# Favoring Collaborative Learning in PBL: An Automated Solution for Semantic Group Formation

Ricardo E. de Santana
*Centro de Informática*
*Universidade Federal de Pernambuco*
Recife, Brazil
res2@cin.ufpe.br

Simone C. dos Santos
*Centro de Informática*
*Universidade Federal de Pernambuco*
Recife, Brazil
scs@cin.ufpe.br

Davi J. Mendes Maia
*Centro de Informática*
*Universidade Federal de Pernambuco*
Recife, Brazil
djmm@cin.ufpe.br

***Abstract*— This research-to-practice full paper proposes an automated solution for forming groups of students in computing higher education with the Problem-Based Learning (PBL) approach. PBL is a pedagogical model that promotes the development of professional skills (knowledge, skills, and attitudes) for problem-solving through collaborative learning. In this context, effective team building is essential to the success of PBL, as it can enhance student learning and development. The formation of teams using the PBL approach involves effort and observation of several aspects, which makes manual grouping inefficient in formatting balanced groups. In addition, both the instructor and the student can participate in the grouping decision process, considering the constraints defined by the instructor and the student's satisfaction with the recommendations made. In this sense, when using an automated approach that considers these criteria and possibilities for forming groups, it is possible to streamline the process of structuring teams and mitigate the formation of groups with low potential for success. This study proposes an automated solution for forming groups of students using the DSR (Design Science Research) method applied in evolutionary cycles. It considers the following research question: RQ) How can balanced groups in computing higher education with PBL be formed automatically, considering predefined criteria, attributes, and affinities between their members? This solution considers multiple attributes of the individuals involved and the organizer's criteria for dividing groups, using a semantic structure for group formation. The solution also allows students to negotiate based on affinity with peers and satisfaction with the group. System prototypes were created and evaluated throughout the DSR cycles to evaluate the proposal, demonstrating compliance with the defined restrictions and indication of balanced teams. As primary limitations, we point out challenges in student data collection that hinder group formation. The DSR method iteratively improved solutions, but more application cycles will be necessary to ensure the solution's robustness. In future work, further analysis should link team performance to group formation, focusing on criteria for balanced groups in PBL courses.**

***Keywords—Computing Higher Education, PBL, Collaborative Learning, Group Formation, Web- semantics***

## I. Introduction

Problem-based learning (PBL) [1]; [2], as its derivation from Project-Based Learning (PrBL) [3], is an instructional model that places the student as an active and interactive actor at the center of the learning process. This model aims to develop knowledge, skills, and attitudes to solve real-world problems, stimulating critical thinking and problem-solving skills through group work [4]. In this context, one of the main demands for adopting PBL is the formation of teams or work groups, considering that, in general, the problems presented are complex and require collaboration between students to be solved. Furthermore, when working in groups, students can learn from each other, one of the principles of PBL, by sharing knowledge, skills, and experiences. Thus, group formation is an essential part of the pedagogical process and can be crucial for the success or failure of the performance of this educational approach [5]. In practice, the formation of teams in PBL-based and PrBL-based education must consider several personal and interpersonal factors, such as age group, personality, individual student skills, and affinities with colleagues, among others [6].

Research suggests that personal characteristics combined with the team's internal relationships interfere with the group's final result and members' learning [5][6][7]. Additionally, according to [7], the "ideal formation" of teams is the one with more heterogeneity within each group (intra-heterogeneous) and more similarities between several of them (inter-homogeneous), promoting balanced groups. When well planned and executed, team formation can enhance student learning, favoring collaboration and allowing them to develop essential skills, such as communication, teamwork, and conflict resolution.

The formation of teams using the PBL approach involves effort and observation of several aspects, which makes manual grouping inefficient in formatting balanced groups. Besides, [5] suggests that the instructor and the student can participate in the grouping decision process, considering the constraints defined by the instructor and the student's satisfaction with the recommendations made. In this sense, when using an automated approach that considers these criteria and possibilities for forming groups, it is possible to streamline the process of structuring teams and mitigate the formation of groups with low potential for success. In this context, we were motivated by the following central research question: RQ) *How can balanced groups be automatically formed in computing higher education with PBL, considering pre-defined criteria, attributes, and affinities between its members?*

Thus, this study proposes an automated solution for team formation motivated by the collaborative learning characteristics of the PBL approach. It considers the multiple attributes of the individuals involved and the organizer's criteria for dividing groups so that the teams formed have the same potential for success in the tasks that the methodology requires. The solution is based on a semantic group formation framework defined by [5], allowing the student to negotiate based on affinity with colleagues and satisfaction with their group.

As a result, this study proposes a group formation process and an algorithm based on the PBL criteria to assist this process through prerequisites established by the instructor, personality profile classification methods, and student affinity. Real-world experiments prove that the prototype met the constraints for the defined set of criteria, indicating balanced teams. Thus, the solution is promising despite still needing improvements, particularly in the user interface.

