of social influence, i.e. the process of influence in a group came to be shared among the members of the group or organization, making this process more diffuse or sparse, Therefore, in the globalized world of today, the traditional notion of leader no longer accurately reflects the reality, with the new term the distributed leadership [5], [28], [29].

The distributed leadership can be defined as a division of tasks or processes between various leaders. This concept of leadership within groups, whether they are working or studying, becomes important, as well as valuing the individuals of the team, seeks to eliminate patterns of behavior that are usually present in teams with a single leader, like authoritarianism, will enforcement dominatrix of the leader, questions by the led on the tasks to be performed and who should assume the leading role.

Harris et al. [5] understand that distributed leadership consists of a social distribution of leadership, where each function of leadership is implemented on the work to be developed and shared between the participants of the group so that the task be accomplished through interaction and collective action. This post, it is understood that the practice of leadership move between those participating in the group, however, for the distributed leadership really happens, is confidence, reciprocal commitment and empathy between team members by encouraging a authentic collaboration, the sharing of information and generation of interdependent ideas [5].

Based on this concept of leadership, which has been changing over time and migrating to the leadership with powers distributed between the members of the group, the profile of each person is important to define and direct the role that each leader will play within the group. In this context, the work of [30] and Northouse Gressick and Derry [31] can be used for the mapping of the cognitive skills of students, as shown in the table I.

TABLE I
NORTHOUSE CONTEXTS AND GRESSICK PROFILES AND DERRY

Northouse [30]	Gressick e Derry [31]	Leader's name
Social Context	Developing Arguments	Leader 1-A
	Ideas Finder/ Contributions	Leader 1-B
Emotional Context	Recognition / Affective	Leader 2-A
	Organizational Movements	Leader 2-B
Cognitive Context	Contribution of knowledge	Leader 3-A
	Control Topic	Leader 3-B

Such skills can fit within the contexts that best describe the training needs of groups within a MOOC, that are members with social skills, cognitive-behavioral and emotional.

However, when there is a very large demand, considering at the same time these three criteria already mentioned clustering (level of knowledge, interests, and distributed leadership), the ideal training becomes practically impossible without the use of computer resources, that is, an NP-hard problem [16], however the PSO algorithm deals very well with this type of problem.

B. The Particle Swarm Oprimization Algorithm

The Particle Swarm Optimization (PSO) was founded in 1995 by Kennedy and Eberhart on observations that shape the "social behavior" of birds or shoals of fish in search of food or your nest [32], [33]. Among the many models available Kennedy and Eberhart if interested in the model developed by biologist Frank Heppner [32], [34], [35].

Heppner's birds have a peculiarity that is the formation of flocks until one or more birds fly over the nest or food. As a bird is what you're looking for, be it the rest or food, attracts the other birds, increasing the chance of them also meet. In other words, the social behavior of fish or birds that provided the inspiration for PSO was foraging behavior, in particular, by observing certain species of fish or birds in the search to find a food source collaboratively [16].

In the theory of particle swarm (Particle Swarm), the particles (called birds in the template for the biology) are generated randomly in a search space [36], [37]. Each particle corresponds to one possible solution represented by your position within that space. To get an optimal solution or close to the great, each particle takes your experience gained previously and also uses the experience gained by the group. So that the particles move in the search space the PSO classic makes use of a vector and a vector of position in order to determine the movement of the particle [38]. Thus, each iteration of the particle algorithm is updated using experience for her, the experience gained by the bunch and your speed.

To update the position of each particle the equation 3 is used.

$$x_{k+1}^i = x_k^i + v_{k+1}^i \tag{1}$$

In this equation, x_{k+1}^i determines the position of the particle i in the iteration $k+1$ and v_{k+1}^i determines the speed. Already the speed vector is given by the equation 2.

$$v_{k+1}^i = w * v_k^i + c_1 * r_1 * \frac{(p^i - x_k^i)}{\Delta t} + c_2 * r_2 * \frac{(p_k^s - x_k^i)}{\Delta t} \tag{2}$$

Where, r_1 and r_2 são are random numbers between 0 and 1, p^i indicates the best position found by the particle i in the search space and p_k^s determines the best position of the bunch towards the iteration k.

The three parameters are problem dependent inertia w that controls the ability of a particle to explore the search space, where a high value determines a more global search and a low value determines a more local search, determining how much of the current speed will stay on next iteration, and the parameters of trust $c1$ (self-confidence) and $c2$ (confidence in the bunch).

Each particle swarm has a vector that represents the best position she found within the search space, this vector is called $pbest_{ij}$. Similarly, the PSO has a vector that holds the best position found among all particles of the swarm, the best overall position, this vector is called $gbest_j$. To determine which is the best position of a particle, the fitness (function that represents the objective to be achieved) is calculated.

However, the formation of groups in MOOCs requires a different modeling in relation to the PSO. To this end, it is convenient to use PSO for optimization of combinatorial problems, a specific modeling of particle and a fitness function modeling for this algorithm. The goal is to improve the process of forming groups so that your composition is done in a time doable and increase the knowledge of all participants in the learning process.

