UNIVERSITY OF GLOUCESTERSHIRE

Maqtary, Naseebah, Mohsen, Abdulqader and Bechkoum, Kamal ORCID logoORCID: https://orcid.org/0000-0001-5857-2763 (2019) Group Formation Techniques in Computer-Supported Collaborative Learning: A Systematic Literature Review. Technology, Knowledge and Learning, 24 (2). pp. 169-190. doi:10.1007/s10758-017-9332-1

Abstract Group formation (GF) is an essential process for group development lifecycle. It has been a growing concern to many researchers to be applied automatically in collaborative learning contexts. Forming a group is an atomic process that is affected by various factors. These factors differ depending on the group members characteristics, the context of the grouping process or the techniques used to form the group(s). This paper surveys the recently published work in group formation process by providing a systematic literature review (SLR) in which 30 relevant studies were analyzed. The findings of this review propose two taxonomies. The first one is for the attributes of group formation while the second is for the grouping techniques. Furthermore, we present the main findings and highlight the limitations of existing approaches in computer supported collaborative learning (CSCL) environment. We suggest some potential directions for future research with group formation process in both theoretical and practical aspects. In addition, We emphasize other improvements that may be inter-related with other computing areas such as cloud computing, mobility, etc.

Keywords Group formation (GF), Attributes of group formation, Group formation techniques, Computer supported collaborative learning (CSCL), Systematic literature review

1 Introduction

Education has improved smoothly through developing various approaches and technologies Resta and Laferrière (2007); Stahl et al (2006). It has been upgraded from the individual learning paradigm to collaborative learning where learners can gain more knowledge and skills through learning together from the same learning situation Matazi et al (2014); Resta and Laferrière (2007); Srba and Bielikova (2015); Stahl et al (2006).

Collaborative learning is defined by Rowe et al (2010) as an instructional method that is used by a group of learners to achieve a common goal. This type of learning is conceivably executed through a three dimensional model along the following axes: i) a group of people either in pairs or more; ii) a credible material of learning, (i.e. course content, activity or life-long work experience); and iii) a way of learning through collaborative interaction among group members Dillenbourg (1999).

The environment of collaborative learning is either real or virtual Dillenbourg (1999). Collaborative learning is performed through face to face conversations and meetings or online using computer tools and frameworks Dillenbourg (1999); Resta and Laferrière (2007); Stahl et al (2006) such as computer-supported collaborative learning (CSCL) Matazi et al (2014); Rowe et al (2010); Srba and Bielikova (2015); Stahl et al (2006). CSCL is a pedagogical approach that uses networking technologies to aid the social and instructional interaction among learners in small groups and learning communities Resta and Laferrière (2007); Rowe et al (2010); Stahl et al (2006). It employs generic tools such as e-mail, file attachments, electronic bulletin boards, chat, blogs, and digital audio and videoconferencing systems. Furthermore, it uses specific tools such as asynchronous/synchronous communication tools of Web-based Instructional Management Systems (Course Management System, CMS; Learning Management System, LMS), and virtual learning environments (Blackboard/WebCT, Moodle, Sakai, Claroline, FirstClass) Resta and Laferrière (2007); Stahl et al (2006).

CSCL has emerged during the mid-1990s. As shown previously, various tools have been used and employed to merge collaboration within educational activities Stahl et al (2006). Focusing on collaborative learning has brought groupwork to the fore. Many studies in CSCL environment have been carried out on administrating groupwork activities like group formation (GF), monitoring and evaluation Sun (2013).

Forming a group that collaboratively learns is one of the most challenging tasks in CSCLs context which is attracting the interest of several researchers Amara et al (2016); Khandaker et al (2006); Srba and Bielikova (2015). Many articles were discussing group formation from different aspects. These aspects are mostly discussing group development life cycle Abnar et al (2012); Sun (2013), optimizing the process of group formation Ho et al (2009); Zheng and Pinkwart (2014), or discovering the attributes that optimally affect group formation Graf and Bekele (2006); Yannibelli and Amandi (2011).

Research Objectives, Questions and Structure

Notwithstanding the valuable contributions thus far, there is still no rounded overview of the group formation process with its various effective components like attributes, and the techniques used. In addition, no surveys were found on group formation process. This led us to hold a systematic literature review (SLR) about group formation. This review aims at summarizing the various previous work and reproducing these contributions in an organized manner through using a systematic approach as discussed in the Section 2. Our contribution will present the techniques and attributes of group formation with different classified views. The attributes of group formation are different and chosen differently according to the grouping context. In addition, computerized tools perform the group formation process based on different techniques that are examined in specific grouping environments. Thus, the objectives of this paper are :

1. To discover the recent contributions in group formation in CSCL contexts.
2. To explore the effective attributes and techniques in grouping process from different viewpoints.
3. To summarize and represent the findings in a structural manner.
4. To draw the knowledge gaps, challenges and opportunities of group formation.

To achieve these objectives we have formulated some research questions and synthesized the relevant studies to answer them through conducting SLR. These questions are as it follows:

– **RQ1:** What are the most effectively used attributes and techniques in group formation process?

– **RQ2:** How can the recent CSCL contributions be represented in group formation process within educational context?

– **RQ3:** What are the knowledge gaps and limitations in group formation?

– **RQ4:** What conclusions can be drawn from the existing studies?

    This paper is organized around the following subsections: firstly, methodology of investigating this review is discussed in Section 2. Then our findings are introduced in Section 3. After that, Section 4 presents discussion and encountered limitations. Finally, conclusions and future trends are addressed in Section 5.

2 Research Methodology

To conduct this review, a systematic approach was proposed by following the straightforward and simple logic as described in Okoli and Schabram (2010). This approach is used to construct a systematic literature review (SLR). It is consisted of five sequential steps which are illustrated in Figure 1. We follow up this approach and come up with various results as shown in Figure 2.
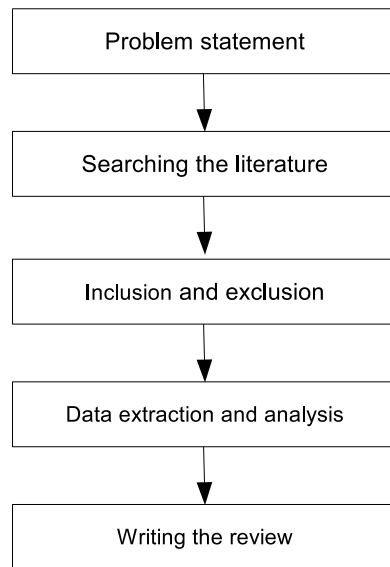
Fig. 1 The main steps of followed methodology for the review

2.1 Problem Statement

This step tackles with specifying the review aim and assigning the work protocol which are discussed as follows.

a. Specifying the purpose of the literature review
   Here, we describe the aim of this study which was discussed previously in Section 1. Briefly, the main purpose is to contribute a systematic literature review in group formation as this field lacked to such contributions.

