VOL. 15, NO. 24, DECEMBER 2020 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

DYNAMIC GROUP FORMATION IN COMPUTER SUPPORTED COLLABORATIVE LEARNING

Ihtesham Hussain[1], Kashif Sultan[2] and Iqra Basharat[2]
[1]Department of Software Engineering, Riphah International University, Islamabad, Pakistan
[2]Department of Software Engineering, Bahria University, Islamabad, Pakistan
E-Mail: kashif.buic@bahria.edu.pk

ABSTRACT

Computer-Supported Collaborative Learning (CSCL) emerges as an instrument of learning and training that can encourage the social nature of learning by adopting range of computer-mediated communication tools and pedagogical methods. These tools are used to facilitate the learning and instructional communication among students and learners in small groups. In this process, group formation becomes painstaking and challenging task. Various factors are involved affecting group formation that includes; personal characteristics, social, cultural, psychological and cognitive diversity. Although, this issue was addressed in various research studies yet an optimal solution for dynamic group formation is not discussed evidently. In dynamic groups, students work or collaborate on a short-term task in a group that change frequently based on the performance of students. In this research study we have proposed a method for dynamic group formation and the impact of dynamic group formation in collaborative learning among peers is demonstrated by conducting an experiment. This experiment is conducted in two phases. In first phase of the experiment, learning styles are assigned to the students and their knowledge level is calculated. Whereas, in second phase of the experiment, the impact of dynamic grouping on collaborative learning of students is determined. Further, two algorithms are proposed, first is used for determining initial number of clusters and second algorithm is used for dynamic grouping after the completion of each permutation. Results of our experiment shows a positive impact of dynamic grouping on learning of students as the performance of collaborative learning among peers is better than individual performances.

Keywords: collaborative learning (CL), computer supported collaborative learning (CSCL).

1. INTRODUCTION

Collaborative learning is an area that has been under discussion for several years. It became prominent due to the support of the technologies and the internet. Collaborative learning is a pedagogical method in which students collaborate to learn and share their experiences while solving the problem [1]. In conventional learning, collaborative learning (CL) is used for group activities and projects. Mentors create groups manually, assign them tasks, students then collaborate and try to solve the assigned tasks mutually. Groups created manually by mentors are not of the same capability, some groups perform better while others are unable to perform the tasks. Such group formation affects the learning of students. The advent of computers and the internet empowered collaborative learning. Students are given scenarios and tasks, which they solve by using the computer, supported shared resources e.g. shared whiteboard, a shared tool for creating diagrams or drawings, etc. [2].

The prime and most important task in CSCL is group formation [1]. Group formation is dividing students into different clusters and assigning them some sort of tasks. They collaborate and complete those tasks. They are clustered into groups based on some attributes like their knowledge level, behavior, learning styles, and interests. From literature, it is evident that there are 3 main ways of group formation [3].

- Random Selection
- Self-selection
- Instructor Selection

In random selection, any computer program or teacher swaps students randomly and create groups out of that, to get heterogeneous groups. Heterogeneous groups have students of different types e.g. different knowledge levels and different personality characteristics. On the other hand, in homogenous groups, all the students are of the same level and capacity. In [4], authors proved that heterogeneous groups perform better than homogenous groups. The self-selection method of group formation allows students to choose their group partners. This type of group formation produces better results but on the other hand, it has a consequence that most of the time students of the same capacity and level make groups, which results in almost all weak students in one group and almost all good students in another group. This type of group formation is also not based on the pedagogical grouping method; it is based on the personal friendship of empathy. So, unbalanced groups are achieved instead of good performing groups [3]. The third way of group formation is instructor-selection method; instructors are allowed to create groups. This is hectic for the instructor to judge and analyze every student's performance and characteristics, so, that he can arrange them in groups. In this method, the instructor can create heterogeneous or homogenous groups. In previous studies, authors have used different attributes for group formation. A proper consensus among authors does not exist, that which attributes are playing a vital role in group formation. In our study, we have adapted the knowledge level as a prime attribute for group

formation. Along with the knowledge level, we are also finding out the learning styles of each student and finding its impact on group formation. These learning styles are adapted from the study of Richard M. Felder [5], where they proposed learning styles for engineering education. The first phase of the experiment of our study is designed based upon the questionnaire of learning styles proposed by Richard M. Felder [5]. Following are the learning styles we are using in this study.

- Sensory
- Intuitive
- Active
- Reflective
- Visual
- Verbal
- Global
- Sequential

Sensory and intuitive refers to the information perception of the students. Do students get information from the surroundings using 6 senses or he/she is getting the information from the intuition (internal) [5]?

Visual/verbal is the mode of channel through which students get the information. Visual means students are getting information in the form of pictures and videos, while verbal means getting information in verbal form or textual form.

Active and reflective are the two terminologies which the author [5] has used for the processing of information. These terms (Active and reflective) specifies how students are processing the information he/she is receiving. They may actively process it by talking about it or practicing it. They may process it intuitively which is also called introspective processing of information. Introspectively processing the information is termed as reflective learning.

Then there comes other terms called global or sequential students. Global students learn by making large jumps, they learn with the abstract eye, while sequential students go in the depth of the information and try to understand each detail [5].

1.1 Research Objectives

In this study our research objectives are:

- How to create dynamic groups in CSCL?
- How to overcome cold start i.e., how to create initial groups?
- Can dynamic group formation enhance the group performance and student learning?