This paper is divided into seven sections. After this brief introduction, Section II discusses concepts related to group formation in education and PBL. Section III presents some related works to this current study. Section IV describes the research methodology. Sections V and VI, respectively, discuss the proposal solutions and how they were constructed along DSR cycles. Finally, Section VII presents the conclusions and future works.

## II. Background

### A. Group Formation in Education

According to [5], each time a domain is assigned to the context of team formation, the definition of a group itself is modified; that is, defining what a group is depends on several contextual factors that vary for each situation, such as the purpose and form of interaction, for example. The labor market frequently uses the formation of groups; they are essential for achieving the departments' objectives. Many characteristics of school group formations are identified in work-oriented ones such as division of tasks and determination of specific functions, however there is a difference between forming teams for the world of work and forming teams in the educational context.

In the work-oriented context, delivering results is the main parameter for evaluating the group's success. The primary focus is on optimizing tasks to achieve predefined goals, so the more specific skills and technical knowledge the group members have, the more likely they are to succeed. According to [8], agile methodologies, such as Scrum, are notable examples of team formation in this sense, detailing the process of executing tasks to maximize effectiveness in completing specific objectives.

Group formation in the educational field focuses on developing skills related to group work, emphasizing the learning process. Members need not be ready to achieve the proposed result. Still, the final objective must be to efficiently provide students with the acquisition, enhancement, and refinement of technical and soft skills, such as teamwork, cooperation, and problem-solving skills, collaboratively throughout their learning process.

In the educational sphere, it is necessary to understand the dynamics of team formation considering this entire scope. In addition to each participant's individual knowledge, the social environment in which they are inserted, together with the interaction they develop with each other, is fundamental for cognitive development.

### B. Group Formation in PBL

PBL is a model that fits within the constructivist view of learning, whose knowledge production is obtained through the learner's interactions with the environment. Thus, the individual is the active agent of their knowledge, constructing meanings and defining their representations of reality based on their experiences in different contexts.

According [6] small teams, as well as the work performed in them, are a fundamental part of PBL, since these teams are the core where problem-solving activities are carried out, and their main focus is the student learning. In these interactions, students develop, individually and in groups, knowledge, skills, and attitudes, such as socio-emotional skills and motivational qualities, in addition to their technical expertise and the networking provided in these interactions.

To develop this set of attributes, it is important that team formation in PBL is rigorous and well-defined and takes into account some dimensions, such as team size, individual differences such as personality, levels of skills and work preference, professional or teamwork experience, and diversity of gender, culture, and society.

Regarding the dimensions mentioned above, it is important to know each of them, as well as their adherence to the process of forming teams in PBL. As a first dimension, team size plays a crucial role in the performance of activities since the number of members can positively or negatively influence the team climate and the execution of work. If the team is large, it can cause an unequal division of activities, indicating overload in some members and lack of activity/work in others. Conversely, tiny teams can cause work overload, harming the team's performance in PBL activities and other disciplines. Given these issues, authors such as [6] indicate an average size of 5-6 students for PBL teams, also considering the size of the class in question, generating an average of 4-6 teams per class.

Another essential characteristic of team formation is personality since it is from this that each team member understands how their relationship with others, the world, and the work they perform in the team occurs. With personality, the manager can better define which type of work suits each person and use the best communication style. The studies in [9], [10], and [11] present a preference for diversity of personalities in PBL teams so that each one contributes some aspect to the work performed, be it communication, organization, creativity, agility and proactivity.

Professional experience or experience with teamwork is an important characteristic in the process since, based on their previous experiences, team members can identify better ways of solving problems and carrying out activities, play roles as instructors and mentors, and strengthen and develop skills such as self-regulation and self-direction [12].

Cognitive skills and work preferences play a fundamental role in the team. Teams with members with different levels of expertise and cognitive skills can act as tutors or mentors, share experiences with other members, participate in the training process, and assist with possible doubts or difficulties encountered in the problem-solving process.

Diversity, in terms of gender, culture, and society, contributes positive aspects to the formation of teams. It can improve performance in activities, encourage motivation and engagement among members, and help members see meaning in their work and its social contribution. Another positive aspect is the ability to evaluate different ways of thinking about a problem based on different social and gender contexts.

Finally, something important to consider when thinking about PBL and team building is that most of the activities in this approach are carried out in teams, and the skills built from these activities emerge from the interpersonal relationships between team members. Therefore, one dimension to be considered when dividing teams is students' affinity with each other.

In conclusion, the diversity in the individualities of each member, which, in its way, contributes to this process, constitutes an essential tool in the formation of PBL teams.

## III. Related Works: Group Formation Automation

From a non-systematic literature review, this section presents related works to this current study, summarized in Table I. These studies were collected using the terms "group formation", "algorithm", and "system" in the Scopus and Google Scholar databases. Based on the satisfactory results found, we chose not to use a rigorous process of a systematic literature review.