Fig. 2 Results of the followed methodology

b. Assigning the protocol of work

Assigning the work protocol deals with formulating the research question which leads the researchers to reach the research objectives. It also comes up with the plan (work protocol) which will organize the methodology of accomplishing the review. As this review aims to contribute a systematic literature review in group formation process, we have formulated the research questions to be as described in Section 1.

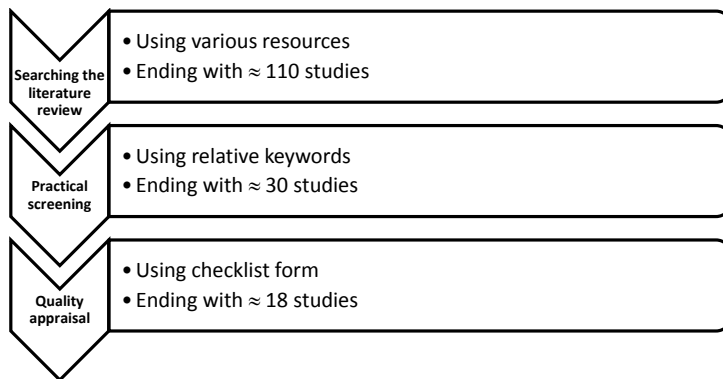The plan drawn by researchers begins with specifying the review aim, sketching the necessary steps of conducting the review and finalizing the output of the review to be ready for publishing. Within the process of setting the protocol, the researchers focus on the following issues:

– Concentrating on group formation process in CSCL systems.

– Searching the recent contributions in the specified field such that including various journal review, conferences papers, theses and dissertations specifically published between 2005 and 2015.

2.2 Searching the literature

Researchers should define the sources of researches and studies that will be used in their review. Therefore, we have investigated more than 110 studies using various available sources such as scholars search engines, reputable journals and conferences proceedings as shown in Table 1. These studies vary in their type and scope. They consist of journal reviews, conference papers, theses and book chapters.

Statistically, conference papers are the most used source in this review. They are about 58% of the total used sources while journal articles and reviews are approximately 40%. Book chapters are the least used sources which present only 2% as illustrated in Figure 3. The researchers used many keywords for searching like collaborative learning, computer supported collaborative learning, group, group formation, coalition formation and group development.

Table 1 Main sources used in this study including conferences, journals and others

Conferences	Journals	Others
Adaptive Hypermedia and Collaborative Web-based Systems	Artificial Intelligence in Education	ACM Digital Library
Advanced Learning Technologies Collaborative-learning: Cognitive and Computational Approaches	Cambridge handbook of the learning sciences Advances in Web-Based Learning Computer Science Research and Application	The Learning and Skills Research Centre
Autonomous agents and multiagent systems	Computers & Education	ScienceDirect
Computational Intelligence and Informatics	Creativity and collaborative learning	Elsiver
Computational Science and Engineering	Educational Psychology Review	IEEE Explore
Computer Science and Information Engineering	Educational Technology & Society	
Forming and Maintaining Coalitions and Teams in Adaptive Multiagent Systems	Frontiers in Artificial Intelligence and Applications	
Information Society	Group & Organization Management	
Intelligent Systems: Theories and Applications	Human Resource Development International	
Intelligent Tutoring Systems	JALT CALL Journal	
Interactive Collaborative Learning	Learning Technologies	
Recent Advances in Information Science: European Conference of Computer Science	Student Centered Learning	
Recent Trends in Information Technology		
Supporting group work: ACM		
Technology Enhanced Education		
Tools with Artificial Intelligence		
User Modeling		
Web Intelligence and Intelligent Agent Technology		

2.3 Inclusion and Exclusion Processes

Collected studies were filtered using multi-level criteria. These criteria were either for inclusion or exclusion as shown in Table 2. The inclusion process is called practical screening while exclusion process is called quality appraisal. These two processes are detailed as follows.

Table 2 Inclusion and exclusion criteria

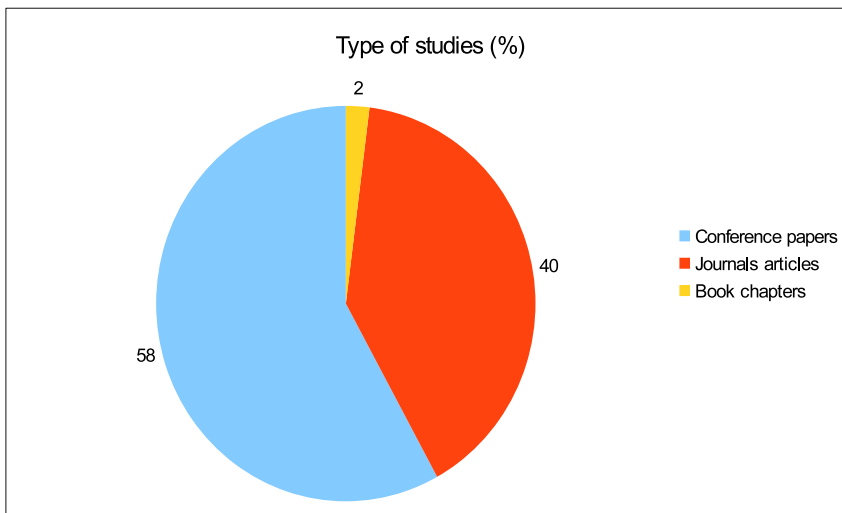| Inclusion	Exclusion
Papers published between 2005 and 2015	Nonacademic papers and gray literature such as reports, technical reports, and working papers.
Using relative keywords	Papers with weak analysis and writing
Prestige and academic papers and articles	Using checklist form

Fig. 3 Percentage of sources types used in the review

a. Practical Screening

This step deals with narrowing down the range of studies by reading each research abstract to ensure the strong relationship to the review aims and questions. Also, studies have been screened through choosing more relative keywords. As a result, the included studies are only those must discuss group formation process in CSCL environment (i.e. 30 studies).

b. Quality Appraisal

After specifying the related studies, another sorting level was executed to ensure strongly correlated studies to the proposed review. This subtask used the exclusion filter on chosen studies through trying to answer the checklist form with seven questions as shown in Figure 4. This checklist is consisted of seven questions. This step is concluded this step with 18 studies that positively answer the proposed checklist form.

2.4 Data Extraction and Analysis

The data required to build the review of group formation are identified to include the following information: type of grouping, number and type of grouping attributes, used technique, special features of grouping. These information were gathered from 18 studies in a tabular form to help us exploring group formation process as presented in Table 3. Then, gathered data should be synthesized using qualitative analysis, quantitative analysis or both. In our review and based on both quantitatively and qualitatively analyzing extracted data, taxonomic maps were used to reorder the ideas and extracted data from the group formation studies. This step comes up with two taxonomies which will be discussed in details in the Section 3.