2. RELATED WORK

The Concept of student space was used by Rahel *et al* [2]. They defined the student space as the group of attributes like work attitude, interest for subject, self-confidence, shyness, etc. The values of these attributes were obtained from easily available indicators like expert opinion and discussion with colleagues. They have proposed a mathematical model where heterogeneity of the groups is calculated using the Euclidean distance. They implemented their mathematical model using ant colony optimization. Christodoulopoulos *et al* were using a fuzzy C-mean algorithm for Homogenous and heterogeneous group formation [6].

IHUCOFS (Integrated human coalition formation and scaffolding) framework is proposed by Soh *et al* [3]. They designed an algorithm called VALCAM as a preliminary implementation of iHUCOFS. VALCAM contains the system agents which were assigned to the human agents. System agents hold an auction and the user agents bid in the auction with the virtual currency they have earned from participating in the previous coalition. The VALCAM is based on certain rules. The Semantic group formation framework is introduced by Ounnas *et al* [4]. Students were asked to enter their data like a list of friends and their preferences. The Authors created an ontology called semantic learner profile. When students submitted their data, it was stored in the RDF file which was then processed using Jena. Their framework had a teacher interface that allows the teacher to select the constraint they care about the most. The authors were using the DLV solver for creating the group based on the constraints set by the instructor.

Ho *et al.* have used particle swarm optimization for the heterogeneous group formation [7]. Their group formation was based on the competence, learning style, and interaction among the student. Neil Rubens *et al* have proposed a group formation method for informal collaborative learning [8]. They collected data from different sources like blogs, social media, and other databases. They created a mash of data and then used data mining to make groups.

Hubscher, R. created groups using the tabu search after the instructor set the preferences [9]. Yannibelli *et al* have worked on the group formation for the software engineering courses. Grouping criteria were based on the Belbin team roles. Students were divided into different groups, which were then decoded and evaluated using the fitness function. The fitness function evaluated the groups to obtain the optimization objective. The optimization objective was to generate the maximum number of balanced groups [10].

Abnar *et al.* have proposed a new method for group formation using the genetic algorithm [11]. Teachers were asked to set different attributes about the students and rate them. Teachers were then displayed with the graph showing the distance. The algorithm was then run to create different groups that were presented before the teacher for acceptance or rejection. Brauer *et al.* have used the social network called diaspora for the research. Users/learners select the topic, potential candidates are then found on the network to form groups [12]. Moreno *et al* have used a genetic algorithm to make inter-homogenous groups. They encode the different attributes of the learners and creates a matrix out of it. Then they apply selection, crossover, and mutation to make groups [13]. Mujkanovic *et al* have worked on the group

formation for the remote laboratories' access, where a student can access the laboratories and solve the lab manual in the group. The authors were using the regression analysis to make groups. Metadata of students were given to the algorithm, then group formation was done based on the rules set by the admin. The algorithm was learning from the student performance to update the rules [14]. Tien *et al* have proposed the concept of TOPSIS (Technique for order performance by similarity to ideal solution). The main steps of the technique were pre-categorization, encoding, initial population, fitness function and if the termination criteria were met then grouping results were shown otherwise selection, crossover, mutation, and elitism were performed to calculate the fitness function again [15]. Ivan Srba *et al.* used and developed an application called popCorm, in which they experimented with the dynamic short term group formation for the online environment. They used the Group technology approach for dynamic grouping. This is the technique used in manufacturing and engineering to find similarities among the products. Input to the method was 2 clusters. The cluster of characteristics and the cluster of assignments were used as input. Data for the input matrix were gathered from questionnaires and external sources. The result of the technique is homogeneous groups [16]. Zhilin Zheng *et al* used the discrete particle swarm optimization for the composition of groups using the gender and the MBTI personality as attributes. They also compared the DPSO with the competing method and the random method and according to them; DPSO performs better [17]. Ullmann *et al* used particle swarm optimization to form groups using the MOOCs platform. Their group formation was based on the knowledge level and interests [18]. YR Chen *et al* have proposed a method for group formation, in which they have used the edX platform for experimentation. Teachers assign a task to the student. Students solve the task and do a discussion on the discussion board. The system fetches the information of interaction on the discussion board and fetches the grades of the students and display the group's list [19]. Hamid Sadeghi *et al* used the undirected weighted graph to model the online e-learning platform. In the graph learners are a node and the relation between them shows the similarity of their interest. They have used the questionnaire to collect data about the students` interests. The similarity was measured by taking a mean of their absolute interest. The graph was also shown in the form of the asymmetric adjacency matrix. The binary integer programming model is introduced to assign students to their respective groups based on their interests [20]. C Yin *et al* have proposed a model called GFS. They were clustering the students into groups based on gender, major, reading pages, reading time, attendance, and content. They were using the educational log and data from Moodle for their research. The teacher had to set the attributes for group formation, the algorithm then makes and display groups created [21]. Y Zheng *et al* used a genetic algorithm for group formation [22]. Bhardwaj et al have used a test-based approach called DISC for the compatibility of employees working in an organization. They divide employees into four type:

- Dominant
- Influence
- Steady
- Compliance

Further, they divide them into two categories called active and passive. They created a matrix where they put values of different attributes of personalities of employees. Based on those values Euclidean distance is calculated if the Euclidean distance is less than the cutoff distance then the compatibility is 1 vice versa [23]. D Jagadish has used the KNN algorithm for grouping. They have used moodle for experimentation [24]. Maina *et al* have worked on group formation. They have used the means and EM clustering algorithms. Group formation was based on the log data of the discussion forum of moodle. Their log file contained several posts, user id, number of replies, and forum ratings. Authors have proposed the method where they get data from the log file and create clusters out of that. Groups were then created based on the cluster formed. Members from each cluster were assigned to the groups based on the high competence level [25]. Yu-Chen Kuo *et al.* performed collaborative learning experiments on students of the English language who study English as a foreign language. The authors performed experiments on three different types of groups. The first form of the group was generated randomly. Kolb's learning styles were used to generate the second and third types of groups. The second group was homogeneously containing students having the same learning styles and thirds group was heterogeneous which had students with different learning styles [26]. Cícero *et al.* generated groups randomly for collaborative activities. Evaluation of group collaborations is carried out using linear regression. Variables for linear regression were self-esteem, and self-efficacy, which was extracted from the student, self-reports [27]. Jigsaw group formation was carried out by Ishari *et al*. Where tasks were assigned to each student, then groups were created in the form of the jigsaw. For example, if there were 5 tasks and 20 students then these tasks are distributed among the 20 students. Groups were created in a way that each group would have 5 students with a different type of questions; each group would have exactly one representation of each task [28]. Ivica at al. Worked on the content independent collaborative learning. They experimented with the 37 school students. Students were given a choice to choose their groups for collaborative learning based upon their needs [29].

The behavior of penguins is mimicked by Zedadra *et al* as they proposed an algorithm based on the natural phenomenon of penguins. They performed dynamic group formation in CSCL by proposing this new approach. Initially, learners used the LMS system where their profile traces are collected which means their profiles are learned by the system. Based on the traces of the

profile, learners are grouped randomly in homogeneous groups. Groups are updated regularly using their dynamic grouping algorithm which is mimicking the natural behavior of penguins [30]. Another approach for dynamic group formation was proposed where students were given individual tasks and based on their performance evaluation; they were placed in different groups which the author called pots. Groups were created with equal participation from each pot and were being updated regularly based on the students' performance. The drawback of this method was that it was not taking into account the negative effect of students' swapping from one group to another group, that degrade group performance.

For example, if one student from group A is swapped to group B, the performance of group B may increase but the performance of group A may drop. Swapping should only happen if it does not affect the performance of the previous group [31].

In literature most of the research studies are doing group formation for collaborative learning in different scenarios. Some studies are using different methods like genetic algorithms [11] [13] [15] [22], particle swarm optimization [2] [7] [17] [18] etc. for group formation but maximum are using random group formation [32] [26] [8] [12] or data mining techniques like K-means [6] [8] [16] [24] [25]. Each study in the literature creates groups based on different attributes like learning styles (Kolbe's or Felder), educational logs, knowledge level, social media or other online logs about students (e.g. blogs, websites).

Therefore, based on a thorough understanding of literature we came up with the hypothesis of dynamic grouping, which means that groups will be updated after the completion of each activity based upon the performance of different groups so that we can get the balanced groups. The subsequent section deliberates our research methodology in detail.

3. RESEARCH DESIGN

In this section, we have discussed our proposed framework. The block diagram in Figure-1 represents the proposed framework. This framework comprises the following blocks.

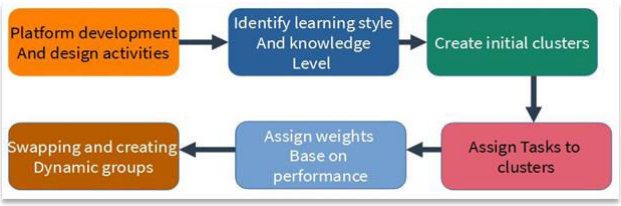

Figure-1. Research design block diagram.

Our methodology is evident from literature [32] [26], most of the researchers used the same methodology to verify their proposed technique for collaborative learning. We proposed our custom algorithm for dynamic grouping in our experimentation instead of using the predefined algorithms because we could not find an algorithm that can be used for dynamic grouping.

4. EXPERIMENT PLATFORM

Our experiment comprised of two phases. In the first phase of the experiment, learning styles are assigned to students and their knowledge level is calculated. In the second phase, based on the data received from phase 1, initial clusters are created. Dynamic groups are then created from the initial clusters based on the performance of students.

To conduct our experiments, we have designed a web-based application. This application is developed using PHP (Laravel framework). MySQL database is used to store the data of the application.

We performed a controlled experiment under the supervision of an instructor. For experimentation purposes, we chose 19 students of the Object-Oriented Programming course as participants for post-test, who belonged to the second semester (Fall 2019) of their degree program and had only taken the programming fundamentals course for beginners in their first semester. They were only able to do structural and procedural programming in C++ using the basic construct of programming such as variables, loops, decisions, arrays, and structures.

They knew each other socially for 6 months. That was a positive point for our successful collaborative learning. They were willing to work in groups and help each other. The activities involved in the experiment were also considered as their regular class assignments, for which they were to be evaluated accordingly. Therefore, they had to take all the tasks seriously and work hard to score well.

They were encouraged to participate in group learning and 5 bonus marks were also announced to be given to the best-performing groups. In this way, they were given an extrinsic motivation as performance-contingent rewards to improve motivation and performance [33].