1) The Particle: The particle is presented as a n-dimensional vector i, whose index indicates the student n and the value for that index, represents the Group g for which the student is inserted as shown in the equation 3.

$$P_i = p_{i0g}, p_{i1g}, p_{ijg}, \cdots, p_{ing} \tag{3}$$

To exemplify this representation, we can see that, according to Figure 1, the student 2 belongs to the Group 3, as well as the student 4 belongs to group 2.

Student	0	1	2	3	4	...	n
Group	3	5	3	2	2	...	g

Fig. 1. Representation of a particle

2) PSO For Combinatorial Problems: Using the concept presented by Jarboui et al., [14], [15], the particle represented previously consists of an additional vector whose values represented by it are within the set {-1, 0, 1}. This vector is shown in Equation 4.

$$Y_i = y_{i0}, y_{i1}, y_{ij}, \cdots, y_{in} \tag{4}$$

Additional vector presented by 4 Equation consists of, Y_i which represents the additional vector for the particle i and n represents the size of this n-dimensional vector, which refers to the number of students, that is, has the same size as the particle. Remember that p_{ij} indicates a position j on i particle. The restrictions for the filling in of the additional vector are shown in Equation 5.

$$y_{ij} = \begin{cases} 1, & \text{if } p_{ij} = gbest_j, \\ -1, & \text{if } p_{ij} = pbest_{ij}, \\ -1 \text{ ou } 1, & \text{random if}(p_{ij} = gbest_j = pbest_{ij}), \\ 0, & \text{else.} \end{cases} \tag{5}$$

The calculation of the speed differs slightly from the classic PSO (equation 2). The new equation to calculate the speed of each particle (equation 6) takes into account the additional vector.

$$v_{ij} = w * v_{ij} + r_1 * c_1 * (-1 - y_{ij}) + r_2 * c_2 * (1 - y_{ij}) \tag{6}$$

In Equation 6, v_{ij} indicates the speed calculation for the position j of the particle i. The other variables in the equation are the same variables of the equation 2, differing only in the use of additional vector in two parts of the formula ("$-1-y_{ij}$" and "$1 - y_{ij}$"). After the computation of the velocities of the particle position update uses the equation 7.

$$\gamma_{ij} = y_{ij} + v_{ij} \tag{7}$$

From this formula the value of additional vector is updated again, but now the restrictions imposed in the equation 8.

$$y_{ij} = \begin{cases} 1, & \text{if } \gamma_{ij} > \alpha, \\ -1, & \text{if } \gamma_{ij} < \alpha, \\ 0, & \text{else.} \end{cases} \tag{8}$$

These restrictions, α is a variable that must be informed before the execution of the algorithm. The values are changed in a particle obeying the rules imposed in the equation 9.

$$x_{ij} = \begin{cases} gbest_j, & \text{if } y_{ij} = 1, \\ pbest_{ij}, & \text{if } y_{ij} = -1, \\ \text{random number}, & \text{else.} \end{cases} \quad (9)$$

Thus, it is understood that when the value of the additional vector for the same position of the particle is 1, the value of x_{ij} will be changed to the value that is in the same position of the vector that contains the global positions ($gbest_j$), that is, those that received the best fitness value between all particles. The same case happens when the value of the additional $vector-1$, but the value that will be placed in the particle will be the relation to best position found by the particle ($pbest_{ij}$). Otherwise, a random value within the number of existing groups must be assigned the position referred to.

III. Modeling of Groups in MOOCs

The proposed method is represented by the flowchart in Figure 2, in which a series of questionnaires are applied to users in order to verify the level of knowledge and the identification of the interests of each student, in addition to identifying the your leadership profile. Application of PSO algorithm seeks a great composition of groups according to the criteria established, which are forming groups composed of six members with the same interests, with different levels of knowledge and each group must include all three profiles of leaders addressed in research.

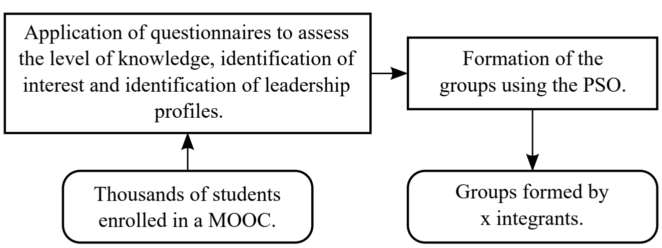

Fig. 2. Diagram for formation of groups at MOOCs

The three phases of the proposed method for formation of MOOCs groups are presented below.

A. Phase 1- Preliminary Questionnaire

In the first phase of the method is answered a questionnaire on the part of the students involved in the studies. This questionnaire serves for the level of knowledge, interests, and the profile of distributed leadership of each student is calculated.

The questionnaires on the level of knowledge and interests shall be adapted by the facilitator of the course to meet the pedagogical needs of the course in which you will use. The third questionnaire to be answered preliminarily by the students concerns the distributed leadership. The questionnaire developed by Northouse [30] can be used in this questionnaire the sum of the answers of the questions indicate which is the leadership profile for each student.

B. Phase 2- Training of Groups

In the second phase of the method, you must enter the data in the proposed algorithm. As described in section II-B1, each position of the vector that represents the particle corresponds to a student, and the value in that position represents the group to which the student is inserted. Analogously, we have three more vectors, one for each category, which are, NIV_n for the levels of knowledge, INT_n to the interests and LID_n for profiles of leadership, as explained in the section III-B1 below.