No.	Question	Yes	No
1	Does the study apply the group formation in educational context?		
2	Does the study use a scientific approach in proving its contribution?		
3	Is the type of the group formation defined? (homogenous, heterogeneous or mixed)?		
4	Does the study identify specific attributes for group formation?		
5	Is the used technique clear and approved through experiments?		
6	Does the study use suitable datasets to prove the experiments' results?		
7	Does the study add valuable contribution to the state of the art of group formation?		

Fig. 4 Checklist form used to analyze the relevant studies in SLR on group formation

Table 3: Overview of extracted data from relevant studies in group formation

Author	GF Attributes					No. of Attributes	Techniques	Contributions
	Knowledge	Learning Style	Social Interaction	Team Role	Personality Traits			
Graf and Bekele (2006)	✓				✓	2	– Ant colony algorithm	– Heterogeneous groups – Introducing the measure called Goodness of Heterogeneity
Christodoulopoulos and Papanikolaou (2007)	✓	✓				2	– Fuzzy C-Means – Random selection algorithm	– Homogeneous and heterogeneous groups – Manual instructor interference – Equality is provided in group size
Soh and Khandaker (2007)	various					-	– Multi-agents	– Introduces I-MINDS system which is infrastructure of collaborative learning – Models the group formation problem as rules and policies

Table 3 – continued from previous page

Author	GF Attributes					No. of Attributes	Techniques	Contributions
	Knowledge	Learning Style	Social Interaction	Team Role	Personality Traits,			
Ounnas et al (2008a)			✓	✓	✓	3	– Semantic Web Ontologies	– Introduces a framework of group formation based on teacher preferences – Solves the problem of orphan student
Ho et al (2009)	✓	✓	✓			3	– Particle swarm optimization	– Using Student Feature Model to model attributes
Rubens et al (2009)	Informal (non-structured) data					-	– Mash-up technique	– Proposing a data-driven automatic group formation model and corresponding methods for mash-up, mining and inference utilizing data originating from multiple data sources
Hubscher (2010)	various					-	– Tabu search	– It forms groups after instructor assignment of the desired attributes
Yannibelli and Amandi (2011)				✓		1	– Knowledge base evolutionary algorithm	– Introducing different models of team roles, especially Belbins team role – First research dealing with team role

Table 3 – continued from previous page

Author	GF Attributes					No. of Attributes	Techniques	Contributions
	Knowledge	Learning Style	Social Interaction	Team Role	Personality Traits,			
Abnar et al (2012)		✓				1	– genetic algorithm – Greedy algorithm	– Homogeneous and heterogeneous grouping – Mixed groups are formed with equal size – Iterative process for satisfying the tutor about grouping
Mujkanovic et al (2012)					✓	1	– Regression analysis optimization	– Grouping is based on idea of no prior rules or data about members because they are from different environments – Using updating rules to adapt the approach of forming groups
Brauer and Schmidt (2012)	✓	✓	✓			3	– Genetic algorithm	– It concentrated on forming group after gathering attributes from OSN.
Moreno et al (2012)	various					-	– Genetic algorithm	– Formulating group problem into multi-objective optimization problem – Inter-homogeneous and intra-heterogeneous groups – It compares between 3 algorthims (random, exhaustive and proposed method)

Table 3 – continued from previous page

Author	GF Attributes					No. of Attributes	Techniques	Contributions
	Knowledge	Learning Style	Social Interaction	Team Role	Personality Traits,			
Tien et al (2013)	various					-	– Genetic algorithm with TOPSIS technique	– Inter-homogeneous and intra-heterogeneous groups – It compares between 3 algorthims (random, genetic and proposed method)
Jozan and Taghiyareh (2013)	various					-	– Genetic algorithm	– They used inversion concept to represent the priority of attributes of group members
Srba and Bielikova (2015)	✓					1	– Clustering with applying the theory of group technology	– Dynamic group with short term and iterative formation – Applying group technology
Zheng and Pinkwart (2014)					✓	1	– Discrete particle swarm optimization	– Heterogeneous groups – It compares between 3 algorithms (exhaustive method, random selection and DPSO based on time complexity and stability)
Amara et al (2016)	various					-	– K-means	– Held the whole process of group-work – Dynamic and customized group formation through MCSCL

2.5 Writing the Review

The final step of conducting systematic literature review is writing the review with its methodology, results and explanations in a scientific manner. The findings of this paper were reported as they were extracted based on a systematic approach. Next sections discuss the contributions founded in group formation studies in CSCLs. They concentrate on the techniques of grouping and the effective attributes.

3 Findings

The findings from reviewing 18 group formation studies in CSCL contexts are presented in this section. This section is organized in a structural manner. Thus, group formation process will be introduced and discussed with its effective attributes. Also, the relevant literature will be presented in a historical order so that most noticed features and contributions are clarified. Later, two taxonomies will be illustrated and discussed to represent group formation attributes and techniques, respectively.

3.1 Group Formation

As mentioned above, various research studies were dedicated to explore new provisions in group formation to make sure that all group members achieve the learning outcomes smoothly and easily Khandaker et al (2006). Group formation is the first process of the group development life cycle in which efforts should be devoted to ensure effectiveness and efficiency of the process Bonebright (2010). The group development life cycle is divided into many phases that translate the process of forming and monitoring the performance of the group. There are various models and theories about group development. The most known model that expresses group development life cycle is Tuckmans model, which has been frequently reviewed and extended by researchers. One example of these extensions is the Tuckman and Jensens model. It divides group development life cycle into five stages: forming, storming, norming, performing and adjourning Bonebright (2010); Srba and Bielikova (2015); Tuckman and Jensen (1977). First stage is forming the group, which considers introducing the backgrounds, preferences and experience of each group member to each other to form the first impression. So the group leader should be aware about group goals, roles and responsibilities to clarify them for the members Bonebright (2010). The next stage, storming, is dealing with setting the rules of group management to minimize, or better still avoid, conflict. The third stage, norming, is the phase during which agreement is reached about how members work together to maximize group performance and achievementBonebright (2010). The performing stage is about the group functioning towards the stated goals, which are finally evaluated in the adjourning stage Bonebright (2010).
As stated in the methodology, 18 studies were selected to explore the contributions in group formation process. These contributions are briefly discussed in Section 3.2.

3.2 Related Literature

This section reviews the state of the art in group formation area. It is written in a historical order to summarize the contributions of each study and its specific used technique. It gives a

brief discussion how this technique was executed within the context of group formation and what are the exactly used attributes as identified in Table 3.

Graf and Bekele used Ant colony optimization algorithm (ACO) to build heterogeneous groups. These groups were built according to members performance and personality traits. Researchers introduced a quality parameter called goodness of heterogeneity (GH) to measure the level of heterogeneity of the group members. They proved the scalability of the used strategy by iterative experiments with different group sizes Graf and Bekele (2006).