4.1 Experiment Phase I

The purpose of this phase of the experiment is to identify the learning styles of students how one student loves to learn and what are his/her personality traits towards learning. Our second purpose of the experiment is to calculate the knowledge level of students. This first phase of the experiment is comprised of certain steps, in each step students perform certain activities, based on those activities, learning styles, and knowledge level of students is identified and calculated. In step 1 of the first phase of the experiment, students are presented with the application where they have to create their accounts. After registering on the system, they log in to the system using their email and password. In the 2nd step of phase 1, students are presented with the interface where students have to select the type of content they want to learn from. After selecting the content type (videos, text file or presentation slides) they will be presented with content/lessons in that format based on their selection.

Once, students select the type of content they are migrated to another interface where they see lessons. Students can also see the "go-to quiz" option and "perform practical" option. If a student clicks on the "perform practical" button they are redirected to some external resources for practice. This link of practice is set by the instructor while designing the activity. But if student choose to click on the "go-to quiz" option then they are redirected to another interface where they solve the quiz. After they submit a quiz, they are displayed with another image where students have to name all the objects, they see in the picture to find out whether they are sensory or intuitive.

4.2 Identify Learning Styles and Calculate Knowledge Level

M. Felder's learning styles are adopted in this study [5].

Following are the list of attributes, values of which are extracted in experiment phase I. 0 or 1 is set to these attributes against each student.

Visual, verbal, sensory, intuitive, active, reflective, global, and sequential.

Algorithm 1 Initial Group Formation

```
studentData ← getDataFromDataBase()
for studentData do
    total ← calculateKnowledgeLevel(student[i])
end for
mean ← total/counter
cluster1[]
cluster2[]
for data do
    if studentData[i]− > KnowledgeLevel() < mean
    then
        cluster1 ← studentData[i]
    else
        cluster2 ← studentData[i]
    end if
end for
totalGroups ← 5
for cluster1 do
    Assign group number to students
end for
for cluster2 do
    Assign group number to students
end for
```

Figure-2. Initial Group Formation Algorithm.

4.3 Experiment Phase II

In the second phase of the experiment, initial groups are created using our proposed algorithm 1. These groups are created from the data collected from phase I of the experiment. Initial group formation algorithm is very simple; we get the knowledge level of each student and calculate the mean. Students with the knowledge level value less than the mean are placed in one cluster and students with knowledge level value greater than the mean are placed in another cluster.

After initial groups are created. Activities are assigned to each student, which they solve individually. After tracking the individual performance of each student, they are assigned activities in the form of groups. These activities are designed by the instructors. 6 activities are designed in this study, each has 5 MCQs. Students can chat with each other to build consensus on the common answer and submit a collective answer.

We ran 6 permutations, each time student`s groups are swapped using our proposed algorithm for dynamic grouping. In a dynamic group formation algorithm, each group's points are calculated. Then we calculate the mean value of the points of groups. Two clusters are created, one called as smallPoints cluster where groups with smallPoints are saved and another is called as greaterPoints cluster where groups with greaterPoints are saved. This categorization that takes place is based on the mean value; if the points are less than the mean then it is saved in smallPoints group and vice versa.

Now, smallPoints cluster is sorted in ascending order and greaterPoints cluster is sorted in descending order. After the sorting, students of both the clusters are swapped in such a way that weak students become the part of good performing groups and average or the student who is the second-best performing in the good performing groups are swapped to the low-performance group to improve their performance. This swapping takes place keeping in view that the performance of the previous group is not affected.

5. RESULT AND EVALUATION

In this section, we present experimental results. Students performed individually and then in groups so that we can measure their improvement while studying in a group. The individual performance of students is shown in Figure-4.

It is obvious from the Figure-4 that there is a lot of fluctuation of points. Some students are performing well, and some are performing poorly. Now, we have results based on the individual performance of students, it is time to hunt for the group performance result and analyze the impact of dynamic grouping upon the learning experience of students.

As it is mentioned in experiment phase II that 6 activities (each having 5 MCQs) are designed and 6 permutations are executed with different groups for each activity using our proposed algorithm shown in Figure-3.

ARPN Journal of Engineering and Applied Sciences

Algorithm 2 Dynamic Grouping Algorithm

```
groupPoint ← getGroupPoints
mean ← calculateMean(groupPoint)
greaterGroupPoint[]
smallerGroupPoint[]
for groupPoint do
  if groupPoint−>point < mean then
    smallerGroupPoint ← groupPoint
  else
    greaterPointGroup ← groupPoint
  end if
end for
smallerSorted ← AssendingSort(smallerPointGroup)

largeSort ← descendingSort(greaterPointGroup)
toBeSwapped[]
while smallerSorted do
  smallStudent ← smallerSorted[i]−>
  getStudents()
  smallStudent ← descendingSort()
  for smallStudent do
    toBeSwapped ← smallStudent
  end for
end while
toBeSwappedWith
while largeSort do
  LargeStudent ← largeSort[i]−>getStudents()
  largeStudent ← AscendingSort()
  for largeStudent do
    toBeSwappedWith ← largeStudent
  end for
end while
```

Figure-3. Dynamic grouping algorithm.

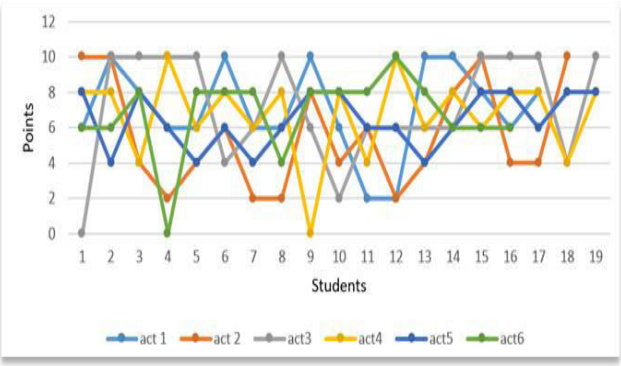

Figure-4. Individual student performance.