1) Algorithm Data: The data used by the algorithm in the early running for the formation of collaborative learning groups have different backgrounds for each vector. The Table II represents the source of each data vector and their acronyms.

TABLE II
Data used by the algorithm

Vector	Origin of data	Distribuition	Acronym
Particles	Random	Index: Student Value: Group	P_n
Knowledge	Questionnaire	Index: Student Value: Knowledge Level	NIV_n
Interests	Questionnaire	Index: Student Value: Category of Interest	INT_n
Leadership	Questionnaire	Index: Student Value: Leadership Profile	LID_n

In the vector P_n, what are the particles of the algorithm, the index of each position represents a student and the values entered in each position of the vector represent the groups in which students are inserted, as described in section II-B1. This vector is populated at random, however the values that are used for the full composition of this vector are components of the set represented by the ceiling shown in equation 10.

$$\left\lceil \frac{n}{6} \right\rceil, \text{ subject a } n \in Z_+ \quad (10)$$

Where n is the number of students and 6 the amount of leaders set for each group.

To compose the vector that represents the knowledge levels of each student (NIV_n) uses a questionnaire. For example, it is possible to obtain through a combination of values for the issues. It depends on the questionnaire and objectives of the course facilitator.

The interests are also defined by the facilitator of the MOOC. As an example, the facilitator can divide affirmative questions into categories and such categories, could represent the sort of student interest, filling so the vector INT_n with the results. The sum of the responses (Northouse questionnaire [30]) that make up each distributed leadership context indicates the value that is assigned to the vector LID_n.

$$A_n = \{f_1, f_2, f_3, f_4\} \quad (11)$$

Where,

- f_1 = represents the group that the student is inserted;
- f_2 = represents the level of knowledge;
- f_3 = represents the interest;
- f_4 = represents the context of leadership.

The values of the set represented by the equation 11 were given by the equation 10 and the vectors that represent the criteria for the formation of the group. Therefore, the values of each variable of the set quoted is subject the terms presented in the equation 12.

$$A_n = \begin{cases} f_1 = & \{x \in N : 0 \leq x \leq n/6\} \\ f_2 = & \{x \in N : 0 \leq x \leq 9\} \\ f_3 = & \{x \in Z_+^* : 1 \leq x \leq 5\} \\ f_4 = & \{x \in Z_+^* : 1 \leq x \leq 3\} \end{cases} \tag{12}$$

In this way, It is observed by the equations 11 and 12 that there are 4 variables for each student, the first being dependent on the number of students and the other for the questionnaires applied.

TABLE III
EXAMPLE OF FILLED VECTORS

	Students					
	0	1	2	3	$\cdots$	n
P_n	6	3	1	5	$\cdots$	f_1
NIV_n	4	2	0	2	$\cdots$	f_2
INT_n	2	4	5	3	$\cdots$	f_3
LID_n	2	2	1	3	$\cdots$	f_4

To illustrate the vectors and variables for each student, note the table III. The vectors have the variables of all students, for example, the student 2, is placed in Group 1, has knowledge level 0, 5 interest and leadership profile number 1.

Each of the variables is independent of the other, but they coexist to form the set of attributes of the student. These attributes are used by the fitness function for composition of groups, therefore, the only attribute that can change during the execution of the algorithm is the f_1. The wording of the fitness function is shown below.

2) The Fitness Function: After the data is loaded in the algorithm, the 13 Equation is minimized by the algorithm PSO to achieve the best possible group training for the given set of students.

$$\begin{aligned} & Minimize f(P_i) = 1/\tfrac{(\sum_1^g \sum_1^c DNIV_g) - Tot_g}{n - Tot_g} \\ & + \tfrac{(\sum_1^g \sum_1^c DINT_g) - Tot_g}{n - Tot_g} + 1/\tfrac{(\sum_1^g \sum_1^c DLID_g)}{g} \\ & + Penalty \end{aligned} \tag{13}$$

In this function, n represents the total number of students, Tot_g refers to the total number of groups, c is the number of students within the Group g, $DNIV_g$ indicates the diversity of levels within the group g, $DINT_g$ indicates the diversity of interests within the group, $DLID_g$ represents the number of times that the group g has more or less two leaders of each profile within the Group and the $Penalty$ is calculated according to the equation 16.

The diversity of levels of knowledge ($DNIV_g$) within each group is calculated by performing the intersection between the sets f_2, that has all levels of knowledge and, NIV_g, that has all levels of knowledge covered by the group that is under review, as demonstrated by the equation 14. The diversity of interests ($DINT_g$) is calculated similarly to the diversity of levels of knowledge. Fitness function 13, the values of $DNIV_g$ e $DINT_g$ all groups are added up and the result is normalized on a scale of 0 to 1, therefore, that is why the $n - Tot_g$ division is added to the equation 13.

$$DNIV_g = |f_2 \cap NIV_g| \tag{14}$$

To calculate the diversity of leadership made necessary the use of restrictions exhibited by the equation 15. When the number of a group leadership profiles g ($L1_g$) is different from 2, increases in diversity 1 $DLID_g$, likewise with the leadership profiles 2 and 3 ($L2_g$ and $L3_g$). Fitness function 13, the sum of all the diversity of leaders $DLID_n$ is divided by the number of groups so that the maximum value of diversity is 3.