Reference [13] proposes a theoretical approach to group formation, which considers personality as a relevant factor in collaboration between members. The article presents an algorithm for forming groups that support new learning roles, called Affective Collaborative Learning roles, based on the relationship between collaborative learning theories and students' personality traits. The evaluation consisted of a series of experiments with university students, where groups were formed based on personality criteria and were evaluated on performance and member satisfaction. The results indicated that the personality-based approach can improve collaboration effectiveness and increase group member satisfaction. The study concludes that forming groups based on personality characteristics may be a promising approach to enhancing collaboration and group performance in collaborative learning environments. The study points to personality as a critical characteristic for team building; this aligns with the present study, as it also uses this characteristic to focus on collaborative learning environments. Despite this, it does not adhere to the context of PBL or consider the team size.

Reference [5] presents a system that uses ontologies to represent knowledge about students and their characteristics, allowing the system to automatically select group members based on their skills, knowledge, and interests. Additionally, the system will enable teachers to customize group formation criteria to meet the specific needs of their class. The authors concluded that considering the diversity of students' characteristics can help make group formation in educational environments more efficient and personalized. The work is in line with the present study concerning the personalization of attributes by the teacher, in addition to taking advantage of the student's interests and skills. Despite this, it does not adhere to the PBL context.

Reference [14] presents an ontology-based approach to group formation in educational environments based on pedagogical theories. The authors propose a theoretical approach to group formation based on a set of pedagogical theories represented through an ontology that uses a reference model for forming heterogeneous groups in some dimensions and homogeneous in others based on specific criteria. The work is exciting and aligns with the present study using students' interests and skills. Despite this, it does not adhere to the PBL context and does not consider team size, gender, and personality.

Reference [15] presents a framework for group formation based on an algorithm (GAGFS-PF) that considers levels of knowledge and learning roles and types of interactions in social networks. The study considers three training schemes: completely random, training with self-selection, and calibrated by interviews. The results showed that the algorithm works to form random teams with self-selection. Although necessary and in line with the context of this study regarding the PBL approach and the use of some characteristics such as cognitive levels and learning roles, the study did not take into account the context of higher education in computing; in addition, the way it used the attitudinal part was through social networks and not personality.

The current research is a case study presenting an automated process for forming PBL teams in higher education in an Information System (IS) course. The model ranges from procedures for collecting student data through the distribution of teams according to defined criteria and evaluation. The algorithm was modeled using the evaluation dimensions: learning and behavior profile, demographic data (age, experience), technical skills (programming, modeling, managing), and affinities, in addition to considering the size of the team. These dimensions were formulated based on the principles of PBL and later presented. Although the algorithm was implemented based on these dimensions, the restriction criteria can be customized to meet the organizer's purpose.

TABLE I. Comparative Table of Related Studies

| **Refs** | **Characteristics** | | |
|---|---|---|---|
| | ***Proposal*** | ***Context*** | ***Formation Criteria*** |
| [13] | Group formation algorithm using personality traits | Computing Higher Education Students | Context, learning activities, group structure and personality |
| [5] | Ontology-based group formation system | Statistics & Computing Higher Education Students | Team size, interests, and personal skills |
| [14] | Ontology-based group formation algorithm | Higher Education Students | knowledge, skills, learning roles and teacher preferences |
| [15] | Group formation scheme based on genetic algorithm | K-12 Students | Knowledge levels, learning roles and interactions on social networks |
| *Current Research* | *Web-semantics-based group formation algorithm* | *PBL-based Computing Higher Education Students* | *Context, team size, personality, diversity (gender, age, professional experience), student preferences, affinities and learning roles* |

## IV. Research method

The methods used in scientific research were based on the DSR methodology [16], according the steps shown in Fig 1.

The research began with identifying the problem during the tutoring work of a Management Information System (MIS) subject in an undergraduate Information Systems (IS) course based on PBL.

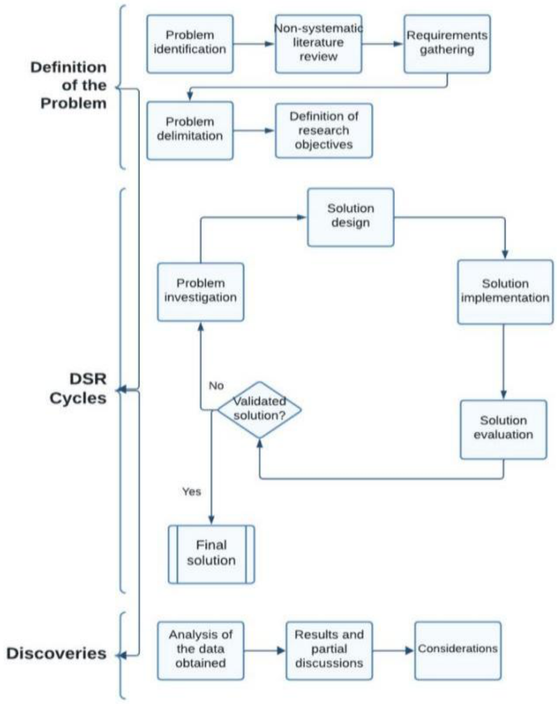


Fig. 1. Methodology scheme

The primary motivation originated from the interaction of the authors of this article with students in their work environment.