Table-2 to Table-7 represents the group formation for different activities. Text in the bold represents the swaps that happened.

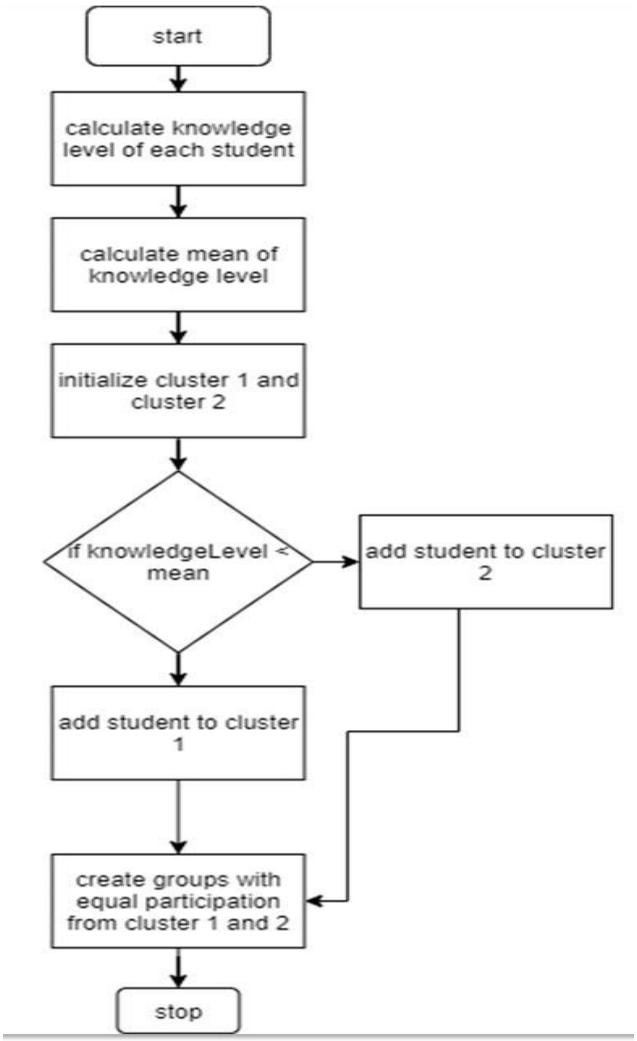

Figure-5. Detail working of initial groping algorithm.

Table-1. Activity 1 group combination.

Group 1	Group 2	Group 3	Group 4	Group 5
Student 6	Student 18	Student 12	Student 10	Student 14
Student 1	Student 8	Student 5	Student 16	Student 4
Student 17	Student 15	Student 1	Student 13	Student 19
Student 3	Student 11	Student 2	Student 7	

Figure-6. Detail working of dynamic grouping algorithm.

Table-2. Activity 2 group combination.

Group 1	Group 2	Group 3	Group 4	Group 5
Student13	Student9	Student 12	Student 6	Student 4
Student 3	Student 15	Student 2	Student 10	Student 5
Student 1	Student 18	Student 19	Student 16	Student 14
Student 7	Student 11	Student 8	Student 17	

Table-3. Activity 3 group combination.

Group 1	Group 2	Group 3	Group 4	Group 5
Student 1	Student 11	Student 8	Student 6	Student 16
Student 3	Student 15	Student 19	Student 17	Student 10
Student 7	Student 18	Student 2	Student 14	Student 5
Student 13	Student 12	Student 9	Student 4	

Table-4. Activity 4 group combination.

Group 1	Group 2	Group 3	Group 4	Group 5
Student 3	Student 18	Student 8	Student 14	Student 10
Student 13	Student 15	Student 19	Student 4	Student 5
Student 7	Student 9	Student 1	Student 17	Student 16
Student 11	Student 12	Student 2	Student 6	

Table-5. Activity 5 group combination.

Group 1	Group 2	Group 3	Group 4	Group 5
Student 6	Student 18	Student 8	Student 3	Student 5
Student 10	Student 15	Student 19	Student 14	Student 16
Student 13	Student 9	Student 1	Student 4	Student 7
Student 11	Student 12	Student 2	Student 17	

Table-6. Activity 6 group combination.

Group 1	Group 2	Group 3	Group 4	Group 5
Student 6	Student 18	Student 8	Student 3	Student 5
Student 10	Student 15	Student 19	Student 14	Student 16
Student 13	Student 9	Student 1	Student 4	Student 7
Student 11	Student 12	Student 2	Student 17	

On successful execution of each activity, the swapping of students among groups took place, but there is no swapping after the completion of activity 5 because almost all the groups scored the same points. Now, let`s have a look at the Figure-7 which is showing student`s group performance.

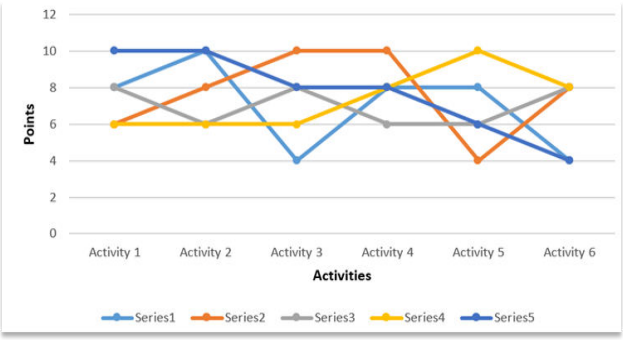

Figure-7. Students group performance graph.