In Christodoulopoulos and Papanikolaou (2007), the authors implemented a web-based group formation tool. This tool has the ability of grouping members homogenously and heterogeneously. Researchers used a fuzzy c-Means algorithm to homogeneously assign a member to the most appropriate group. The appropriate group was selected by looking at the different probabilities of members belonging to different groups. In addition, the proposed strategy used the random selection algorithm in order to heterogeneously create groups. Knowledge level and learning style of the members were used as criteria of the grouping process.

Soh and khandaker presented a multi agent framework for group formation of students. They implemented VALCAM, an algorithm that groups students based on the idea of iterative auctions. This framework was applied in a computer supported collaborative learning environment called I-MINDS Soh and Khandaker (2007). I-MINDS is a distributed computing infrastructure that uses intelligent multiagent information system for education Soh (2004); Soh et al (2008). It reforms the problem of group formation through using rules and policies. These rules and policies are realized by agents while the process of group formation is executed Soh (2004); Soh and Khandaker (2007).

Another approach of group formation was presented by Ounnas et al (2007b) who have modeled a semantic framework to represent the interaction data of learners through using FOAF ontology. Moreover, researchers have introduced the term orphan student which means the left or unassigned student to a group Ounnas et al (2008b). They have used web semantic technologies and logic programming Ounnas et al (2008a).

Particle swarm optimization algorithm was the used technique in Ho et al (2009). Researchers have chosen social interaction, competences and learning style attributes to form the groups. They have involved the time complexity of the problem that considers more than one attribute while forming groups.

Some contexts of collaborative learning in informal environments lack prior information about learners. This issue attracted Rubens and his research team to propose a data-driven model for extracting information of learners from various data sources such as blogs, wikis, forums, etc. The extracted information was built in a mash-up way that led to automatic group formation through grouping the learners with sharable knowledge Rubens et al (2009).

Yannibelli and Amandi applied the evolutionary algorithm to form groups based on the attributes of teams roles. To form a well-balanced group with various team roles, the group formation process should be heterogeneous. Heterogeneity raises the time complexity of the problem, which was solved by using this algorithm. Researchers evaluated the result of formation process by ensuring that each outcome group has highly diverse team roles Yannibelli and Amandi (2011).

Brauer and Schmidt have proposed another approach to capture data about members attributes. They developed a graph model for modeling members data from online social networks (OSN). The data included attributes such as knowledge, learning style and social interaction. They used various graph traversal algorithms to capture the candidate members of groups. These candidate members were grouped through using a genetic algorithm which handles the group formation process execution and scalability Brauer and Schmidt (2012).

Genetic algorithms were also used to form heterogeneous groups in Sukstrienwong (2012). The researcher has modeled a fitness function with fairness and equity in terms of members performance to ensure the fair formation, which means each group has various knowledge levels of the members.

Moreno Moreno et al (2012) with his group suggested using genetic algorithms to form groups with multiple attributes. They have formulated the grouping problem into multi-objective optimization problem under combinatorial scenario. To validate the result, comparative study was done in contrast with exhaustive and random algorithms. They have stated that their proposed strategy has contributed better results in both computational and pedagogical directions.

A different approach was applied by Abnar et al (2012) where groups were formed iteratively to reach the continuously tuned fitness threshold of the genetic algorithm. This approach was featured by the flexible facility of using different attributes with ranking (prioritizing) them according to the group task. Same feature was handled in Hubscher (2010) so that group formation was iteratively adaptable to the context of the used strategy in teaching and learning. A new criterion called evenly skilled, which depends on reciprocal teaching method, was proposed to assign students to groups. Researcher has employed the tabu search algorithm to form groups because of its stable and systematic use of memory.

The situation of forming groups that have members who are geographically distributed and no prior rules are known about them were investigated in Mujkanovic et al (2012). The research approach used self-learning algorithm called regression analysis optimization that adapted rules of forming groups over time yielding a progressively improving group performance.

Hui-Wen Tien and his research group examined the effectiveness of group formation process through proposing new strategy that adopted genetic algorithm with TOPSIS technique. They formulated the fitness value to achieve the goal of obtaining inter-homogeneous and intra-heterogeneous groups. After comparison study between the random algorithm, genetic algorithms and the proposed strategy, researchers concluded that the proposed method recorded better grouping results than random and genetic algorithms against different number of characteristics Tien et al (2013).

A similar approach was conducted in Jozan and Taghiyareh (2013) where a genetic algorithm was reapplied to group formation process. Researchers also applied the idea of priority and weights of the members attributes with the concept of inversion. They evaluated the quality of the formed groups through inter-group fitness and intra-group fitness measures. The concluded remarks stated that the proposed strategy returned better results in group formation process with inter-group fitness criteria while it showed weak performance with intra-group fitness criteria.

Discrete particle swarm optimization algorithm (DPSO) was used to group heterogeneous learners according to their personality traits and gender in Zheng and Pinkwart (2014). It was also evaluated in contrast with exhaustive and random algorithms. The researchers argued that DPSO algorithm gave better performance and stability results in a reasonable time than other evaluated algorithms for group formation.

In Srba and Bielikova (2015), a group was formed through applying group technology (GT). It is a concept in manufacturing and engineering management sector, whereby its most applied task is cell manufacturing. Researchers proposed a novel method by using clustering algorithms to involve the collaboration feedback of the members dynamically and iteratively each time the group was formed. They simplified the group lifecycle to suit the short-term groups in virtual domains.

Recently, a research was applied to form homogeneous groups in mobile collaborative learn-

ing environment (MCSCL) Amara et al (2016). It held all the activities of groupwork with adding the dynamism of the group formation at any level of groupwork and customized selection of forming attributes. The research enabled the instructors to determine type, number and weight of grouping criteria. The used technique for group formation was K-means algorithm.

From the above discussion and literature we can notice that many of the issues raised in group formation process and still need scientific investigations by researchers. For example, the formation process was conducted under various contexts to cover all collaborative learning aspects. The group formation can be characterized by different parameters such as the group size, duration of the groupwork, the ideal method of grouping, the authority of grouping, type of the formed groups, etc. In addition, attention is needed to the educational and psychological characteristics of members involved in the grouping. These characteristics vary from one group to another if the group goal and tasks are different. In some of the previous work, groups were formed using member characteristics such as knowledge, skills or competences while in other work grouping were based on learning styles, personality traits or other characteristics. These characteristics are discussed in details and classified within a proposed taxonomy in Section 3.3 below.