A non-systematic literature review was conducted to delimit and understand the elements limiting the problem. It was based on studies on the importance of forming teams in the PBL methodology to formulate its resolution.

The problem was defined during the application of the PBL method in MIS discipline taught and tutored by the authors. The need to speed up this process was realized during the selection of teams, as it was long and laborious for the teacher, and the possibility of generating unbalanced teams was much greater due to human limitations.Thus, theoretical references that supported the team formation automation process from the PBL perspective were searched.

Stakeholder demands were analyzed based on the formation of groups from previous classes and dialogues with the subject teacher. Among the characteristics of the solution, we sought one that formed heterogeneous teams and met the restrictions presented by the teacher. These restrictions were adjusted according to the organizer's needs and the class's characteristics after data collection. In this point, the research question was defined: *How can balanced groups be automatically formed in computing teaching with PBL, considering pre-defined criteria, attributes, and affinities between its members?*

Subsequently, the team formation automation algorithm was created using the Design Science Research (DSR) method. DSR is a method applied in cycles, each composed of stages. In this research, three cycles were carried out, each consisting of four stages: problem investigation, solution design, implementation, and evaluation.

The four aforementioned stages are carried out in different classes concerning DSR cycles. This study describes the three iterations and the partial results obtained from them.

In the first stage of investigating the problem, the teacher initially had an ideal set of restriction criteria. However, before the set of restrictions was finalized, the students' profiles and affinities were surveyed, and the necessary adaptations were made. The problem is understood within the cycle as forming balanced teams with students from a given class.

In the solution design stage, an algorithm was thought of that used the data collected from students as a database to distribute it among teams following the satisfaction of proposed restrictions in order of priority.

The implementation was the stage of building and improving the computational code of the PBL clustering algorithm.

The last step, evaluation, compares the results of implementing the solution with the validation parameters.

In cycle 1, the algorithm was generated. However, the application was made with data from a class that had already been completed, as the academic period had already ended. Therefore, the evaluation was carried out based on checking the satisfaction of the restrictions only. In cycles 2 and 3, the artifact was evaluated based on the defined criteria, and the teams were finalized after the class had completed the negotiation stage.

The system makes relevant contributions towards speed and efficiency in forming PBL teams.

## V. Group Formation Proposal

Agreeing with [7] definition of balanced groups, we sought a team formation model that was as diverse as possible, mixing similar profiles between them so that the possibility of success for each team was equal.

Based on the author's experience (class teacher), an initial set of restrictions was created before interacting with students based on the profiles that are historically most common in the discipline. These criteria are presented and justified according to Table II.

TABLE II. Criteria for Group Formation

| Group Formation Criteria | | |
|---|---|---|
| ***Criteria*** | ***Justification*** | ***Refs*** |
| 1. A mix of MBTI personalities | Forming groups with diverse MBTI personalities ensures a variety of skills and perspectives, leading to better problem-solving and decision-making. Teams with MBTI personality diversity tend to be more creative, innovative, and productive. | [17] |
| 2. Minority MBTI personalities distributed in teams | It is crucial to have representation of minority MBTI personalities in teams to prevent marginalization and leverage their unique perspectives. | [17] |
| 3. Minority gender distributed in teams | It is crucial to have minority gender representation in teams to prevent marginalization | [18] |

| Group Formation Criteria | | |
|---|---|---|
| ***Criteria*** | ***Justification*** | ***Refs*** |
| | and leverage their unique perspectives. | |
| 4. At least one member with professional experience | Team members with professional experience can bring real-world knowledge and skills to the team, as well as help mentor and support other team members. | [6] [12] |
| 5. Affinities identified from the students' interaction | Teams with pre-existing affinities can collaborate more fluently and productively. | [19] |

Fig. 2 presents a process of group formation. In the first interaction with the class, students are introduced to the PBL methodology, and their profiles and affinities are surveyed.

Students initially fill out an online form that allows them to classify their personality type based on their MBTI (Myers-Briggs Type Indicator) profiles. With their MBTI profile, students are directed to fill out a spreadsheet with data relating to the criteria selected by the teacher, in this case, name, age group, gender, Keirsey profile (converted from MBTI), experience time, preferential activity (Program - P, Modeling - Mo or Managing - Ma) and affinities (participants with whom you want to work). The perfil Keirsey is used to facilitate understanding skills based on the two dimensions communication and action [20].

