Hochschule für Angewandte Wissenschaften Hamburg
Hamburg University of Applied Sciences

Anwendungen 2 SoSe 2012

Steffen Brauer

Group Formation in eLearning-enabled Online Social Networks

Fakultät Technik und Informatik
Department Informatik

Faculty of Engineering and Computer Science
Department of Computer Science

Steffen Brauer

Group Formation in eLearning-enabled Online Social Networks

Eingereicht am: 31. August 2012

Contents

1 Introduction **2**

2 Related Work **3**

2.1 Social Network-Based Course Material Transformation 3

2.1.1 Approach 3

2.1.2 Discussion 4

2.2 Multi-Objective Team Composition 5

2.2.1 Methodology 5

2.2.2 Discussion 6

2.3 Searching for Expertise in Social Networks 6

2.3.1 Results 8

2.3.2 Discussion 9

3 Approach **10**

4 Conclusion and Outlook **12**

1 Introduction

Nowadays education does not stop by finishing school or university. Lifelong, self-paced learning is needed to stay up-to-date in a specific field. This makes it necessary to embed the learning process in the daily life of the learner, even if there is high workload, learning have to be done ubiquitously whenever it is possible. These requirements cannot be satisfied by the current educational system, including current eLearning environments. eLearning systems are often based on a classroom concept moderated by an instructor, who selects content, builds groups for collaboration and manages the learning progress. Even if these eLearning environments are approved they are limited in scalability in the technical as in the educational meaning.

Online Social Networks (OSN) offer a scalable way to socialize, communicate and share content among user. This happens in a self-organized way without any moderator or instructor. By designing a social network for eLearning, the instructor's tasks have to be achieved by the system including 1) how to provide content that suits the user's needs 2) How to guarantee a consistent learning progress and 3) How to build groups for collaborative work. Learning in groups creates motivation for a user by offering social interaction. This paper concentrate on how to form a group for collaborative learning in an eLearning-enabled OSN.

There are many of possible factors which can influence the quality of a learning group. Often criteria like knowledge, learning style are taken into account, but via OSN it is also possible to account for the relationships between the users by build groups. In this context, possible metrics are the weights of an edge between two users indicating former collaboration [1] or creating a representation of trust between users [2]. In our approach we concentrate on availability, knowledge, learning style and the group density in the social network by forming a group for collaborative learning.

To handle the problem of group formation easier, our approach is divided into two parts. First we browse the social network and try to find a given number of suitable candidates for group formation based on an initiator and a chosen topic. Based on the candidates the second part tries to find the best candidate constellation for a successful group learning experience.

The remainder of the paper is organized as follows: The next section covers an overview of related work and presents three works in-deep. The third section introduces our approach of group formation in eLearning enabled Online Social Networks and the last section gives an outlook and conclusion.

2 Related Work

Group formation in eLearning environments is a field with a lot of research done. Current approaches are often based on the classical eLearning model of a classroom and an instructor. In this context, the approach of Ounas et al [3],[4] is popular. The authors stress the group formation problem as a constraint-satisfying-task, which is solved by anthologies and given rules, that define the desired group constellation. Given our focus on OSN, the field of social learning is interesting. Vassileva [2] shows future trends of eLearning by focusing on the needs of digital natives. She gives two implications for the design of social learning environments. The first implication is the learner-centric nature of social learning environments. That means that decisions and adoptions done by the system should be invisible to the user to maintain the experience of control to the user. The requirements described in this context cover our problems by removing an instructor from an eLearning environment. She stresses the problem of how to find the right content for a user and how to find people, the user can collaborate with. The second implication for the design of social learning software is that learning should be more gratifying. This point will be considered in our future work.

The rest of this section focuses on presenting and discussing three related publications.

2.1 Social Network-Based Course Material Transformation

Arndt and Guercio [5] introduce a course material transformator, which generates eLearning content based on learning style and on social network data of the students. Goal of this approach is a shared ubiquitous eLearning experience.