3.3 Taxonomy of Group Formation Attributes

The group formation process can be achieved either manually or automatically Srba and Bielikova (2015). Manual formation of the group is either self-selection or instructor assignment Resta and Laferrière (2007); Srba and Bielikova (2015); Ounnas et al (2007a). In the self-selection approach the member has the right to choose the most suitable group for him/her. This approach does not guarantee a balanced grouping and thus violates the ideal group formation Abnar et al (2012); Zheng and Pinkwart (2014). The second approach is managed by the instructor decisions about which member will form part of which group Srba and Bielikova (2015). This kind of selection guarantees better results, in terms of a balanced grouping, but it is a fairly complex process when large numbers of members are grouped manually Mujkanovic et al (2012); Srba and Bielikova (2015). In order to assign members to groups automatically, there exist many CSCLs environments that offer the option of creating groups automatically with or without human intervention Abnar et al (2012). Random selection is one way of achieving group formation automatically Srba and Bielikova (2015). Other approaches are used to form groups according to the context of the group.

During the formation of groups, various attributes are taken into consideration to ensure that groups will achieve their goals Coffield et al (2004). These attributes can be categorized into two classes: member attributes and group attributes. Member attributes are the attributes that describe the who person will be included in groups while the group attributes describe the group characteristics as a whole.

3.3.1 Member Attributes

Examples of member attributes are knowledge, skills, learning styles and personality traits. They are used to decide about the most suitable group for each member Abnar et al (2012); Graf and Bekele (2006). Other member attributes such as social interaction and team roles are also considered in some situations of group formation Yannibelli and Amandi (2011). Different studies are conducted that tackle at grouping students based on their knowledge

and skills for a specific task or assignment. These attributes are measured through Likert scale stated by the instructor himself according to his knowledge about students Graf and Bekele (2006); Ho et al (2009) or collected from different learning management systems which store student information and their academic progress Brauer and Schmidt (2012). Learning styles or personality traits are arbitrarily used in different works to check its effectiveness for group formation Abnar et al (2012); Brauer and Schmidt (2012); Christodoulopoulos and Papanikolaou (2007); Ho et al (2009); Martin and Paredes Barragán (2004). There is a strong intuitive appeal to consider learning styles as an indicator of the speed, manner and confidence of picking information and data Coffield et al (2004); Martin and Paredes Barragán (2004). There are various models of learning styles such as Kolbs learning style Inventory (LSI),Herrmann whole brain model (HDBI) and Myers-Briggs Type Indicator (MBTI) Coffield et al (2004). To capture the learning style of a group member, a questionnaire is run and then an index is used to specify the learning style of the member after classifying his responses to the questionnaire Coffield et al (2004); Martin and Paredes Barragán (2004).

Some researchers used the personal information of the group members, such as age, gender, IQ, race etc., as an attribute that affects the group formation process Graf and Bekele (2006); Mujkanovic et al (2012); Ounnas et al (2008b); Zheng and Pinkwart (2014).

Social interaction is another attribute that attracts attentions of psychological and educational researchers. They agreed that social interaction and negotiation contribute to the way with which people learn how to develop shared understanding about certain concepts or tasks Kreijns et al (2002). This attribute consists of different social skills that learners should have while working collaboratively within the group. These social skills are participation, social grounding, active learning conversation skills, performance analysis and group processing and promotive interaction Soller (2001).

Members within a group should play different roles according to the mission of the group and their behavior. A role is the way a person is expected to behave, contribute and interrelate with others throughout collaborative work. Several team roles models were proposed and investigated in the literature about group formation. These models are based on the concept of well-balanced groups, which should be formed with members having heterogeneous roles. The most known team role model is Belbins which is applied in training activities by different organizations, consulting firms and executive education programs Yannibelli and Amandi (2011).

3.3.2 Group Attributes

The group attributes relate to the context of the group goal and task. For example, the nature of the task identifies the homogeneity of the group members. Accomplishing some tasks need homogeneous characteristics of the group members while other tasks need diverse characteristics of group members in order to force the tasks to be completed Christodoulopoulos and Papanikolaou (2007); Srba and Bielikova (2015). In addition, the duration of completing a group task is another attribute which affects the group formation process. Thus, there are short term and long term groups Huang et al (2009); Srba and Bielikova (2015). Moreover, the process of assigning members can be static for the duration of the task or can benefit from previous information about group members and their abilities in accomplishing group tasks. The latter type of formation is called dynamic or adaptive formation Mujkanovic et al (2012); Srba and Bielikova (2015).

All the aforementioned review of the group formation attributes are reorganized and classified within a proposed taxonomy that reflects the above presentation. Figure 5 depicts this
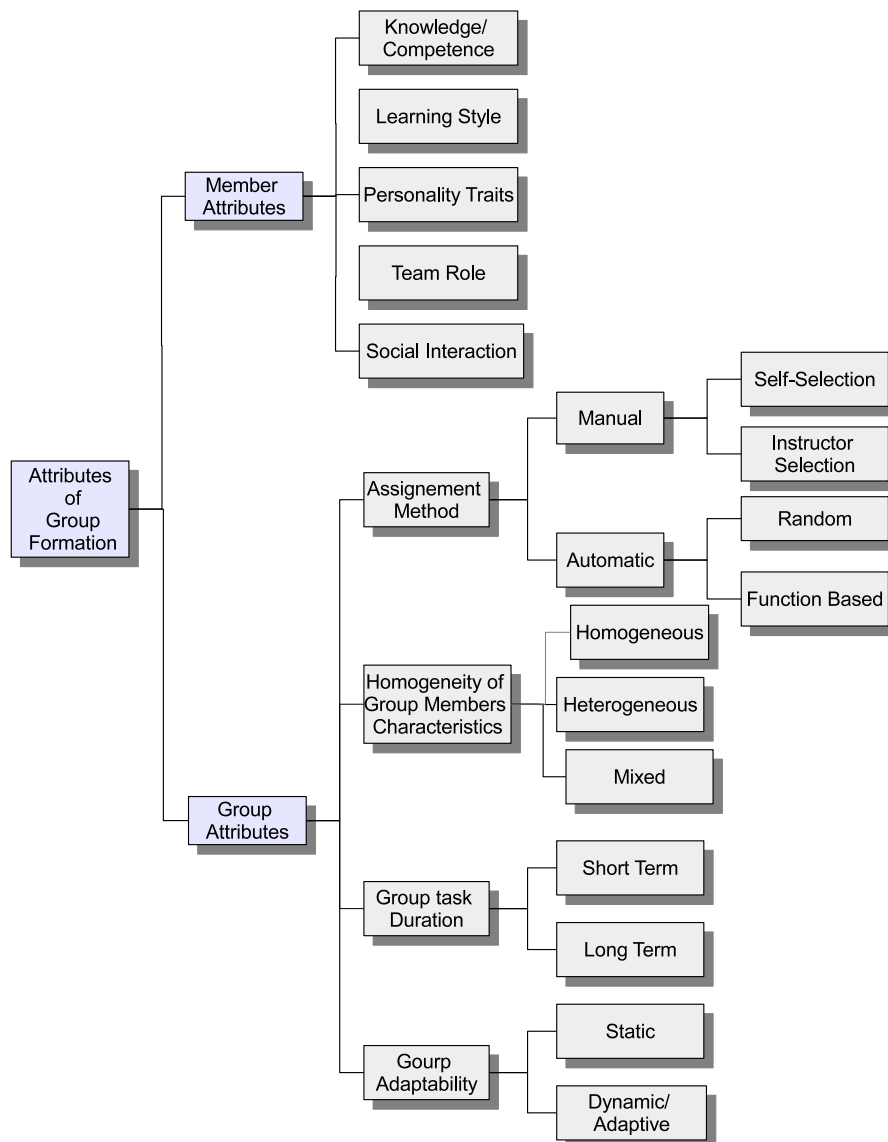
Fig. 5 Taxonomy of group formation attributes