After students fill out the two forms, data preparation occurs in step 1 of the cycle. Based on the data obtained, the necessary adjustments are made to the pre-established restriction criteria to adapt them to the reality of each class. For example, if no women are enrolled in each class, restriction 3 makes no sense and will soon be removed from the list.

In step 2, solution design, it is defined how the algorithm behaves in more complex cases and the priority order between the criteria. The algorithm adopted for team formation operates in an environment developed in Python using the Pandas library for data manipulation. The algorithm logic was structured around customizing specific priorities and criteria, such as personality profile (profile_keirsey), gender, experience, preference, age group, and affinities.

In step 3, the selection and distribution of students into groups is implemented team by team. As a group was formed using the proposed restrictions, all groups are formed following the same criteria.

The code uses control structures to filter and allocate students based on these criteria. The algorithm maintains flexibility by allowing personalized choice of priority personality profiles. Additionally, the technical approach employs the manipulation of data structures such as lists and dictionaries to store and manage student information.

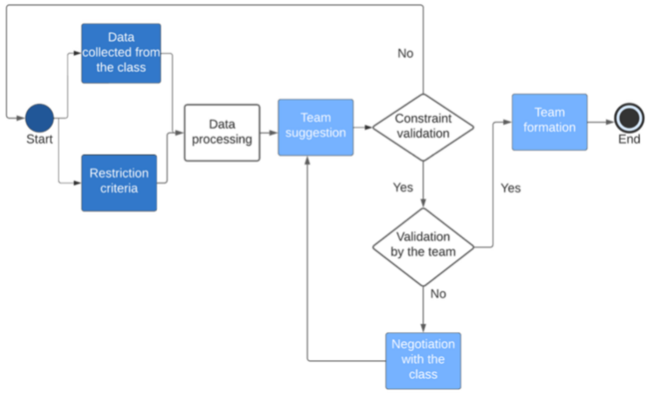


Fig. 2. Group formation process

Finally, the suggested distribution of students by the algorithm is carried out, seeking to form balanced teams, optimizing the heterogeneity of the groups formed. Still, the constraints are checked after the teams' first suggestion to ensure satisfaction. If any restriction has not been satisfied, the process begins again. If all restrictions have been met and there is no objection from any member, then the teams are finalized; otherwise, the negotiation process with the group begins.

## VI. Results & Discussions

In this section, the results of the system will be presented for the three cycles applied. Cycles 1, 2 and 3 were applied in three classes followed by the MIS discipline. The way the data was collected significantly influenced the results obtained.

### *A. Cycle 1*

The formation of teams in Cycle 1 was carried out manually; however, even after finishing the course, the data from these students were used to carry out the first iteration of the algorithm.

This class had 28 students. As the academic period for this class had already ended, it was necessary to work exclusively with the data recorded in the spreadsheet that they filled out, resulting in many inaccuracies.

The data in Fig. 3 shows the profile of students in this class. The age groups were separated into intervals: A - 18 to 20 years old; B - 21 to 25 years old; C - over 25 years old.

| | | | TOTAL |
|---|---|---|---|
| **Gender** | **M** | 25 | 28 |
| | **F** | 3 | |
| **Age Range** | **A** | 12 | 24 |
| | **B** | 7 | |
| | **C** | 5 | |
| **Keirsey Profile** | **Craftsman** | 5 | 26 |
| | **Guardian** | 8 | |
| | **Idealistic** | 8 | |
| | **Rational** | 5 | |
| **Work experience** | **No experience** | 10 | 25 |
| | **With experience** | 15 | |
| **What do you like to do most? (P)rogram, (Mo)del or (Ma)nage?** | **P** | 17 | Not applicable |
| | **Mo** | 11 | |
| | **Ma** | 13 | |

Fig. 3. Student profiles in Class 1

In this cycle, the affinities of the students were not considered; there were a smaller number of "Artisan" and "Rational" profiles and a mostly male class, so, after collecting the data, the set of restrictions was as follows:

1. Each team comprises a mix of profiles.

2. One "Artisan" or "Rational" per team.
3. Female members distributed across teams.
4. At least one team member with some professional experience.

The first distribution was obtained after implementing the team formation algorithm, as shown in Fig. 4.

| | Criterion 1 | Criterion 2 | Criterion 3 | Criterion 4 |
|---|---|---|---|---|
| Team 1 | ✔ | ✔ | ✔ | ✔ |
| Team 2 | ✔ | ✔ | ✔ | ✔ |
| Team 3 | ✔ | ✔ | ✔ | ✔ |
| Team 4 | ✔ | ✔ | ✔ | ✔ |
| Team 5 | ✔ | ✔ | ✔ | ✔ |
| Team 6 | ✔ | ✔ | ✔ | ✔ |

Fig. 4. Verification of constraints satisfaction in Cycle 1.

All teams had a mix of profiles related to practically all possible criteria; however, a point of improvement would be to add the age group as one of the profile variation factors within the team. Although some teams were without women, three female representatives were distributed into different groups.