2.1.1 Approach

The students in a class are modelled as vertices in an undirected graph G. The edges between the vertices describe friendship relations. Each vertex in G has an associated vector, which contains the learning preferences for a student. The result of the group formation progress is, that all members of the class are assign to a group. Given a minimum size of members in a learning group the algorithm builds groups as follows:

Initialization phase In the initialization phase of the algorithm, an empty set is created, which contains the sub graphs of the learning groups. Also G is defined as the social network of the classroom and MIN as the minimum number of students in a group.

Clique detection phase Based on G, all cliques are found with a minimum size of MIN. Each of these cliques is removed from G and added to the set of groups. Cliques in graph theory describes a set of vertices, which are all connected to each other. These cliques represent a good learning group, because everyone knows each other.

Relaxation phase After all cliques are removed, the relaxation phase tries to find k-edge-connected components in G, where k is stepwise reduced from $|G| - 1$ to 0. The phase also stops, if G is empty or no more components are found. If a k-edge-component is found, it is removed from G and added to the set of learning groups. K-connected-components are weaker connected than cliques.

Coalescence phase The coalesce phase checks, if G contains more vertices as MIN. If this is the case, G is added to the set of groups. Are less vertices in G as MIN, a vertex is removed from a group with a higher number of members than MIN and added to the group of left vertices. This is done until $|G| > MIN$.

Consensus learning phase In this last phase of the algorithm, each vector of a group member is replaced by an averaged vector of all group members, to transform the course material based on the group preferences to grantee a shard ubiquitous learning experience.

2.1.2 Discussion

The presented approach does not facilitate the features required for our goal, because it is bound to a small size of a classroom. Moreover, the groups created by the algorithm have some weaknesses. Even tough the students provide a vector of their learning preferences, it is not considered in the group formation process. Only in the last phase, the vector of all group members is averaged to grantee a shared ubiquitous eLearning experience. Another weakness is that the groups have different qualities. The groups generated in the first phase are very highly connected. This feature decreases during the algorithm and the last group is created by the left vertices with possible no connection among each other.

2.2 Multi-Objective Team Composition

Relevant work in the field of group formation can also be found in the field of team composition in a professional context. Dorn et al [1] try to build expert teams based on the skill of each user and their relation within the social network.

2.2.1 Methodology

The presented approach has three major stages, which are visualized in figure 1.

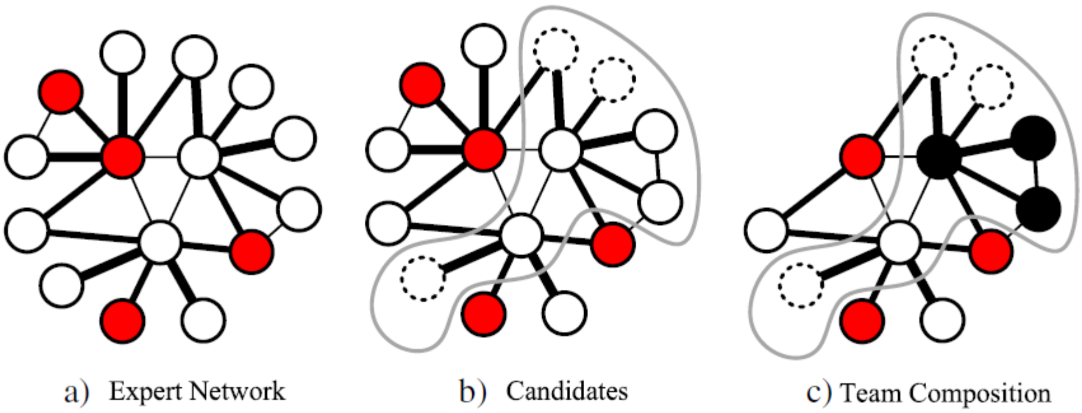

Figure 1: Overview of Group Formation approach by Dorn et al [1]

Network Establishment The first step is to establish a social network with user profiles covering information about the provided skills and the expert availability. The edges in the network are weighted based on former interaction of two experts. Figure 1a shows an example network with different edges weights and unavailable experts, who are represented by red marked vertices.

