A Case-Based Reasoning Approach for Facilitating Online Discussions

Wen Gu[1,2(✉)], Ahmed Moustafa[1], Takayuki Ito[1], Minjie Zhang[2], and Chunsheng Yang[3]

[1] Department of International Collaborative Informatics, Nagoya Institute of Technology, Nagoya, Japan
{gu.wen,ahmed.moustafa}@itolab.nitech.ac.jp, ito.takayuki@nitech.ac.jp
[2] School of Computing and Information Technology, University of Wollongong, Wollongong, Australia
minjie@uow.edu.au
[3] Digital Technology Research Center, National Research Council Canada, Ottawa, Canada
chunsheng.yang@nrc.gc.ca

Abstract. In online discussion platforms, human facilitators are introduced in order to facilitate the discussions to proceed smoothly and build consensus efficiently. However, problems such as human bias and scalability are becoming critical with increasing sophistication of these online discussion platforms. In order to address these problems, online discussion facilitation support becomes more and more essential. Towards this end, in this paper, a novel case-based reasoning (CBR) based online discussion facilitation support approach, which consists of a case definition method and a case retrieval algorithm, is proposed to support online facilitation in large-scale discussion environments. The proposed approach models the online discussions using the issue based information system (IBIS) discussion style, where complex problems are modelled as a conversation amongst several stockholders. In the proposed approach, discussion cases are generated and retrieved based upon the structure features of their discussions. The experimental results show the proposed discussion case generation approach is able to reflect more precise discussion features than those approaches that are based only on the quantitative features, and the ability of the proposed case retrieval algorithm to retrieve the most similar case from the case base.

Keywords: Online discussion platforms · Facilitation support · Case-based reasoning · Issue based information system

1 Introduction

One effective approach to solve critical social problems is to collect the wisdom from a crowd of participant people. However, it grows difficult to organize a large number of people to discuss in one particular place during a particular time. With the development of the Internet, online discussion forums have attracted much attention as the platform of gathering a crowd of people together to solve

A. C. Nayak and A. Sharma (Eds.): PRICAI 2019, LNAI 11672, pp. 582–592, 2019.
https://doi.org/10.1007/978-3-030-29894-4_47

common problems. The core advantage of online discussion is that people can join the discussion via the Internet from different places whenever they are free. Therefore, platforms such as Climate CoLab [1], Deliberatorium [2] and Collagree [3] have been developed to encourage people to discuss in online environments. These platforms have been utilized to organize people to participant in online discussions about topics such as global climate change, law reform and city planning. In order to facilitate the discussion to proceed smoothly and achieve consensus efficiently, many platforms [2,3] introduce human facilitators into the discussions to conduct facilitation. Facilitators can promote the development of the discussion, integrate ideas and opinions, and help the group to build consensus [4]. However, as there is no general definition of online discussion facilitation, a human bias cannot be avoided. In addition, with the increase of the participants number in a certain discussion, scale issues and schedule issues also become critical problems for human facilitators.

As a result, it becomes more and more necessary to develop online discussion facilitation support techniques to help human facilitators relieve their burdens. The challenging part of facilitation in online discussion forums is that it is a very complicated problem, which changes significantly with the discussion development. It means that plenty of information need to be considered for facilitation and it is difficult to describe the method in a number of specific rules. Existing research [5] emphasizes the importance of experiences for human facilitators, since they reuse the successful experience they had in the past to solve new similar problems.

In this paper, we propose using CBR to support online discussion facilitation. CBR is one of the famous artificial intelligence techniques that have been successfully used in real-world applications [6,7]. In this regard, it provides an effective reasoning paradigm for solving new problems by adopting similar solutions that have been proposed for similar problems in the past [8]. This is very similar to the human facilitator thinking paradigm. Experienced human facilitators facilitate better than novice human facilitators because they have more experience that is derived from the past facilitation they have done. And experienced human facilitators are able to utilize these sorts of experience when they try to conduct facilitation in new discussion situations. Towards this end, in this paper, we propose a CBR based approach to support online discussion facilitation. When using CBR to solve a problem, the first and most important part is to find out the essential characteristics which can be used to express this problem in order to define the problem as a case. As a result, this proposed approach introduces a novel method of defining online discussion cases from their structures. In addition, a case retrieval algorithm is implemented to retrieve the most similar cases from the case base.