Some improvements were made, such as improving the data collection form, which was considered for adjustment in Cycle 2. In addition, Cycle 2 improved the issue of considering the students' affinities.

### *B. Cycle 2*

To collect data from the 38 students in the second class, we used a spreadsheet with the appropriate validations for all questions relating to the restriction criteria. Students could indicate only a single answer, solving possible doubts or inconsistencies in the data. The last one was the unique exception.

The last question of the questionnaire dealt with affinities. In it, students typed the names of colleagues with whom they wanted to work, so in this cycle, we sought to consider the affinities between students.

Fig. 5 shows a summary of the profile of this class.

| | | | TOTAL |
|---|---|---|---|
| Gender | M | 30 | 38 |
| | F | 8 | |
| Age Range | A | 16 | 38 |
| | B | 19 | |
| | C | 3 | |
| Keirsey Profile | Craftsman | 10 | 38 |
| | Guardian | 17 | |
| | Idealistic | 8 | |
| | Rational | 3 | |
| Work experience | No experience | 10 | 38 |
| | With experience | 28 | |
| What do you like to do most? (P)rogram, (Mo)del or (Ma)nage? | P | 20 | 38 |
| | Mo | 8 | |
| | Ma | 10 | |

Fig. 5. Student profiles in Class 2

In this class, the number of women was still well below half of the students, although this number has increased proportionally to more than 26% of the class compared to 12% in the previous one. Thus, distributing women across groups remains a goal in this cycle.

Another topic that deserves attention is the Keirsey profile. The minority was "Idealist" or "Rational," so these are the rarest profiles that should be distributed as a priority in teams.

Therefore, the set of final restrictions for this class was:

1. Each team is made up of a mix of profiles.
2. One "idealist" or "rationalist" per team.
3. Female members distributed across teams.
4. At least one team member with some professional experience.
5. Affinities identified from the students' interaction.

Fig. 6 presents the evaluation of the distribution of teams from the perspective of satisfying the established restrictions.

| | Criterion 1 | Criterion 2 | Criterion 3 | Criterion 4 | Criterion 5 |
|---|---|---|---|---|---|
| Team 1 | ✔ | ✔ | ✔ | ✔ | ✔ |
| Team 2 | ✔ | ✔ | ✔ | ✔ | ✔ |
| Team 3 | ✔ | ✔ | ✔ | ✔ | ✔ |
| Team 4 | ✔ | ✔ | ✔ | ✔ | ✔ |
| Team 5 | X | X | ✔ | ✔ | ✔ |
| Team 6 | ✔ | ✔ | ✔ | ✔ | ✔ |

Fig. 6. Verification of constraints satisfaction in Cycle 2

In all groups there was at least one person whose affinity was met, however, this is a point for improvement of the algorithm, as there were several incompatibilities due to errors in spelling, formatting, or ambiguities in the typing of names.

As shown, criteria 1 and 2 were not satisfied in Team 5, as the team's composition did not include an Idealist or Rational profile, in addition to the type C age group and people with a preference for Modeling (Mo). Therefore, the team also violated criterion 1 (mix of profiles).

Thus, improving the distribution of profiles between the groups was also necessary for the following cycle.

### *C. Cycle 3*

Class 3 had 36 students, predominantly male again; the vast majority are very young and have no professional experience, as shown in Fig. 7.

| | | | TOTAL |
|---|---|---|---|
| Gender | M | 27 | 36 |
| | F | 9 | |
| Age Range | A | 26 | 36 |
| | B | 7 | |
| | C | 3 | |
| Keirsey Profile | Craftsman | 6 | 36 |
| | Guardian | 9 | |
| | Idealistic | 8 | |
| | Rational | 13 | |
| Work experience | No experience | 27 | 36 |
| | With experience | 9 | |
| What do you like to do most? (P)rogram, (Mo)del or (Ma)nage? | P | 17 | 35 |
| | Mo | 12 | |
| | Ma | 6 | |

Fig. 7. Student profiles in Class 3

One of the students did not indicate his technical preference, as completing this item was not mandatory. This constitutes a pending issue to be corrected in the system.

In this class, a minority of "Artisan" and "Idealist" profiles were obtained; therefore, analyzing the specific characteristics of the class, the final set of restrictions was formulated in this case:

1. Each team is made up of a mix of profiles.
2. One "craftsman" or "idealist" per team.

3. Female members distributed across teams.
4. At least one team member with some professional experience.
5. Affinities identified from the students' interaction.

As in the previous cycle, identifying affinities was problematic, so it was decided to ask students to select their ID instead of the participant's name.