taxonomy, which has a multi-level categorization. Attributes are divided into two groups: i) member attributes and ii) group attributes. Member attributes are clarified into five different attributes. Group attributes are also classified into four subcategories which are: i) assignment method, ii) homogeneity of the group members' characteristics, ii) group duration and iv) adaptability of the group. Also a tabular mapping is used to map each study with its specified grouping attributes to facilitate grasping information about studies contributions. This mapping is presented in Table 3.

3.4 Taxonomy of Group Formation Techniques

The literature on group formation is quite rich and publications range from gathering data of group members to apply various models on group formation process with different perspectives in various contexts. This section will highlight the distinction between reviewed studies under different points of view as illustrated in the taxonomy shown in figure 6.

Researches made by Brauer and Schmidt (2012); Rubens et al (2009) were dealing with data-driven models to collect data about learners from different environments such as online social networks (OSN). In almost all the work reviewed, the group formation process is discussed in various circumstances.

Some researches depended on either single attribute to form a collaborative group such as Abnar et al (2012); Mujkanovic et al (2012); Srba and Bielikova (2015); Sukstrienwong (2012); Yannibelli and Amandi (2011); Zheng and Pinkwart (2014) or multiple attributes as in Brauer and Schmidt (2012); Christodoulopoulos and Papanikolaou (2007); Graf and Bekele (2006); Ho et al (2009); Ounnas et al (2008a). It is obvious that the more attributes of group members, the more complex is the process of group formation.

In addition, various techniques were used to prove the experiments of forming groups such as evolutionary approach Abnar et al (2012); Brauer and Schmidt (2012); Jozan and Taghiyareh (2013); Moreno et al (2012); Sukstrienwong (2012); Tien et al (2013); Yannibelli and Amandi (2011), swarm techniques Graf and Bekele (2006); Ho et al (2009); Zheng and Pinkwart (2014), clustering algorithms Amara et al (2016); Christodoulopoulos and Papanikolaou (2007); Srba and Bielikova (2015), semantic ontologies Ounnas et al (2008b) or multi-agent Soh (2004).

Moreover, a wide view of recent researches in group formation shows that different approaches were applied to form effective group. These researches can be categorized into three classes based on research goal:

Finding out suitable attributes that affect group formation and achievement Yannibelli and Amandi (2011).

Trying to optimize the process of group formation through using optimization techniques, so that better and faster formation will be achieved Abnar et al (2012); Graf and Bekele (2006); Jozan and Taghiyareh (2013); Zheng and Pinkwart (2014).

Applying new technologies from different scientific area, for example, using the manufacturing theory called group technology, GT Srba and Bielikova (2015).

Many researchers merged more than one goal, for instance, a research may conclude with applying suitable attributes with iterative experiments to obtain the optimal solution Abnar et al (2012); Tien et al (2013); Zheng and Pinkwart (2014).

Looking more deeply into previous work on group formation, it becomes apparent that there are two approaches of formation which compare the similarity of the group members characteristics: either clustering the homogeneous characteristics or constrained optimization for heterogeneous and mixed characteristics Christodoulopoulos and Papanikolaou (2007); Hubscher (2010).

Figure 6 presents a proposed taxonomy on classification of the techniques, which are reviewed and discussed previously. This taxonomy demonstrates various criteria to classify techniques from different perspectives. As shown in the figure, the applied techniques can be classified based on the final formation of the group, which are homogeneous, heterogeneous or mixed. Also these techniques differ in the point of data representation according to the context of the problem. The problem of group formation process has been formulated by different models in order to solve it optimally. Surveyed researches modeled the GF problem through using agents, semantic networks, or graphs. Clearly, these works may be classified
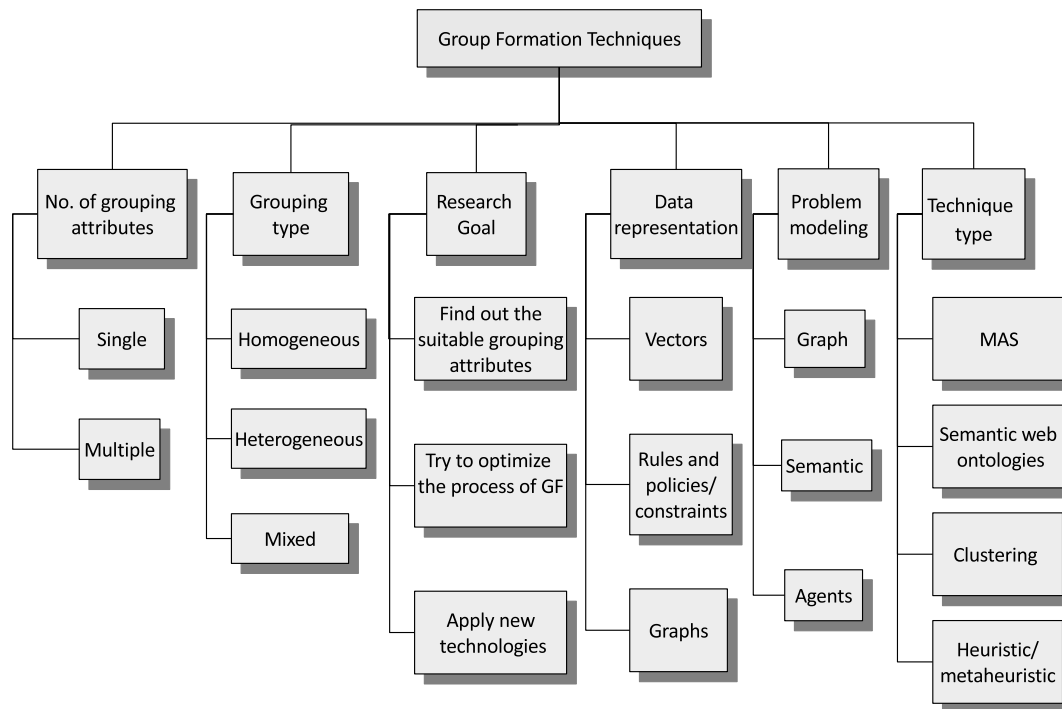
Fig. 6 Taxonomy of surveyed GF techniques

through examining the type of the technique used. The type of technique varies from context to another. Thus, technique could be based on heuristic/metaheuristic, multi-agents, clustering or semantic ontologies.