$$DLID_g = \begin{cases} L1_g \neq 2, & x+1; \\ L2_g \neq 2, & x+1; \\ L3_g \neq 2, & x+1. \end{cases} \tag{15}$$

The purpose of the penalty function (equation 16) is to add 20% to the value of fitness every time a group contains less or more than six students, in order to try to get the optimal number of groups pupils around the six students per group.

$$Penalty = \begin{cases} \text{if } c > 6, & fitness * (1.20)^{qtd_{ocor}} \\ \text{if } c < 6, & fitness * (1.20)^{qtd_{ocor}} \end{cases} \tag{16}$$

Thus, the result of minimizing fitness function results in heterogeneous groups regarding level, because in this format groups provide a more consistent learning [12], [13], homogeneous for the interest of each student, providing groups with a focus on common and with the two leaders for each leadership profile.

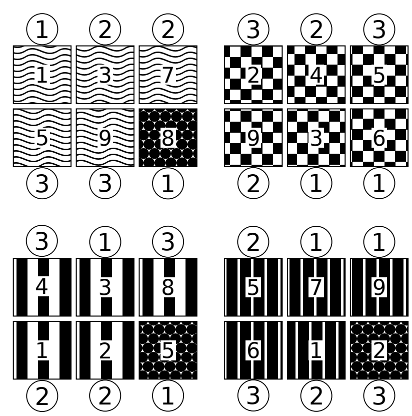

Fig. 3. Groups with different knowledge level (within the square numbers), similar interests (colors) and distributed leadership (numbers inside the circles)

To illustrate what it would be like a heterogeneous group with respect to the level of knowledge, but with the same interests and distributed leadership, see Figure 3. Each square represents one student in a MOOC and grouping of these squares represent the study groups. The texture indicate the type of student interest, the numbers in each square represent the level of your knowledge and the numbers inside the circles indicate the student's leadership profile.

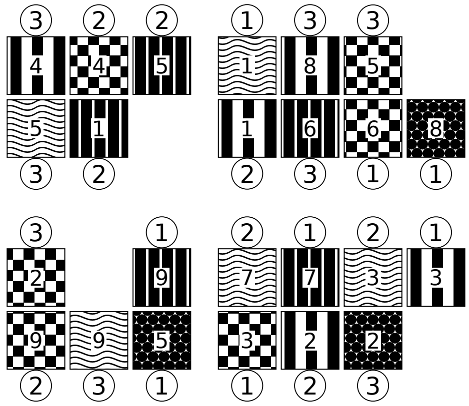

Fig. 4. Groups formed randomly

That way, you can see that the groups in Figure 3 are close to ideal, since there is an obvious diversity of levels of knowledge in each group (represented by the numbers in the square), people have the same type of interest (represented by the textures) is grouped in one group and each group has a pair of each leadership profile. In Figure 4, the groups presented are formed randomly or according to preferences and personal relations of students (considered a random groups for Dascalu et al. [12], [13]) which often causes the groups do not follow the format of educational group proposed in this work.

C. 3 Phase - Interaction of students

In phase 3, students will know which group they belong, which was assigned leader profile and should receive an explanatory text about how every leader must interact within the group. That said, students who share the same context must come into agreement about which role each leader will exercise, for example, the two students who have emotional context profile should talk and decide who will be the leader 2-and who will be the leader 2-B. The demonstration of the contexts and profiles is shown in the table I.

IV. Results

This study employed a series of experiments to analyze the effectiveness and efficiency of the PSO algorithm for group formation in MOOCs. To this end, we used data generated at random and will be described in section IV-A. The algorithm was run on a computer with Intel Xeon 2.0 GHz processor with 4 cores and 4 GB RAM. The implementation was accomplished with the Java language.

Experimental settings of the problem are presented in Section IV-A. Later, in the section IV-B, the robustness of the algorithm with respect to the variation of number of test scenarios. It is also presented a comparison between the PSO and the random algorithm.

A. Experimental Settings

To analyze the effectiveness of the algorithm, 11 data sets were generated randomly varying the parameters of the data. Table IV illustrates each dataset.

In eleven sets of data, the number of students (n) ranged from 12 to 9216. Each dataset was generated only once for the execution of experiments. Students were grouped in 2 the 1536 collaborative learning groups, where g represents the number of groups in each dataset.

TABLE IV
Data set used-Second algorithm

DataSet	N.S. (n)	A.L.	A.I	D.L.	N.G. (g)
1	12	6	5	3	2
2	24	6	5	3	4
3	48	6	5	3	8
4	96	6	5	3	16
5	192	6	5	3	32
6	384	6	5	3	64
7	768	6	5	3	128
8	1536	6	5	3	256
9	3072	6	5	3	512
10	6144	6	5	3	1024
11	9216	6	5	3	1536

Legenda: N.S. (Number of students), A.L. (Amount of levels), A.I. (Amount of interest), D.L. (Distributed Leadership and N.G. (Number of groups).

The PSO used for Combinatorics has four parameters to be adjusted, the inertia (w), the $C1$ that indicates the particle learning, $C2$ that indicates the group learning and alpha (α) what is a parameter used to calculate the new position of the particle. Experimental settings were also varied the number of particles and the number of iterations. To adjust each of these parameters, systematic tests were performed on them, running 30 instances for each variation.