Regarding the satisfaction of restrictions in Cycle 3, Fig. 8 shows the achievement or not of each restriction in each team.

| | Criterion 1 | Criterion 2 | Criterion 3 | Criterion 4 | Criterion 5 |
|---|---|---|---|---|---|
| Team 1 | ✔ | ✔ | ✔ | ✔ | ✔ |
| Team 2 | ✔ | ✔ | ✔ | ✔ | ✔ |
| Team 3 | ✔ | ✔ | ✔ | ✔ | ✔ |
| Team 4 | ✔ | ✔ | ✔ | ✔ | ✔ |
| Team 5 | ✔ | ✔ | ✔ | ✔ | ✔ |
| Team 6 | X | ✔ | ✔ | X | ✔ |

Fig. 8. Verification of contraints satisfaction in Cycle 3

In most groups, the criteria were met, including affinity. As team 6 was the last to be formed and this group had very few people with experience and prone to management, the algorithm ran out of options to meet criteria 1 and 4.

In addition to these improvement points, the algorithm interface needed to be developed after Cycle 3 was applied.

It is worth adding that at the end of each cycle's implementation, the teacher opens a negotiation round with the students to consider the possibilities of meeting more student requests; therefore, the final training is different from that generated by the algorithm.

### *D. Lessons Learned*

Automated group formation can significantly impact students' interpersonal skills. By being grouped based on criteria such as age, professional experience, activity preferences, and affinities, students can develop socio-emotional competencies, such as communication, teamwork, and conflict resolution. The diversity within the groups also promotes empathy and understanding of different perspectives, which is essential for the development of interpersonal skills.

Many of the difficulties encountered in processing the collected data arose from the moment of collection; therefore, this stage of the cycle is definitive for the system's satisfactory result.

The data collection for the automated formation of PBL groups was conducted using a Google Sheets spreadsheet shared with the students and teachers of the course. This tool provided easy and intuitive access to the data, facilitating the students' input.

To ensure the quality of the collected data, several measures were implemented. Initially, clear and objective instructions were provided on how to fill out each field in the spreadsheet, aiming to minimize errors and inconsistencies. Additionally, the spreadsheet was configured with automatic validation for certain fields to prevent the entry of invalid data. After the data collection was completed, the teaching team manually reviewed the data to identify and correct any possible errors or inconsistencies.

The security of the collected data was a central concern throughout the entire process. To ensure the protection of the information, several measures were implemented. Firstly, the Google Sheets spreadsheet was configured with restricted access permissions, allowing only viewing and editing by the students and the teaching team. Furthermore, the spreadsheet was stored in a secure environment, accessible exclusively to the research team members. The entire process of data collection and storage was conducted in compliance with ethical research standards, ensuring respect for the participants' rights.

After the distribution of teams in Cycle 3, the system underwent improvements in the code and interface. Fig. 9 shows the current system in which the Streamlit library was used to build the WEB interface.

In this update, the organizer loads the system with data collected from students in a CSV file, chooses the desired number of teams, chooses the criteria according to the order of priority they want, and determines which Keirsey profiles should be distributed in priority according to with the reality of the class. Once done, click and generate the teams.

After implementing the algorithm, we can understand that the process of students forming teams by teams is not ideal, as the first teams formed are more likely to meet the criteria to the detriment of the last. Therefore, for the continuity of the work, the authors intend to use a new approach, grouping criteria by criteria in all classes per round.

Even so, the system presented generates a good initial formation, so the organizer can compose his balanced team formation based on it and with a few modifications.


Fig. 9. Interface of PBL Team Former

## VII. LIMITATIONS

As in all exploratory research, we face challenges and limitations. In the first cycles, there were some inconsistencies regarding student data collection, making it difficult to form more balanced groups. According to the DSR method, the proposed solution was improved with each round of testing. However, there were few interactions in this complete and adjusted process to obtain more robust results. More application cycles of the system would reinforce the results and considerably increase the reliability of the research.

Another limitation of the research was the non-linking of team performance results in the discipline with the formation of teams. A successful team in PBL is understood to be one that manages to develop the planned educational skills throughout the problem-solving process. Validating the extent to which the choice of teams influences this success is a good parameter to include in a future analysis. In this work, we focus on the method of satisfying criteria in group formation as a first step of ongoing research.

Finally, the research's group formation criteria were designed to create balanced groups to solve specific problems in courses based on PBL, considering its characteristics of intense collaboration and teamwork, problem-solving based on discussions (stimulating critical analysis, creativity), and consensus (decision-making). Thus, considering other teaching and learning approaches can demand another set of criteria not considered in this study.

## VIII. Conclusions

The PBL methodology demands intense and fine-tuned teamwork; therefore, forming well-adjusted and balanced teams can impact both the result and the collective and individual learning process.

Hence, the organizer must be doubly careful when choosing working groups, and this requires immense effort. Therefore, automating the team formation process using the PBL methodology is a relevant project.

This push toward automation represents an evolution in pedagogical practice and a fundamental step toward driving the continued effectiveness and popularity of PBL methodology in the current educational landscape.