Candidate Selection Based on the network, candidates are selected based on their availability and the required skills. Figure 1b shows a set of selected candidates. These candidates are top ranked experts for the skills, but are just loosely connected in the network.

Heuristic Optimization To find a team, which is better connected than the top ranked experts, the team composition tries to identify a group with a high skill coverage and a high connectivity in the network. An important role in the approach takes the skill-dependent recommendation model. If a selected user is not available, the algorithm selects another based on the interaction structure of current team members customized for a given situation. To minimize manual management, a self-adjusting trade-off model is introduced which determines the trade-off between interaction distance and recommendations. The presented work includes a proof, that this step of the approach is NP-hard. Because of this challenge two heuristics are introduced: Genetic algorithms and simulated annealing. While Genetic Algorithms apply crossover and mutation operations on a given team configuration and evaluates the team fitness, Simulated Annealing maps each team configuration to a temperature, where a smaller temperature indicates a better team configuration. The evaluation of Dorn et al shows that both heuristics are able to find good team configurations, but Genetic Algorithm performs better and produce more stable configurations.

2.2.2 Discussion

The work of Dorn et al [1] presents a solid approach for team composition. The main evaluation focus on the comparison of simulated annealing and genetic algorithms and on the mathematical modelling. The presented approach is able to form groups on large networks with low computational cost because of the applied heuristics. It is mentioned that the algorithm should be able to search the network very fast, but there is a lack of possibilities to do so.

2.3 Searching for Expertise in Social Networks

The requirement of group formation for searching in large social networks with millions of users demands effective search algorithms to find suitable group members. Zhang and Ackermann [6] compare several search algorithms specialized for social networks. As test data, they generate a social network form an email dataset provided by the Enron company. In this network, a vertex is an employee with a vector of his/her knowledge. This knowledge is statistically determined based on words used in email communication. A directed edge between two vertices is created, when node a has sent an email to node b. To validate their test data, some metrics are evaluated like degree distribution. Table 1 shows the evaluated search strategies grouped in three families.

Family	Name	Heuristic
General computational	Breadth First Search (BFS)	Broadcasts query to all neighbours
	Random Walk (RWS)	Selects the next node randomly
Network structure based	Best Connected (BCS)	Selects next node based on the degree
	Weak Tie (WTS)	Based on Granovetter's weak tie concept [7]
	Strong Tie (STS)	Also based on weak tie concept
	Hamming Distance (HDS)	Picks the neighbour who has the most uncommon friendlist
	Cosine Similarity (CSS)	Like HDS, but divide the Hamming distance by the degree
Similarity based	Information Scent (ISS)	Picks the neighbour with the highest match score between the query and the profile