To adjust the values of inertia (w), $C1$, $C2$ and Alpha (α) the number of iterations was set at 100, the number of particles was 20, the number of dimensions or pupils was 192. Each one of the parameters was varied according to the interval described in 17.

$$[-1, -1 + a * 0,05], \text{ where } 1 \leq a \leq 40, \ a \in Z_+ \quad (17)$$

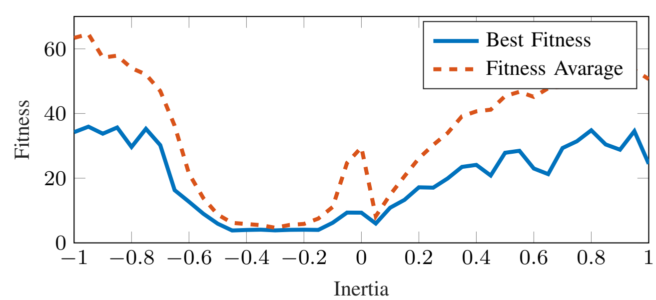

Fig. 5. Variation of inertia

The first variation performed to adjust the algorithm was the inertia, which obtained the best average of *fitness* (the best of each execution) when your value was -0.3 (Figure 5).

The first variation performed to adjust the algorithm was the inertia, which obtained the best average of *fitness* (the best of each execution) when your value was -0.3 (Figure 6).

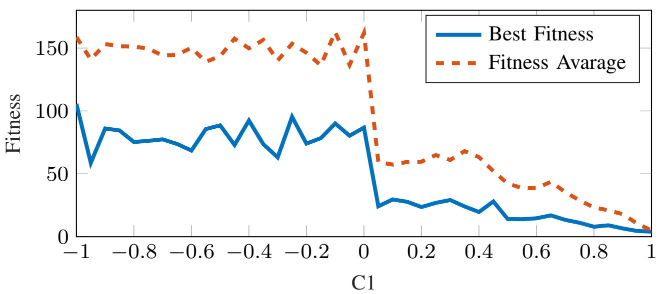

Fig. 6. Variation C1

The next parameter tested was the $C2$, that, similarly to the above, the values of inertia and $C1$ were fixed, making your best result found in $C2$ 0.05 (Figure 7). The fourth parameter tested for the calibration of the algorithm was the alpha, reaching the mark of 1 for best results (Figure 8).

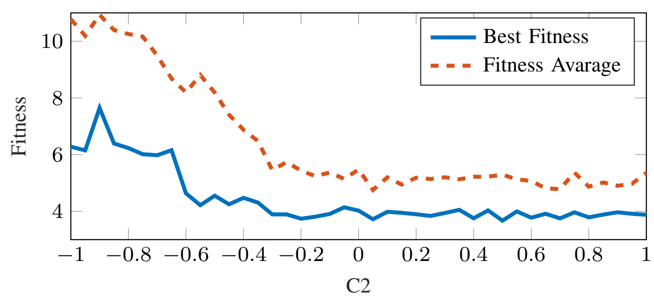

Fig. 7. Variation of C2

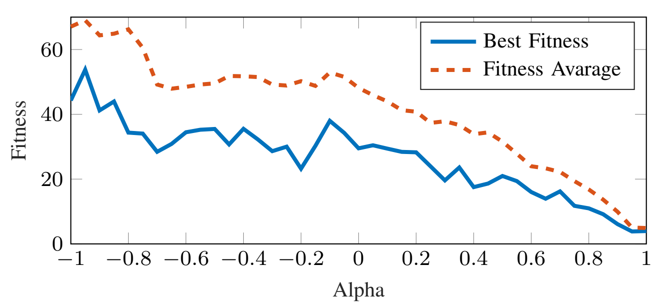

Fig. 8. Variation of Alpha

The number of particles was the fifth parameter to be adjusted. However, as this is an integer, was varied according to the interval submitted by 18. The number of particles selected was 25, so that the average value of *fitness* stabilized when the number of particles obtained this value, and by the assumption that this parameter is critical for the algorithm, that is, the higher the number of particles, the greater the cost and the computational time. The last parameter to be tested was the amount of iterations. The value selected for this parameter was 300 iterations, since the algorithm began to stabilize when reached this value.

$$[5, a * 5], \text{ where } 2 \leq a \leq 10, \; a \in Z_+ \tag{18}$$

The tests in this section had the purpose to find out the best values for each of the parameters in order to optimize the next tests in section IV-B. The best results of these parameters are presented in table V.

TABLE V
SUMMARY OF INITIAL TESTING OF THE ALGORITHM

Inertia (w)	$C1$	$C2$	**Alpha** (α)	**Particles**	**Iterations**
-0.30	1	0.05	1	25	300

B. Experiments

In the study presented in this section of the performance of the PSO for groups with six pupils was compared with the random algorithm. Comparisons are made using the average time of execution, the average of the *fitness* and best *fitness* found (*bestFitness*) for each dimension. Test scenarios on the PSO were run 30 times using 25 particles, while the random method was executed according to Equation 19.

$$(N_{part} * Nmax_{It} + N_{part}) * 30 \tag{19}$$

Where N_{part} is the number of particles and $Nmax_{It}$ is the maximum number of iterations, components of the equation 19, were configured from the initial tests presented in section IV-A and have as values 25 and 300, respectively. As the normal random algorithm test case generates many infactable solutions. The same repair function used in the PSO was executed by him for the amount of students in the Group were respected, leading in a runtime and larger computational cost.