As future work, improvements to the interface, database integration and customization of criteria based on class contexts are suggested. Efforts to optimize the distribution of student profiles within teams, enforce mandatory fields, integrate negotiation mechanisms, and enable team archiving will further improve system functionality and user experience. These improvements aim to promote more equitable and effective team compositions, thus enriching collaborative learning outcomes and individual growth within PBL.

## References


[1] Yu, W., 2005. Work in progress-a mobile computing collaborative framework for problem-based learning environment. In Proceedings Frontiers in Education 35th Annual Conference. IEEE, pp. F4E–12.

[2] Ibrahim, N. and Halim, S. A., 2014. Generic framework design of project-oriented problem-based learning (popbl) for software engineering courses. In 8th. Malaysian Software Engineering Conference (MySEC). IEEE, 2014, pp. 359–364.

[3] Bell, S., 2010. Project-Based Learning for the 21st Century: Skills for the Future. In The Clearing House, 83: 39–43, Taylor & Francis Group, LLC. ISSN: 0009-8655print. DOI: 10.1080/00098650903505415.

[4] Dos Santos, Simone C.; Soares, Felipe SF. Authentic assessment in Software Engineering education based on PBL principles a case study in the telecom market. In: 2013 35th International Conference on Software Engineering (ICSE). IEEE, 2013. p. 1055-1062.

[5] Ounnas, Asma. Enhancing the automation of forming groups for education with semantics. 2010. Tese de Doutorado. University of Southampton.

[6] Fonteijn, Herco TH; Dolmans, Diana HJM. Group work and group dynamics in PBL. The Wiley handbook of problem-based learning, p. 199-220, 2019.

[7] Moreno, J.; Ovalle, D. A.; Vicari, R. M. A genetic algorithm approach for group formation in collaborative learning considering multiple student characteristics. In: Computers & Education, v. 58, n.1, p.560-569, 2012.

[8] Schwalbe, K. (2012, September). Managing a project using an agile approach and the PMBOK® guide. In Proceedings of the Information Systems Educators Conference ISSN (Vol. 2167, p. 1435).

[9] Driskell, J. E., Goodwin, G. F., Salas, E., & O'Shea, P. G. (2006). What makes a good team player? Personality and team effectiveness. Group Dynamics: Theory, Research, and Practice, 10(4), 249.

[10] Holen, A., Manandhar, K., Pant, D. S., Karmacharya, B. M., Olson, L. M., & Koju, R. (2015). Medical students' preferences for problem-based learning in relation to culture and personality: a multicultural study. International journal of medical education, 6, 84.

[11] Ferreira, T. B., Buiar, J. A., Fernandes, M. A., Pimentel, A. R., & Oliveira, L. S. (2020). Regras para formação de grupos de colaboração utilizando detecção automática de traços de personalidade. Revista Brasileira de Informática na Educação, 28, 273-296.

[12] Henry, H. R., Tawfik, A. A., Jonassen, D. H., Winholtz, R. A., & Khanna, S. (2012). “I know this is supposed to be more like the real world, but...”: student perceptions of a PBL implementation in an undergraduate materials science course. Interdisciplinary Journal of Problem-based Learning, 6(1), 43-81.

[13] Reis, R. C. D., Lyra, K. T., Reis, C. D. G., Penteado, B. E., & Isotani, S. (2020). The use of personality traits to enhance theory-driven group formation. Revista Brasileira de Informática na Educação-RBIE, 28, 796-818.

[14] Isotani, S., Inaba, A., Ikeda, M., & Mizoguchi, R. (2009). An ontology engineering approach to the realization of theory-driven group formation. International Journal of Computer-Supported Collaborative Learning, 4, 445-478.

[15] Chen, C. M., & Kuo, C. H. (2019). An optimized group formation scheme to promote collaborative problem-based learning. Computers & Education, 133, 94-115.

[16] Weringa, J. R. “Design Science Methodology for Information Systems and Software Engineering”. Springer, 2014.

[17] Montequín, V. R., Balsera, J. V., Fernández, J. M. M., & Juez, J. D. C. (2010, April). Using Myers-Briggs Type Indicator (MBTI) for assessment success of student groups in project based learning. In International Conference on Computer Supported Education (Vol. 2, pp. 156-160). SCITEPRESS.

[18] Michailidou, A., & Economides, A. (2008). Gender and diversity in collaborative virtual teams. In Computer-Supported Collaborative Learning: Best practices and principles for instructors (pp. 199-224). IGI Global.

[19] Ardaiz-Villanueva, O., Nicuesa-Chacón, X., Brene-Artazcoz, O., de Acedo Lizarraga, M. L. S., & de Acedo Baquedano, M. T. S. (2011). Evaluation of computer tools for idea generation and team formation in project-based learning. Computers & Education, 56(3), 700-711.

[20] Kobusingye, C. (2006). The role of personality in developing high performance teams case study: faculty of education: Kampala International University (Doctoral dissertation, Kampala International University, College of Economics and Management).