The rest of this paper is organized as follows. Section 2 introduces the related work of this research. In Sect. 3, the proposed CBR based online discussion facilitation support approach is introduced. Section 4 presents the experimental settings along with the experimental results.

2 Related Work

Many research efforts have been attempted in order to support group discussion facilitation. In this regard, Dickson et al. [9] explored three human-based facilitation modes in Group Decision Support System(GDSS) meeting and showed that group consensus improved in all the three modes. In addition, Anson et al. [10] showed human facilitators maintain high quality group interactions and improved group processes and greater cohesion. Also, an automated facilitation technique has been developed to support group discussion. Limayem et al. [11] showed both human-facilitated decision making groups and automated-facilitated decision making groups experienced significantly higher post-meeting consensus and perceived decision quality than non-facilitated groups in GDSS meeting. Aiken and Vanjani [12] showed that automated facilitator is better than human facilitator for simple idea generation and voting tasks. Wong and Aiken [13] showed both expert-human facilitated groups and automated facilitated groups perform significantly better than novice-human groups in electronic meetings when faced with relatively simple idea generating and ranking tasks. Derrick et al. [14] demonstrated that automated facilitation of system requirement generation is possible and showed that the agent-facilitated groups generate more complete requirements than non-facilitated groups. As shown in the above mentioned research works, just like human facilitators, automated facilitation techniques can also support group discussions.

However, there are still some problems in group-discussion facilitation. In specific, most of the group-discussion facilitation techniques can only support tasks such as agenda preparer, timekeeper, simple idea generation and voting. Therefore, it becomes difficult to use these techniques to support high level online discussion facilitation, since high level discussion facilitation, such as proper facilitation time detection and facilitation pattern decision, needs to be generated on the basis of the dynamically changing discussion situation. As a result, it is highly critical to develop novel discussion facilitation techniques to support high level facilitation in online discussion forums.

On the other hand, Gu et al. [15] proposed a CBR based online discussion facilitation support approach that is able to adapt to different discussion situations. Specifically, they proposed to use IBIS style to model the discussion structure and to define discussion cases from the quantitative perspective of the IBIS structure elements. IBIS style is based on the principle that the design process for complex problems is fundamentally a conversation amongst several stakeholders [16]. This has been used as a visual aid to help participants and facilitators to understand the discussion structure. However, in this IBIS based discussion process, a lot of inner characteristics cannot be reflected if we just consider the quantitative features. For example, we consider the similarity between $Case(a)$ and $Case(b)$ that are demonstrated in Fig. 1. In this situation, if we use the quantitative features to define a case base, the differences between $Case(a)$ and $Case(b)$ cannot be distinguished because their quantitative features are the same. Even though, they are constituted by different structures, these differences cannot be reflected because they have the same number of issues, ideas, and arguments. On the other hand, it is critical to consider the structure characteristics in the case definition because the discussion structure reflects the relationship among the discussion contents. For example, the connections among different vertexes need to be considered because they reflect the consistency in the discussion while those vertexes which are not connected cannot reflect the consistency in the discussion. As a result, it is critical to design new approaches to define the discussion case in order to reflect the structure characteristics of the discussion.

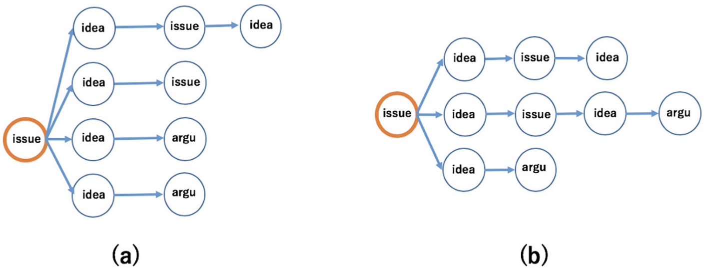

Fig. 1. Two IBIS style discussion cases

3 Proposed Approach

3.1 Problem Description

In this research, we aim to develop CBR based online discussion facilitation support approach that can be used to help human facilitators conduct facilitation in online discussion platforms.