In table VI are shown the tests carried out using the eleven sets of data broken down in section IV-A, for each set of data the PSO was executed with 300 iterations, in addition to the presentation of the results of the random algorithm.

The dimensions were divided into three groups, low, medium scale scale and large scale. To the problems of low scale, or smaller ones, like for example the 12 and 24 dimension, the random algorithm obtained values worse than when we look to the PSO *fitness* average, however best values and very close when observed values of *bestfitness*.

In other small-scale problems, 48 until 768, notes that the average *fitness* is considerably larger for the random algorithm, as well as some of the best value of *fitness* (*bestFitness*) found from all executions performed by the random algorithm. Figure 9 demonstrates graphically the benefit of PSO on the random algorithm.

In addition to the comparison of the PSO algorithm and random, through the Table VI shows that in smaller problems, called low scale, up to 768 dimension, the runtime of the algorithm random was far superior to the PSO algorithm, that due to Organization function that prevents and reorganizing groups with fewer than six students and seven. Without this

TABLE VI
EXPERIMENTAL RESULTS OF THE ALGORITHMS

	PSO			Random		
Dimensions	**Average Time**	**Average Fitness**	**BestFitness**	**Average Time**	**Average Fitness**	**BestFitness**
12	3,161E-02	3,693E+00	3,167E+00	5,764E-01	3,883E+00	3,067E+00
24	3,665E-02	4,200E+00	3,217E+00	1,148E+00	5,855E+00	3,117E+00
48	8,897E-02	4,084E+00	2,964E+00	2,329E+00	8,336E+00	3,680E+00
96	2,248E-01	4,305E+00	3,564E+00	5,390E+00	2,058E+01	6,307E+00
192	6,022E-01	4,106E+00	3,384E+00	1,416E+01	1,330E+02	2,954E+01
384	1,748E+00	4,643E+00	3,847E+00	4,152E+01	6,356E+03	4,686E+02
768	5,607E+00	1,219E+01	6,243E+00	1,356E+02	1,784E+07	9,929E+05
1536	1,965E+01	2,177E+03	5,203E+02	4,827E+02	1,885E+14	2,090E+12
3072	6,049E+01	1,770E+11	2,356E+09	1,814E+03	3,175E+28	2,000E+25
6144	2,313E+02	1,109E+30	5,008E+28	7,048E+03	1,599E+57	1,041E+53
9216	5,206E+02	5,379E+50	5,465E+48	1,571E+04	1,111E+86	1,603E+81

function, the random algorithm would be smaller, but would generate many infactíveis groups.

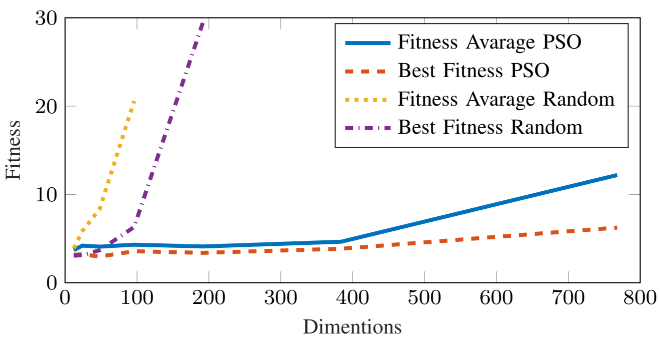

Fig. 9. Average results of fitness and bestFitness-Low range

In medium scale problems, ranging from 1536 and 3072 dimension, it is observed that the average time continues to increase considerably, where, for example, the average time of execution of the 1536 dimension more than tripled compared to the previous dimension, the same occurred for the size 3072, for both algorithms. However, we can observe that the PSO algorithm gets much better values than the random algorithm. The gain that the PSO algorithm has about the random algorithm can be seen with more clarity, for example, when we observe the values of the averages of the *fitness*.

For large scale problems, which are the dimensions 6144 and 9216, one realizes that the average time continued to increase, but the rate of increase between these two dimensions decreased when compared with the rate of increase in the time between the size 3072 and 6144. For the random algorithm, this rate has increased. The *fitness* value has increased considerably for both algorithms. However, the random algorithm continued to get much higher values to the PSO.

Observing the values of the averages of *fitness* and runtime, it can be concluded that the PSO algorithm gets better results in all datasets used for the analysis. Some *fitness* values may appear high due to the penalty imposed on the function algorithm to avoid infactíveis solutions (explained in section III-B2), but can represent a major advance in the formation of groups for collaborative learning in MOOCs using three grouping criteria.

V. Conclusions and Future Work

In this work a method to group formation in MOOCs, based on the criteria of level of knowledge and interest of each student and distributed leadership concepts, using the PSO algorithm. The proposed method allows the formation of groups from a large number of students involved in the context of providing a MOOC improvement in the quality of interactions and advancement of the proven knowledge through a case study.

Thus, the proposed method can be of great value in *online* environments with a large number of participants where the teacher can't meet everyone and also when students are geographically far apart, making it impossible for the group study attendance. The method consists of three clearly defined phases and can be used in the context of the MOOCs as well as in computer supported learning environments, such as long distance courses.