Figure-7 shows the positive impact of dynamic grouping on students learning. If both the graph of individual and group performances are compared then the difference of fluctuation of points can be observed. The graph for group performance has low fluctuation and 3 groups out of 5 groups have scored the same points in the last activity.

We conducted the verification of the efficiency of the algorithm using statistical comparison with the result of the pre-test conducted based on the random grouping and K-mean clustering. Random grouping and k-means clustering are widely used in the literature for group formation [32] [26]. We experimented with the other 19 students for the pre-test. All the participants in the pre-test were also from the Object Programming Course and they were all from the second semester (Spring 2020). Same questions were used with the same difficulty level in the pre-test as well. Summary of our verification is:

- Dynamic groups generated using our proposed algorithm performed better than the groups generated randomly and k-mean clustering.
- Groups generated using k-mean clustering performs better than the groups generated randomly.

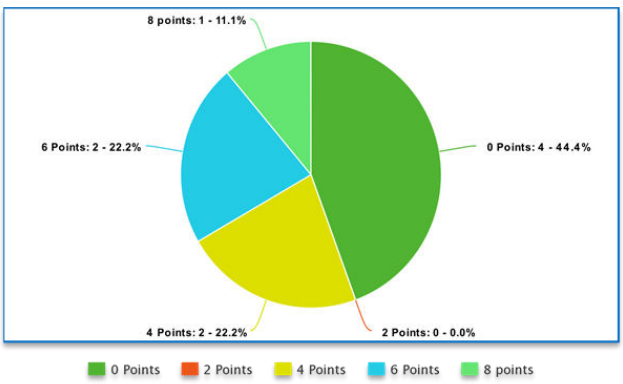

Figure-8. Result of groups generated randomly.

As it can be seen in Figure-8, four groups scored 0 points in different activities, which is 44% of the whole result. Only one group scored the highest point.

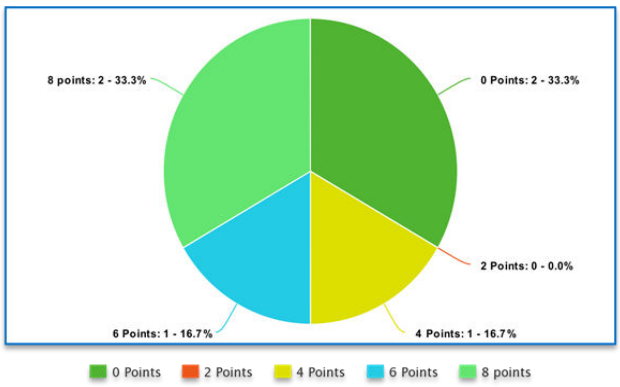

Figure-9. Result of groups generated using K-means.

Figure-9 above shows that two groups scored 0 points in different activities, which is 33% of the whole result. Two groups have reached the highest points. The result of the k-means clustering is better than the result of randomly generated groups.

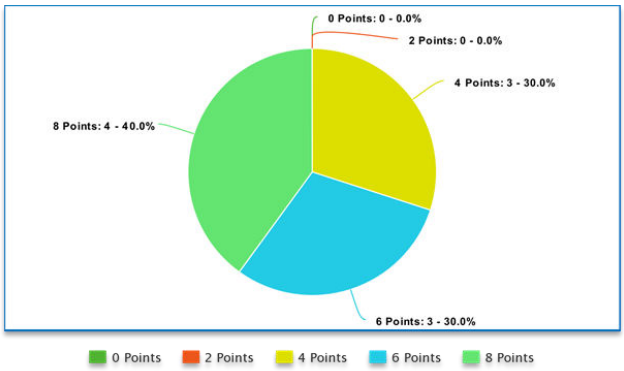

Figure-10. Result of groups generated using our proposed method.

Group performance has increased in the case of dynamic group formation using our proposed algorithm. Four groups reached the highest points and none of the groups scored 0 points in any activity as it can be seen in the pie chart in Figure-10.

The group formation is an essential part of collaborative learning. In literature, many researchers have proposed different techniques for group formation. We came up with our hypothesis that dynamic group formation can enhance the learning of students. Dynamic group means groups that are frequently changing based upon the performance. We experimented with pre and post-tests. Our methodology is evident from literature [32] [26]. In pre-test, we created groups using random grouping technique and k-mean clustering [32] [26]. In post-test, we used our proposed technique and we get the following results.

- In randomly generated groups 11% of students were able to get to the highest marks and 44% scored 0 marks.
- In groups created using k-mean based on the learning styles of students 33% student scored the highest marks and 33% scored 0 marks.
- Groups created using our technique 40% students scored highest marks, 30% scored 4 marks and 30% scored 6 marks.

Our results are supported by literature also as it states that heterogeneous groups perform better than homogeneous groups [4]. K-Means clustering create homogeneous groups while we are creating heterogeneous groups using our dynamic group formation method.

To further validate our result and make it trust able, we applied t-student test (statistical test). We have applied t-test because the sample size is less than 30. In order to apply t-student test we need null hypothesis. In our case null hypothesis is;

H0: dynamic group formation has no effect on the performance of groups in CSCL.