In order to design a CBR based online discussion facilitation support approach, as the first step, we need to build a reasonable case base that represents the system's experience in online discussion facilitation. For each case, the definition consists of a problem description part and a problem solution part. In this paper, we propose a novel structure perspective method to describe the online discussion facilitation problem. The solution of the problem is a result that whether facilitation is necessary to be added or not. In addition, we propose a case retrieval algorithm in order to find the most similar case from the case base.

Specifically, we consider the discussion process that is represented in issue based information structure(IBIS) style [16]. We generate IBIS structures of the discussions by using three sorts of elements which are issue, idea, and argument. Issues are defined as the questions that need to be answered during the discussions. Ideas are defined as the possible answers to the issues. Arguments contains both pros and cons. Pros are defined as the support to an issue or an idea and cons are defined as the object to an issue or an idea. Each element in the IBIS style discussion structure is represented as a vertex. Facilitations are the posts that facilitators generated to promote the discussion. Each association between two elements is represented as a directed edge. For example, discussion in Fig. 2, (a) can be represented in IBIS style as demonstrated in Fig. 2, (b). FA represents facilitation posts and argu represents argument posts.

Fig. 2. IBIS style discussion expression

3.2 Discussion Case Definition

Discussion case definition has been proposed from the quantitative perspective which focuses on the numeric IBIS style discussion features such as issue number, idea number and argument number [15]. However, in order to reflect more precise features of the generated IBIS style discussion graphs, the characteristics of the structure in the graph also need to be considered. One of the examples is demonstrated in Fig. 3.

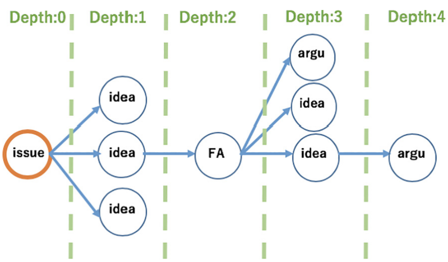

Fig. 3. IBIS style discussion expression

As shown in Fig. 3, one of the salient characteristics in the IBIS style discussion structure is the original vertex. The original issue vertex, which equals to the root topic of the discussion, can be considered as the parent vertex of all other vertexes. All other vertexes show the discussion details which are generated to solve the original discussion issue. Considering the discussion structure, it is obvious that the depth of the vertexes in the discussion structure can reflect status information in the discussion. For example, if two issues hold the same number of ideas, the issue which has more deep ideas can be considered as better discussed than the issue which has fewer deep ideas.

In this research, we define the original vertex's depth as 0. All other vertexes are in a depth which is more than 0. And we define the discussion case as a sort of labeled graph on the basis of the IBIS style discussion structure. Specifically, we define a labeled discussion graph as a directed graph in which each vertex and each edge is associated with one label. A labeled graph is defined by a triple G $= <V, r_V, r_E>$, where.

- V is a finite set of vertexes.
- $r_V \subseteq V \times L_V$ shows the relation between vertexes and labels. Each vertex has only one related label, which can be issue, idea or argument. The situation that v_i has label l_{v_i} is represented by a tuple (v_i, l_{v_i}). r_V is the set of tuples v_i.
- $r_E \subseteq V \times L_V \times V' \times L_{V'}$ shows the relation between edges and labels. One edge $(v_i, l_{v_i}, v_j, l_{v_j})$ is defined as the combination of two labeled vertexes (v_i, l_{v_i}) and (v_j, l_{v_j}). r_E is the set of quaternaries $(v_i, l_{v_i}, v_j, l_{v_j})$.

For example, one labeled IBIS style discussion case graph as demonstrated in Fig. 4(a) can be formulated by a labeled group as demonstrated in Fig. 4(b).