Table 1: Evaluated search strategies

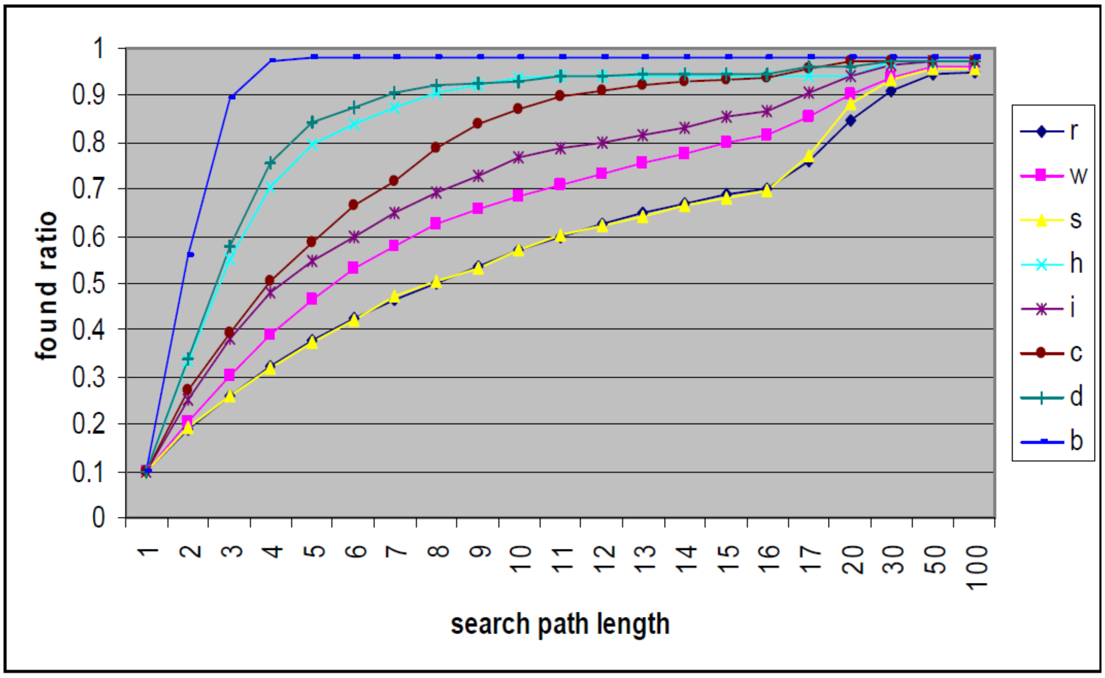

Figure 2: Search path length of the social network search strategies from Zhang and Ackermann [6]

2.3.1 Results

Figure 2 shows the percentage of successful queries as a function of different search lengths. While Breath First Search(BFSlength of) has the lowest search path length, although Hamming Distance Search(HDS) and Best Connected Search(BCS) performs very well. The search strategies with the highest averaged search path are Random Walk Search(RWS) and Strong Tie Search(STS).

Zhang and Ackermann also analyse the relevance of the out-degree of a node. They found that the out-degree is important for all search algorithms. Specially HDS and BCS have a high number of nodes with a high out-degree, because their metrics select nodes with a high out-degree.

Besides the computational cost and the importance of the out-degree, Zhang and Ackermann evaluate the social cost of the algorithms. One introduced metric to measure the social cost is the number of people used per query. While BFS, RWS and STS use eight to nine people per query in median, BCS, HDS and CSS only need three to four people to find an expert. WTS and ISS median lies between six and five.

Based on the result that STS performs worse than WTS. Zhang and Ackermann focus on the role of weak ties by searching on social networks. They remove the weak ties and run a new evaluation. The removal of weak ties lead to 23% of queries which could not be finished because of the less connected network.

A second observation shows the sensitivity of the search strategies to the out-degree. The 10 users, who have the highest out-degree were removed and a new evaluation was run. A high difference was found in the average path length by BCS, HDS and CCS, which is caused by their metric for choosing a next node. All other search strategies show no significant difference.

2.3.2 Discussion

The work of Zhang and Ackermann is very relevant for the purpose of group formation, as we need for an efficient search in the social network. But the results are not directly transferable. In their approach, the authors try to find one expert within the network, but in our approach we are seeking a group, which is built on an impact of their relation in the social network. So a search algorithm, which tends to has a deep-first search nature, would find good candidates with a high distance to the initiating node.

Another interesting contribution of this work is the method of creating of a test network. Zhang and Ackermann generate their test data from an E-Mail data set and extract the required information. This approach enables them to generate suitable real test data without implementing an own network or crawling a commercial one.

3 Approach

This section will give a short introduction of our approach of group formation in an eLearning-enabled Online Social Network. An in-depth presentation can be found in [8].

The goal of our group formation approach is to find a group of collaborators. These collaborators are modelled by their availability, learning style and knowledge. If a user wants to start a collaborative task, he / she selects a topic and the group formation is started. It tries to find suitable collaborators and ranks possible team configurations.

The availability of a user is represented by a simple flag, which can be true or false. To represent the learning style of a user, we use the Felder and Silverman Theory [9]. It assigns learning preferences according to four defined dimensions:

- Active or Reflective (Processing)
- Visual or Verbal (Input)
- Sensing or Intuitive (Perception)
- Sequential or Global (Understanding)

In each of these dimension the user can have three different strengths of the score: fairly well-balanced, moderate preference and a very strong preference.