Data used in the pre-test is the result of activities performed by groups generated using k-means and for post-test we are using the result of activities performed by the groups generated using our algorithm. We are applying the t-test on the results of the groups so n = 5. Where n is the total number of groups generated during pre and post-test. These results are calculated with the confidence level of 95% (a = 0.05). Degree of freedom is n-1 which is 5-1 = 4. It is mentioned in the student's distribution table that the value of t, when degree of freedom is 4 and confidence is 95\%, is 2.7764. The t score we calculated for our pre and post-test is 3.5. Therefore, value of t score is greater than the value of t 0.05, which means that our null hypothesis is rejected and the other hypothesis is correct that states dynamic group formation has positive impact on the performance of groups.

Table-7. T-student test analysis table.

	N	Mean	SD	t score
Pre-test	5	3.6	10.24	3.5
Post-Test	5	6.4	3.84	

5.1 Research Findings

Q1. How to create dynamic groups in CSCL?

We have discussed in previous section 4 about dynamic group formation in CSCL. Complete details of algorithm have been discussed there in the previous section.

Q2. How to overcome the cold start i.e., how to create initial groups?

Initial group formation is very challenging. We solved this problem by letting student solve the first activity of our experiment phase 1 individually. Once, students solved the activity, we then calculated their knowledge level and based on their knowledge we created clusters of average and best students. After the creation of

clusters, groups are created with equal participation from each cluster of average and best students.

Q3. Can dynamic group formation enhance the group performance and student learning?

Yes, dynamic group formation enhances the learning of students. If we look at the figure 7 again, we can see that 3 out of 5 groups have performed better. So, we can conclude that dynamic grouping enhances the learning of students.

6. CONCULSION AND FUTURE WORK

We aimed to find out the impact of dynamic grouping on student`s learning. In this paper, we have focused on three objectives (1) how to create dynamic groups? (2) How to make initial groups? (3) what is the impact of dynamic group formation on student`s learning? To achieve our objective, we experimented in two phases. In the first phase, initial group formation is carried out, that was quite a challenging task and we solved this problem by letting students solve the first activity of our experiment phase 1 individually. Once, students solved the activity, their knowledge level is calculated, and based on their knowledge level clusters of average and best students are created using our proposed initial grouping algorithm. After the creation of initial clusters, groups are created with equal participation from each cluster of average and best students using our proposed dynamic grouping algorithm. Activities are assigned to these groups, which they solve in collaboration. After the completion of each activity, groups swapping took place. Our experimental results show that dynamic group formation has a positive impact on student learning. Student`s performance is better when groups are balanced. In the real time class environment, instructors evaluate students based on the individual assessments and group projects. This research study will help instructors create balanced groups based on the individual assessments. During the project-based assessment dynamic groups will enhance the learning and performance of students. The potential limitation of our study is that it is only feasible in a controlled environment like collaborative activities of classrooms. Therefore, the validity of this research can be improved in the following ways:

- In this study we are not measuring the interaction of each student, so we can extend this study to propose a method for measuring the participation of each students and their interaction with each other.
- Include open ended questions in our experiment and use Natural Language Processing (NLP to find out the impact of dynamic group formation in CSCL.
- We can add gamification in collaborative learning where students can be rewarded with points and ranks badges on their profile [34].

REFERENCES

[1] S. Borges, R. Mizoguchi, I. I. Bittencourt, S. Isotani. 2017. Group formation in cscl: A review of the state of theart, in: Researcher Links Workshop: Higher Education for All, Springer. pp. 71-88.

[2] S. Graf, R. Bekele. 2006. Forming heterogeneous groups for intelligent collaborative learning systems with ant colony optimization, in: International conference on intelligent tutoring systems, Springer. pp. 217-226.

[3] L.-K. Soh, N. Khandaker. 2007. Forming and scaffolding hu-man coalitions with a multi-agent framework, in: Pro-ceedings of the 6th international joint conference on Autonomous agents and multiagent systems. pp. 1-3.

[4] A. Ounnas, H. Davis D. Millard. 2008. A framework for semantic group formation, in: 2008 Eighth IEEE inter-national conference on advanced learning technologies, IEEE.

[5] R. M. Felder, L. K. Silverman, *et al.* 1988. Learning and teaching style in engineering education. Engineering education. 78: 674-681.

[6] C. E. Christodoulopoulos, K. A. Papanikolaou. 2007. A group formation tool in an e-learning context, in: 19th IEEE international conference on tools with artificial intelligence (ICTAI 2007), Vol. 2, IEEE, pp. 117-123.

[7] T.-F. Ho, S. J. Shyu, F.-H. Wang, C. T.-J. Li. 2009. Compos-ing high-heterogeneous and high-interaction groups in collaborative learning with particle swarm optimization, in: 2009 WRI World congress on computer science and information engineering, Vol. 4, IEEE, pp. 607-611.

[8] N. Rubens, M. Vilenius, T. kamoto. 2009. Automatic group formation for informal collaborative learning, in: 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology, Vol. 3, IEEE, pp. 231-234.

[9] R. Hubscher. 2010. Assigning students to groups using general and context-specific criteria. IEEE transactions on learning technologies. 3: 178-189.

[10] V. Yannibelli, A. Amandi. 2011. Forming well balanced collaborative learning teams according to the roles of their members: An evolutionary approach, in: 2011 IEEE 12th international symposium on computational intelligence and informatics (CINTI), IEEE. pp. 265-270.