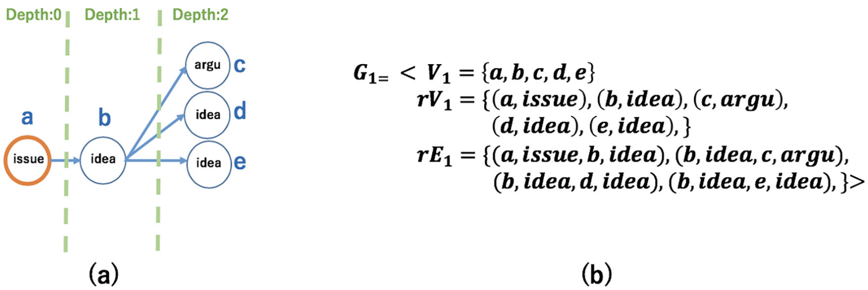

Fig. 4. IBIS style discussion expression

3.3 Discussion Case Retrieval

In this research, we calculate the similarity between two discussion cases on the basis of labeled graph similarity algorithm [18].

When we measure the similarity between two labeled IBIS style discussion graphs, we use the idea that comparing the number of features which are common to both objects, to the total number of the features [19]. In our situation, the total number of the features of a labeled IBIS style discussion graph can be represented as a triple G. When we compare the similarity of two graphs $G_1 = <V_1, r_{V_1}, r_{E_1}>$ and $G_2 = <V_2, r_{V_2}, r_{E_2}>$, such that $V_1 \cap V_2 = \emptyset$, one additional thing that we need to consider is the relation that connects the two graphs. In this research, we define the relation that connects two graphs as: if two vertexes hold same sort of label and are in the same structure depth, these two vertexes are consider to be similar. Similarity between two graphs G_1 and G_2 can be calculated by Eq. 1.

$$sim_S(G_1, G_2) = \frac{f(descr(G_1) \cap descr(G_2))}{f(descr(G_1) \cup descr(G_2))} \tag{1}$$

$descr(G_1)$ and $descr(G_2)$ are the descriptions of labeled graph G_1 and G_2, respectively. Each description is made up by all the labeled graph vertex features in addition to the edge features. The similarity between the two graphs, i.e., G_1 and G_2, is calculated by using the common features that the two graphs share divided by the set of all the two graph features.

There is a special state where new discussion case is the subgraph of more than two cases in the case base, and the number of common features of these two cases is the same. If these two cases are retrieved from the case base together, and then the algorithm retains the subbranch that includes the new discussion case and cut off other subbranches from the origin vertex. After that, the similarity is recalculated between these two cases in order to choose the most similar case.

4 Experiments

In this section, we introduce the experimental results of comparing the proposed structural perspective discussion case definition with the quantitative discussion case definition in case retrieval results.

In order to demonstrate the ability of the proposed CBR-based approach to retrieve similar discussion cases efficiently, we built a synthetic test case base that includes seven discussion cases. The quantitative information of the case base is shown in Table 1.

Table 1. Synthetic cases in test case base

Case ID	Case1	Case2	Case3	Case4	Case5	Case6	Case7
Number of issues	1	1	1	1	1	3	3
Number of ideas	1	3	3	1	5	5	5
Number of arguments	2	4	3	1	5	2	2
Idea depth	1	1	2	1	3	3	3
Facilitation	1	0	0	1	0	1	0

We choose four sorts of quantitative information as the parameters, number of issues, number of ideas, number of arguments and idea depth. The idea depth means the depth of the deepest idea in the discussion case. And the structural information of the case base is demonstrated in Fig. 5.

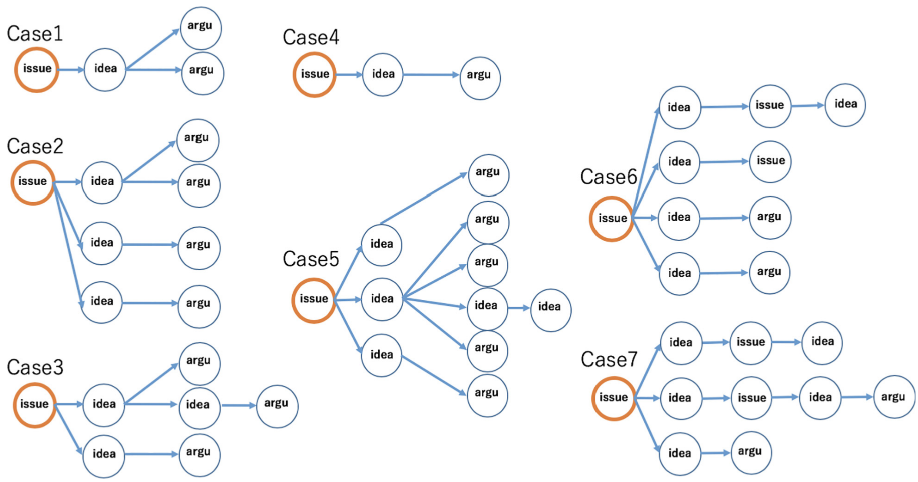

Fig. 5. Experiment case base

In addition, we built two synthetic test cases whose information is demonstrated in Table 2 and Fig. 6.