4 Discussion and Future Trends

This section will discuss the obtained results from this SLR. It has reviewed 18 studies in group formation area to achieve the research aims. The aims were: discovering the recent contributions in group formation in CSCL contexts, exploring the effective attributes and techniques on grouping process, summarizing and representing our findings in a structural manner and drawing knowledge gaps, challenges and opportunities.

It is obvious that group formation process has been investigated from two important perspectives. First one is the attributes that affect the grouping process and the second is about techniques used in CSCL contexts. These perspectives will be discussed next in details.

The obtained results after studies analysis have led us to classify the contributions according to various viewpoints. Our classification comes up with two novel taxonomic maps to represent both perspectives. These maps were presented in the Section 3 and briefly will be discussed within group formation perspectives.

Finally, challenges and opportunities will be summarized based on the knowledge gaps found in the related literature.

4.1 Attributes of Group Formation

As illustrated in Table 3, group formation process depends on the chosen attributes in each study. It is clear that studies differed in their chosen attributes and their number. Five of them used only single attribute to form groups such as Abnar et al (2012); Sukstrienwong (2012); Yannibelli and Amandi (2011) while two studies used two attributes in their group formation Christodoulopoulos and Papanikolaou (2007); Graf and Bekele (2006). On the other hand, there were 3 studies out of the 18 used 3 attributes Brauer and Schmidt (2012); Ho et al (2009); Ounnas et al (2008a). These studies' aims were to explore the capability of forming group with multiple criteria in more complex situations. The rest of studies used other attributes depending on their available and used data sets Amara et al (2016); Hubscher (2010); Jozan and Taghiyareh (2013); Moreno et al (2012).

The most common used attribute was the knowledge. It was used in 6 studies Brauer and Schmidt (2012); Christodoulopoulos and Papanikolaou (2007); Graf and Bekele (2006); Ho et al (2009); Srba and Bielikova (2015); Sukstrienwong (2012). As we understand, it is the most suitable and important attribute to form educational groups because of its effects on the group outcomes.

In addition, the attributes learning styles and personality traits were discussed similarly. Each one was used in 4 studies Abnar et al (2012); Christodoulopoulos and Papanikolaou (2007); Mujkanovic et al (2012); Ounnas et al (2008b); Zheng and Pinkwart (2014). Educationally, learning style is an effective attribute that play vital role in students learning and subsequently affects the grouping process. While personality traits are the simplest gathered attributes for experiments and studies.

The recent approach has added new attributes. It focused on the learners relationships and their roles within teams. Thus, there were three studies used social interactions as an attribute to form groups Brauer and Schmidt (2012); Ho et al (2009); Ounnas et al (2008a) while other two studies used the team role attribute Ounnas et al (2008a); Yannibelli and Amandi (2011).

The analyzed attributes directed us to represent them using a novel taxonomic map according to their classification and relationship to group formation process. Figure 5 shows this classification and reflects our findings.

According to aforementioned discussion, we think that the more flexible system that offers many choices for the instructor to form his groups, the better grouping process will done.

4.2 Techniques Used in Group Formation

As seen in Table 3, it is Commonly and obviously that evolutionary algorithms were used in group formation process Abnar et al (2012); Brauer and Schmidt (2012); Graf and Bekele (2006); Ho et al (2009); Jozan and Taghiyareh (2013); Moreno et al (2012); Tien et al (2013); Yannibelli and Amandi (2011). They were about 60% of the reviewed studies. For example, genetic algorithm was the dominant used technique. It was used in more than 5 studies Brauer and Schmidt (2012); Jozan and Taghiyareh (2013); Moreno et al (2012); Sukstrienwong (2012); Tien et al (2013).

Clustering algorithms were also frequently used, especially for homogenous grouping. Approximately, four studies used these algorithms Amara et al (2016); Christodoulopoulos and Papanikolaou (2007); Mujkanovic et al (2012); Srba and Bielikova (2015). While other studies varied in their techniques because of their aims (e.g. trying new approach for group-

ing). Soh and Khandaker (2007), Ounnas et al (2008b) and Srba and Bielikova (2015) used multiagents, semantic ontologies and group technology, respectively.

4.3 Opportunities and Challenges of Group Formation

Based on the related literature, trends are observed in the area of group formation and it is obvious that the reviewed work covered the automated group formation process from the viewpoint of education and collaboration learning. It is worthwhile to try reapplying this process in other contexts, which need people working in groups. Such contexts are training, business, psychology, etc.
However, there are still many issues that are not sufficiently discussed and solved in group formation. These issues are gaps and shortcomings in the reviewed literature. They are used as a basis for defining the directions for future investigations, as follows:

1. Choosing specific technique to form a group is oriented by the context of the group formation problem. However, in some situations there are many suitable techniques to be applied in the group formation process. This concept leads us to ask about the reason behind applying the chosen technique, which is not clear in some of the previous work.

2. Using local datasets for evaluating the proposed strategies may be a shortcoming in the situation of comprehensive comparison among the computational and pedagogical results of the applied group formation techniques.

3. A comprehensive paradigm that expresses all the details of group formation process in different situations should be developed in conjunction with other disciplines such as education, training, psychology, etc.

4. Incomplete solutions were introduced to solve the group formation in various contexts. Thus, there is a need to develop an autonomous system that holds main grouping operations and learning preferences in the field of group formation.

5. Poor contributions in the field of quality metrics that measure the quality of group formation process from different viewpoints. So, quality of service (QoS) as an evaluation framework for group formation is a fertile area to deal with comprehensively for identifying the success of group formation process.

6. Mobile and cloudy environments are the new trends for developed systems. Thus, incorporating group formation process within these environments would offer facilities to various disciplines.

5 Conclusions and Future Work

Automated group formation process has become an important issue in terms of talking about collaborative learning. The state of the art of group formation has shown the various and widespread works from different viewpoints. In this study, a systematic literature review has been introduced on group formation process in CSCL contexts. It highlights the attributes affecting the process of group formation through presenting taxonomy of these attributes

which were categorized according to different criteria. The techniques of group formation were reorganized in another proposed taxonomy which tackles the way of grouping, the problem definition, the data representation and the type of the applied technique. These taxonomies were constructed to answer our research questions: What are the most effectively used attributes and techniques in group formation process? How can the recent CSCL contributions be represented in group formation process within educational context? What are the knowledge gaps and limitations in group formation?