We model the knowledge of a user as tags. If a user starts a group formation process, he /she selects some tags to define the required knowledge and assigns each tag a weight. Each user has although a set of tags with a corresponding activity index, which represents the relevance in the given fields. Now a user can be ranked by his /her knowledge tag set and the weighted required knowledge by calculating the inner product.

In agreement with Dorn et al [1], we split our approach into two parts:

- candidate selection and
- group formation

In the candidate selection phase, we try to find likely group candidates by searching the social network, starting at a user, who initiate the group formation. From this node, the network is traversed and at each user node, the distance between the initiator in learning style and the knowledge rank is evaluated. If the values are below a given threshold, the node is added to the candidate set.

If a given number of candidates is found, the group formation part tries to rank possible group constellations. We rank the group based on the distance in learning style, which should be minimized, the knowledge rank, which should be maximized and the density of the group in the social network, which also should be very low. To preform this task in time, we employ Genetic Algorithms. Here a team configuration is a chromosome and a set of team configurations represents a population. Given a fitness function which measures the quality of a configuration, in each generation some team configuration are changed using a crossover or mutation operation.

Performing a crossover operation on two team configurations exchanges the head and tail of these configurations on a random position. A mutation exchanges only one candidate from a configuration with another given in the candidate set.

Based on the list of ranked possible group constellations, the user can choose one and invite the other collaborators.

4 Conclusion and Outlook

This work discusses three related approaches in the field of group formation in an eLearning-enabled social network. The first paper introduces a group formation approach based on a social network between the members in a classroom. Because of the limited size of a classroom an algorithm with high computational cost is applied. The second work shows a group formation approach in a professional team composition field. The group is formed based on former interaction and the skill of a user. Main contribution of this paper is the application of genetic algorithms and simulated annealing to find suitable team configuration. In the third discussed paper several search algorithms are applied on a social network to find experts. Based on this related work, an approach was developed, which is able to form groups of learners in a social network based on learning style, knowledge and the position in the social network. In future research, the approach will be implemented and evaluated on synthetic and - if possible - on real data.

References

[1] C. Dorn, F. Skopik, D. Schall, and S. Dustdar, "Interaction Mining and Skill-dependent Recommendations for Multi-objective Team Composition," *Data & Knowledge Engineering*, vol. 70, pp. 866–891, 2011.

[2] J. Vassileva, "Toward Social Learning Environments," *Learning Technologies, IEEE Transactions on*, vol. 1, no. 4, pp. 199 –214, oct.-dec. 2008.

[3] A. Ounnas, H. C. Davis, and D. E. Millard, "A Framework for Semantic Group Formation in Education," *Educational Technology & Society*, vol. 12, no. 4, pp. 43–55, 2009.

[4] A. Ounnas, H. Davis, and D. Millard, "Towards Semantic Group Formation," in *Advanced Learning Technologies, ICALT 2007. Seventh IEEE International Conference on*, july 2007, pp. 825 –827.

[5] T. Arndt and A. Guercio, "Social Network-Based Course Material Transformations For A Personalized And Shared Ubiquitous E-Learning Experience," in *The 5th int. conf. on Mobile Ubiquitous Computing, Systems, Services and Technologies (UBICOMM 2011)*, 2011, pp. 218–222.

[6] J. Zhang and M. S. Ackerman, "Searching for expertise in social networks: a simulation of potential strategies," in *Pro. of the 2005 int. ACM SIGGROUP conf. on Supporting group work*, ser. GROUP '05, 2005, pp. 71–80.

[7] M. Granovetter, "The Strength of Weak Ties," *The American Journal of Sociology*, vol. 78, no. 6, pp. 1360–1380, 1973.

[8] S. Brauer and T. C. Schmidt, "Group Formation in eLearning-enabled Online Social Networks," in *Proceedings of the International Conference Interactive Computer aided Learning ICL 2012*, September 2012.

[9] R. Felder and L. Silverman, "Learning and teaching styles in engineering education," *Engineering education*, vol. 78, no. 7, pp. 674–681, 1988.