Table 2. Synthetic test cases

Case ID	Case_t1	Case_t2
Number of issues	1	1
Number of ideas	1	6
Number of arguments	2	6
Idea Depth	1	2

In this experiment, firstly, we retrieve the most similar case to our test cases from the case base by using the Nearest Neighbor (NN) algorithm [17]. In NN algorithm, similarity is calculated on the basis of the euclidean distance of each feature parameter. Two cases are more similar if the euclidean distance between them is smaller. If the euclidean distance between two cases is 0, it means that the two cases are identical, i.e., the two cases hold the same number of quantitative perspective features. Secondly, we retrieve the most similar case to our test case from the case base by using the proposed labeled graph similarity algorithm we

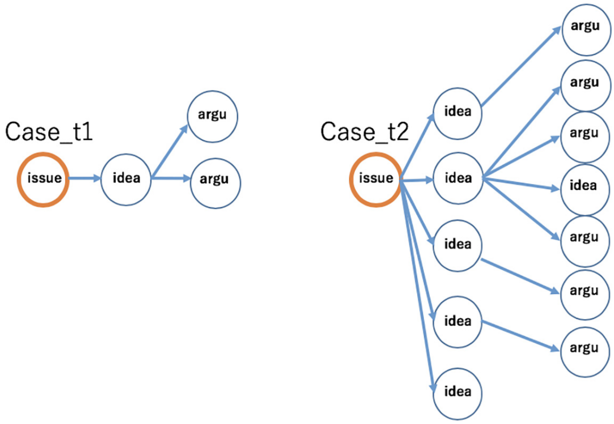

Fig. 6. Synthetic test cases

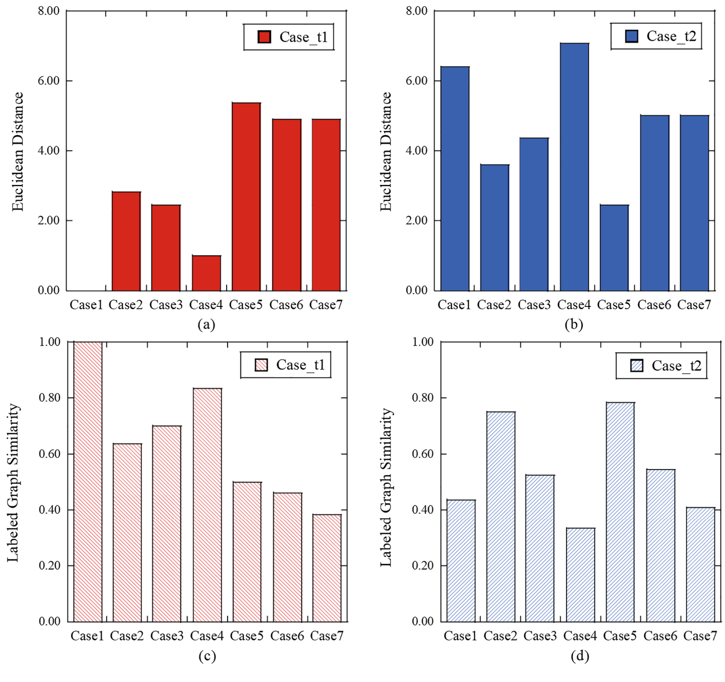

Fig. 7. Experiment results

introduced in Sect. 3. Two cases are more similar if the labeled graph similarity between them is higher.

Specifically, two test cases, *Case_t*1 and *Case_t*2 are designed to test our proposed approach. *Case_t*1, which is identical to *Case*1 of the test case base, is designed to test whether identical cases can be calculated as identical by our proposed approach. $Case\_t2$, which is similar to $Case5$, $Case6$ and $Case7$ of the test case base, is designed to test whether structure details can be reflected by our proposed approach.