The model of formation of groups for collaborative learning in MOOCs uses the *Particle Swarm Optimization algorithm* as method for the composition of the groups, allowing the automatic creation of the groups, given the huge volume of students in a MOOC. Were created an algorithm form groups with six members, with different levels of knowledge, interests and with concepts of distributed leadership. The teacher can not only perform work according to the criteria mentioned, but also add other criteria for adequate training of their groups adapting the *fitness* function of the algorithm according to your goals pedagogic.

At the end of the experiments, the algorithm has demonstrated that it can meet the criteria for grouping in a computation time and being more efficient than random groups. The tests also showed that the algorithm is robust considering the various data sets and variations of iterations. However, despite the algorithm does not guarantee that the end result is the best solution, due to the computation time be limited, this research contributes towards a method that turns out better than the commonly used, which is the formation of random groups.

As future work, computational point of view, we want to compare the PSO algorithm with other techniques of composition of groups, such as the genetic algorithm. The comparison of algorithms can help define what the circumstances are more favorable for use of each algorithm, e.g., number of criteria and number of students. From the pedagogical point of view, we intend to analyze the content of the discussions with other methods of analysis of speech, but also compare the groups individually and compare the speeches made by each leadership profile for perspective more thorough about the quality of the interactions. We want to use this method in a MOOC of a larger scale, which has a greater diversity in order to have a broader view on its results.

REFERENCES

[1] A. McAuley, B. Stewart, G. Siemens, and D. Cormier, "The mooc model for digital practice, sshrc knowledge synthesis grant on the digital economy," *Recuperado de http://www. edukwest. com*, 2010.

[2] C.-H. Tu, C.-J. Yen, L. Sujo-Montes, and K. Sealander, "Digital lifelong-learning literacy," *The wiley handbook of educational policy*, pp. 531–550, 2018.

[3] G. Kennedy, C. Coffrin, P. de Barba, and L. Corrin, "Predicting success: how learners' prior knowledge, skills and activities predict mooc performance," in *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge*. ACM, 2015, pp. 136–140.

[4] A. Anderson, D. Huttenlocher, J. Kleinberg, and J. Leskovec, "Engaging with massive online courses," in *Proceedings of the 23rd international conference on World wide web*. ACM, 2014, pp. 687–698.

[5] A. Harris, M. Jones, and S. Baba, "Distributed leadership and digital collaborative learning: A synergistic relationship?" *British Journal of Educational Technology*, vol. 44, no. 6, pp. 926–939, 2013. [Online]. Available: http://dx.doi.org/10.1111/bjet.12107

[6] M. K. Chandrasekaran, M.-Y. Kan, B. C. Tan, and K. Ragupathi, "Learning instructor intervention from mooc forums: Early results and issues," *arXiv preprint arXiv:1504.07206*, 2015.

[7] L. V. Morris, "Moocs, emerging technologies, and quality," *Innovative Higher Education*, vol. 38, no. 4, pp. 251–252, 2013.

[8] A. C. Robinson, J. Kerski, E. C. Long, H. Luo, D. DiBiase, and A. Lee, "Maps and the geospatial revolution: teaching a massive open online course (mooc) in geography," *Journal of Geography in Higher Education*, vol. 39, no. 1, pp. 65–82, 2015.

[9] S. K. Ch and S. Popuri, "Impact of online education: A study on online learning platforms and edx," in *MOOC Innovation and Technology in Education (MITE), 2013 IEEE International Conference in*. IEEE, 2013, pp. 366–370.

[10] I. Claros and R. C. Leovy Echeverria, "Towards moocs scenaries based on collaborative learning approaches," in *Global Engineering Education Conference (EDUCON)*, 2015, pp. 955–958.

[11] W.-N. Chen, J. Zhang, H. Chung, W.-L. Zhong, W. gang Wu, and Y. hui Shi, "A novel set-based particle swarm optimization method for discrete optimization problems," *Evolutionary Computation, IEEE Transactions on*, vol. 14, no. 2, pp. 278–300, April 2010.

[12] M.-I. Dascalu, C.-N. Bodea, and A. Burlacu, "Platform for creating collaborative e-learning communities based on automated composition of learning groups," in *Engineering of Computer Based Systems (ECBS-EERC), 2013 3rd Eastern European Regional Conference on the*, Aug 2013, pp. 103–112.

[13] M.-I. Dascalu, C.-N. Bodea, M. Lytras, P. O. De Pablos, and A. Burlacu, "Improving e-learning communities through optimal composition of multidisciplinary learning groups," *Computers in Human Behavior*, vol. 30, pp. 362–371, 2014.

[14] B. Jarboui, N. Damak, P. Siarry, and A. Rebai, "A combinatorial particle swarm optimization for solving multi-mode resource-constrained project scheduling problems," *Applied Mathematics and Computation*, vol. 195, no. 1, pp. 299–308, 2008.

[15] B. Jarboui, S. Ibrahim, P. Siarry, and A. Rebai, "A combinatorial particle swarm optimisation for solving permutation flowshop problems," *Computers & Industrial Engineering*, vol. 54, no. 3, pp. 526–538, 2008.