[11] S. Abnar F. Orooji, F. Taghiyareh. 2012. An evolutionary algorithm for forming mixed groups of learners in web based collaborative learning environments. 1-6.

[12] S. Brauer, T. C. Schmidt. 2012. Group formation in e-learning-enabled online social networks, in: 2012 15th international conference on interactive collaborative learning (ICL), IEEE. pp. 1-8.

[13] J. Mor eno, D. A. Ovalle, R. M. Vicari. 2012. A genetic algorithm approach for group formation in collaborative learning considering multiple student characteristics. Computers & Education. 58: 560-569.

[14] A. Mujkanovic, D. Lowe, K. Willey, C. Guetl. 2012. Unsuper-vised learning algorithm for adaptive group formation collaborative learning support in remotely accessible laboratories, in: International Conference on Information Society (i-Society 2012), IEEE. pp. 50-57.

[15] H.-W. Tien, Y.-S. Lin, Y.-C. Chang, C.-P. Chu. 2013. A genetic algorithm-based multiple characteristics grouping strategy for collaborative learning, in: International conference on web-based learning, Springer. pp. 11-22.

[16] I. Srba, M. Bielikova. 2014. Dynamic group formation as an approach to collaborative learning support. IEEE trans-actions on learning technologies. 8: 173-186.

[17] Z. Zheng, N. Pinkwart. 2014. A discrete particle swarm optimization approach to compose heterogeneous learning groups, in: 2014 IEEE 14th international conference on advanced learning technologies, IEEE. pp. 49-51.

[18] M. R. Ullmann, D. J. Ferreira, C. G. Camilo, S. S. Caetano, L. de Assis. 2015. Formation of learning groups incomes using particle swarm optimization, in: 2015 IEEE Congress on Evolutionary Computation (CEC), IEEE. pp. 3296-3304.

[19] Y.-R. Chen, Y.-C. Lin, L. Chu, Y. Chiou, T. K. Shih. 2015. Team formation for collaborative learning with social network consideration based on edx's online discussion board, in: 2015 8th International Conference on Ubi-Media Computing (UMEDIA), IEEE. pp. 146-151.

[20] H. Sadeghi, A. A. Kardan. 2016. Toward effective group formation in computer-supported collaborative learning, Interactive learning environments. 24: 382-395.

[21] C. Yin, N. Uosaki. 2017. Building a group formation system by using educational log data, in: 2017 IEEE 17th International Conference on Advanced Learning Technologies (ICALT), IEEE. pp. 320-321.

[22] Y. Zheng, C. Li, S. Liu, W. Lu. 2018. An improved genetic approach for composing optimal collaborative learning groups, Knowledge-Based Systems. 139: 214-225.

[23] C. A. Bhardwaj, M. Mishra, S. Hemalatha. 2017. An auto-mated compatibility prediction engine using disc theory based classification and neural networks, arXiv preprintarXiv: 1709.00539.

[24] D. Jagadish, Groupin in collaborative e-learning environment based on interaction among students, in: 2014 International Conference on Recent Trends in Informa-tion Technology, IEEE, 2014, pp. 1-5.

[25] E. M. Maina, R. O. Oboko, P. W. Waiganjo. 2017. Using machine learning techniques to support group formation in an online collaborative learning environment. International Journal of Intelligent Systems & Applications. 9.

[26] Y.-C. Kuo, H.-C. Chu, C.-H. Huang. 2015. A learning style-based grouping collaborative learning approach to im-prove efl students' performance in english courses. Jour-nal of Educational Technology & Society. 18(2).

[27] C. Quarto, A. Busson, L. C. Fonseca, M. Bercht, E. Eder. 2019. Formation of groups for collaborative learning: An ap-proach based on profiles of individual self-reports, in: 2019 IEEE 19th International Conference on Advanced Learning Technologies (ICALT), Vol. 2161, IEEE, pp. 134-136.

[28] I. Amarasinghe, D. Hern´andez-Leo, A. Jonsson. 2017. Intel-ligent group formation in computer supported collaborative learning scripts, in: 2017 IEEE 17th International Conference on Advanced Learning Technologies (ICALT), IEEE. pp. 201-203.

[29] I. Boticki, L. H. Wong, C.-K. Looi. 2012. Designing technology for content-independent collaborative mobile learing, IEEE Transactions on Learning Technologies. 6(1): 14-24.

[30] A. Zedadra, Y. Lafifi, O. Zedadra. 2016. Dynamic group formation based on a natural phenomenon. International Journal of Distance Education Technologies (IJDET). 14(4): 13-26.

[31] M. M. J. Wijaya, D. W. Utomo, D. Kurniawan. 2019. Imple-menting dynamic group formation in web-based collaborative learning for high school, in: 2019 International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), IEEE. pp. 528-533.

[32] J. Bernacki, A. Kozierkiewicz-Hetma ́nska. 2014. Creating collaborative learning groups in intelligent tutoring systems, in: International Conference on Computational Collective Intelligence, Springer. pp. 184-193.

[33] R. Hendijani, D. P. Bischak, J. Arvai, S. Dugar. 2016. Intrinsic motivation, external reward, and their effect on overall motivation and performance, Human Performance. 29(4): 251-274.

[34] C. Li, Z. Dong, R. H. Untch, M. Chasteen. 2013. Engaging computer science students through gamification inan online social network based collaborative learning environment. International Journal of Information and Education Technology. 3(1).