Through our survey, group formation process still needs more improvements to be an ideal process in CSCL environment. Some of these improvements are directly related to the process itself. For instance, obtaining complete solutions with optimal performance is the most critical issue. On the other hand, other improvements may be inter-related with other computing areas such as cloud computing, mobility, etc.

References

Abnar S, Orooji F, Taghiyareh F (2012) An evolutionary algorithm for forming mixed groups of learners in web based collaborative learning environments. In: Technology Enhanced Education (ICTEE), 2012 IEEE International Conference on, IEEE, pp 1–6

Amara S, Macedo J, Bendella F, Santos A (2016) Group formation in mobile computer supported collaborative learning contexts: A systematic literature review. Journal of Educational Technology & Society 19(2):258–273

Bonebright DA (2010) 40 years of storming: a historical review of tuckman's model of small group development. Human Resource Development International 13(1):111–120

Brauer S, Schmidt TC (2012) Group formation in elearning-enabled online social networks. In: Interactive Collaborative Learning (ICL), 2012 15th International Conference on, IEEE, pp 1–8

Christodoulopoulos CE, Papanikolaou KA (2007) A group formation tool in an e-learning context. In: 19th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2007), IEEE, vol 2, pp 117–123

Coffield F, Moseley D, Hall E, Ecclestone K (2004) Learning styles and pedagogy in post 16 learning: a systematic and critical review. The Learning and Skills Research Centre

Dillenbourg P (1999) What do you mean by collaborative learning. Collaborative-learning: Cognitive and computational approaches 1:1–15

Graf S, Bekele R (2006) Forming heterogeneous groups for intelligent collaborative learning systems with ant colony optimization. In: International Conference on Intelligent Tutoring Systems, Springer, pp 217–226

Ho TF, Shyu SJ, Wang FH, Li CTJ (2009) Composing high-heterogeneous and high-interaction groups in collaborative learning with particle swarm optimization. In: Computer Science and Information Engineering, 2009 WRI World Congress on, IEEE, vol 4, pp 607–611

Huang Y, Zhu M, Wang J, Pathak N, Shen C, Keegan B, Williams D, Contractor N (2009) The formation of task-oriented groups: Exploring combat activities in online games. In: Computational Science and Engineering, 2009. CSE'09. International Conference on, IEEE, vol 4, pp 122–127

Hubscher R (2010) Assigning students to groups using general and context-specific criteria. IEEE Transactions on Learning Technologies 3(3):178–189

Jozan MMB, Taghiyareh F (2013) An evolutionary algorithm for homogeneous grouping to enhance web-based collaborative learning. International Journal of Computer Science Research and Application

Khandaker N, Soh LK, Jiang H (2006) Student learning and team formation in a structured cscl environment. FRONTIERS IN ARTIFICIAL INTELLIGENCE AND APPLICATIONS 151:185

Kreijns K, Kirschner PA, Jochems W (2002) The sociability of computer-supported collaborative learning environments. Educational Technology & Society 5(1):8–22

Martin E, Paredes Barragán P (2004) Using learning styles for dynamic group formation in adaptive collaborative hypermedia systems

Matazi I, Messoussi R, Bennane A (2014) The design of an intelligent multi-agent system for supporting collaborative learning. In: Intelligent Systems: Theories and Applications (SITA-14), 2014 9th International Conference on, IEEE, pp 1–8

Moreno J, Ovalle DA, Vicari RM (2012) A genetic algorithm approach for group formation in collaborative learning considering multiple student characteristics. Computers & Education 58(1):560–569

Mujkanovic A, Lowe D, Willey K, Guetl C (2012) Unsupervised learning algorithm for adaptive group formation: Collaborative learning support in remotely accessible laboratories. In: Information Society (i-Society), 2012 International Conference on, IEEE, pp 50–57

Okoli C, Schabram K (2010) A guide to conducting a systematic literature review of information systems research. Sprouts Work Pap Inf Syst 10:26

Ounnas A, Davis HC, Millard DE (2007a) Semantic modeling for group formation

Ounnas A, Davis HC, Millard DE (2007b) Towards semantic group formation

Ounnas A, Davis H, Millard D (2008a) A framework for semantic group formation. In: 2008 Eighth IEEE International Conference on Advanced Learning Technologies, IEEE, pp 34–38

Ounnas A, Davis H, Millard D (2008b) Semantic web-based group formation for e-learning

Resta P, Laferrière T (2007) Technology in support of collaborative learning. Educational Psychology Review 19(1):65–83

Rowe JP, Shores LR, Mott BW, Lester JC (2010) Integrating learning and engagement in narrative-centered learning environments. In: International Conference on Intelligent Tutoring Systems, Springer, pp 166–177

Rubens N, Vilenius M, Okamoto T (2009) Automatic group formation for informal collaborative learning. In: Proceedings of the 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology-Volume 03, IEEE Computer Society, pp 231–234

Soh LK (2004) On cooperative learning teams for multiagent team formation. In: of the AAAIs 2004 Workshop on Forming and Maintaining Coalitions and Teams in Adaptive Multiagent Systems

Soh LK, Khandaker N (2007) Forming and scaffolding human coalitions with a multi-agent framework. In: Proceedings of the 6th international joint conference on Autonomous agents and multiagent systems, ACM, p 62

Soh LK, Khandaker N, Jiang H (2008) I-minds: a multiagent system for intelligent computer-supported collaborative learning and classroom management. International Journal of Artificial Intelligence in Education 18(2):119–151

Soller A (2001) Supporting social interaction in an intelligent collaborative learning system. International Journal of Artificial Intelligence in Education (IJAIED) 12:40–62

Srba I, Bielikova M (2015) Dynamic group formation as an approach to collaborative learning support. IEEE Transactions on Learning Technologies 8(2):173–186

Stahl G, Koschmann T, Suthers D (2006) Computer-supported collaborative learning: An historical perspective. Cambridge handbook of the learning sciences 2006:409–426

Sukstrienwong A (2012) Genetic algorithm for forming student groups based on heterogeneous grouping. In: Recent Advances in Information Science: Proceedings of the 3rd European Conference of Computer Science, pp 92–97

Sun G (2013) Teamwork as a service: design of a cloud-based system for enhancing teamwork performance in mobile learning

Tien HW, Lin YS, Chang YC, Chu CP (2013) A genetic algorithm-based multiple characteristics grouping strategy for collaborative learning. In: International Conference on Web-Based Learning, Springer, pp 11–22

Tuckman BW, Jensen MAC (1977) Stages of small-group development revisited. Group & Organization Management 2(4):419–427

Yannibelli V, Amandi A (2011) Forming well-balanced collaborative learning teams according to the roles of their members: An evolutionary approach. In: Computational Intelligence and Informatics (CINTI), 2011 IEEE 12th International Symposium on, IEEE, pp 265–270

Zheng Z, Pinkwart N (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