[16] Y.-T. Lin, Y.-M. Huang, and S.-C. Cheng, "An automatic group composition system for composing collaborative learning groups using enhanced particle swarm optimization," *Computers & Education*, vol. 55, no. 4, pp. 1483–1493, 2010.

[17] J. L. Pierobom, M. R. Delgado, and C. A. Kaestner, "Particle swarm optimization applied to the dynamic allocation problem," in *Neural Networks (SBRN), 2012 Brazilian Symposium on*. IEEE, 2012, pp. 184–189.

[18] M. Rosendo and A. Pozo, "Applying a discrete particle swarm optimization algorithm to combinatorial problems," in *Neural Networks (SBRN), 2010 Eleventh Brazilian Symposium on*, Oct 2010, pp. 235–240.

[19] S.-Q. Wang, L.-H. Gong, and S.-L. Yan, "The allocation optimization of project human resource based on particle swarm optimization algorithm," in *Services Science, Management and Engineering, 2009. SSME '09. IITA International Conference on*, July 2009, pp. 169–172.

[20] N. Dordaie and N. J. Navimipour, "A hybrid particle swarm optimization and hill climbing algorithm for task scheduling in the cloud environments," *ICT Express*, vol. 4, no. 4, pp. 199–202, 2018.

[21] M. Marichelvam, M. Geetha, and Ö. Tosun, "An improved particle swarm optimization algorithm to solve hybrid flowshop scheduling problems with the effect of human factors–a case study," *Computers & Operations Research*, vol. 114, p. 104812, 2020.

[22] I. Amarasinghe, D. Hernández-Leo, and A. Jonsson, "Data-informed design parameters for adaptive collaborative scripting in across-spaces learning situations," *User Modeling and User-Adapted Interaction*, vol. 29, no. 4, pp. 869–892, 2019.

[23] N. M. Webb, "Predicting learning from student interaction: Defining the interaction variables," *Educational psychologist*, vol. 18, no. 1, 1983.

[24] ——, "Testing a theoretical model of student interaction and learning in small groups," *Interaction in cooperative groups: The theoretical anatomy of group learning*, vol. 102, p. 119, 1992.

[25] L. Linnenbrink-Garcia, K. J. Pugh, K. L. Koskey, and V. C. Stewart, "Developing conceptual understanding of natural selection: The role of interest, efficacy, and basic prior knowledge," *The Journal of Experimental Education*, vol. 80, no. 1, pp. 45–68, 2012.

[26] P. N. Karamolegkos, C. Z. Patrikakis, N. D. Doulamis, P. T. Vlacheas, and I. G. Nikolakopoulos, "An evaluation study of clustering algorithms in the scope of user communities assessment," *Computers & Mathematics with Applications*, vol. 58, no. 8, pp. 1498–1519, 2009.

[27] S. J. Yang, "Context aware ubiquitous learning environments for peer-to-peer collaborative learning," *Journal of Educational Technology & Society*, vol. 9, no. 1, pp. 188–201, 2006.

[28] R. J. S. Lourenço, "Um estudo exploratório na faculdade de engenharia da universidade do porto sobre a materialização da liderança entre os diferentes agentes," 2012.

[29] E. Ossiannilsson, "Leadership in global open, online, and distance learning," in *Online Course Management: Concepts, Methodologies, Tools, and Applications*. IGI Global, 2018, pp. 2212–2240.

[30] P. G. Northouse, *Introduction to leadership: Concepts and practice*. Sage Publications, 2014.

[31] J. Gressick and S. J. Derry, "Distributed leadership in online groups," *International Journal of Computer-Supported Collaborative Learning*, vol. 5, no. 2, pp. 211–236, 2010.

[32] J. Kennedy and R. Eberhart, "Particle swarm optimization, ieee international of first conference on neural networks," 1995.

[33] S. Sengupta, S. Basak, and R. A. Peters, "Particle swarm optimization: A survey of historical and recent developments with hybridization perspectives," *Machine Learning and Knowledge Extraction*, vol. 1, no. 1, pp. 157–191, 2019.

[34] J. Kennedy and R. Eberhart, "A discrete binary version of the particle swarm algorithm," in *Systems, Man, and Cybernetics, 1997. Computational Cybernetics and Simulation., 1997 IEEE International Conference on*, vol. 5, Oct 1997, pp. 4104–4108 vol.5.

[35] M. Salehi and H. Rezaei, "A fuzzy multi-objective model for allocating orders to suppliers under shortfall and price-quantity discounts: An mpso and nsga-ii with tuned parameters," *International Journal of Industrial Engineering & Production Research*, vol. 30, no. 2, pp. 225–239, 2019.

[36] C. A. Peckens and C. Fogg, "Bio-inspired iterative learning technique for more effective control of civil infrastructure," in *Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 2019*, vol. 10970. International Society for Optics and Photonics, 2019.

[37] J. Al-Sawwa and S. A. Ludwig, "Centroid-based particle swarm optimization variant for data classification," in *2018 IEEE Symposium Series on Computational Intelligence (SSCI)*. IEEE, 2018, pp. 672–679.

[38] M. Li, S. Lian, F. Wang, Y. Zhou, B. Chen, L. Guan, and Y. Wu, "Neural network modeling based double-population chaotic accelerated particle swarm optimization and diffusion theory for solubility prediction," *Chemical Engineering Research and Design*, 2020.