The results of the experiments are shown in Fig. 7. From these results, we can see that proposed approach can work as good as the quantitative perspective discussion case definition method. In specific, it finds the most similar case to $Case\_t1$ and $Case\_t2$ from the case base, which are $Case1$ and $Case5$, respectively. $Case\_t1$ is also identified to be identical with $Case1$ because the labeled graph similarity is 1. In addition, for the similarity results of $Case6$ and $Case7$, differences are not reflected when using the quantitative features, as shown in Fig. 7(a) and (b). However, this sort of difference can be reflected when using the structure perspective features and the proposed algorithm, as shown in Fig. 7(c) and (d). This reflects the situation that the quantitative features are the same, but the structures are different.

5 Conclusion

In this paper, we proposed a novel CBR based online discussion facilitation support approach to support online discussion facilitation. In the proposed approach, IBIS style discussion structure format is employed to define discussion cases on the basis of their discussion structure. Labeled graph similarity algorithm is utilized to retrieve the most similar discussion cases from the case base. Experimental results demonstrated the ability of the proposed case definition method to reflect structural discussion cases differences that the quantitative features cannot reflect and the ability of the proposed case retrieval algorithm to retrieve most similar cases from the case base. One of the directions for future work is to improve the efficiency of the proposed case retrieval algorithm when handling large-scale case base. The possible solution is to consider comparing the extracted features from each case instead of comparing two graphs. Another direction of future work is to introduce more information into the case definition in order to reflect more precise discussion situation. One possible solution is to consider the semantic information of each IBIS vertex. These additional information can be added as sub-vertexes that can ensure more precise case retrieval.

Acknowledgment. This work was supported by JST CREST Grant Number JPMJCR15E1, Japan.

References

1. Introne, J., Laubachar, R., Olson, G., Malone, T.: The climate Colab: large scale model-based collaborative planning. In: Proceedings of the International Conference on Collaboration Technologies and Systems (CTS 2011) (2011)
2. Klein, M.: How to harvest collective wisdom on complex problems: an introduction to the MIT deliberatorium. CCI working paper (2011)

3. Ito, T., Imi, Y., Ito, T.K., Hideshima, E.: COLLAGREE: a faciliator-mediated large-scale consensus support system. In: Collective Intelligence 2014, 10–12 June 2014. MIT, Cambridge (poster) (2014)
4. Ito, T.: Towards agent-based large-scale decision support system: the effect of facilitator. In: The 51st Hawaii International Conference on System Sciences, Hilton Waikoloa Village, USA, 3–6 January 2018 (2018)
5. McLean, M.: What can we learn from facilitator and student perceptions of facilitation skills and roles in the first year of a problem based learning curriculum. BMC Med. Educ. **3**, 1–10 (2003)
6. Yang, C., Orchard, R., Farley, B., Zaluski, M.: Authoring cases from free-text maintenance data. In: Perner, P., Rosenfeld, A. (eds.) MLDM 2003. LNCS, vol. 2734, pp. 131–140. Springer, Heidelberg (2003). https://doi.org/10.1007/3-540-45065-3_12
7. Lopes, E.C., Schiel, U.: Integrating context into a criminal case-based reasoning model. In: The Proceedings of 2nd International Conference on Information, Process, and Knowledge management (2010)
8. Schank, R.C.: Dynamic Memory: A Theory of Reminding and Learning in Computers and People. Cambridge University Press, New York (1983)
9. Dickson, G.W., Partridge, J.-E.L., Robinson, L.H.: Exploring modes of facilitative support for GDSS technology. MIS Q. **17**(2), 173–194 (1993)
10. Anson, R., Bostrom, R.P., Wynne, B.: An experiment assessing group support system and facilitator effects on meeting outcomes. Manage. Sci. **41**, 189–208 (1995)
11. Limayem, M., Lee-Partridge, J.E., Dickson, G.W., DeSanctis, G.: Enhancing GDSS effectiveness: automated versus human facilitation. In: Proceedings of the Twenty-Sixth Hawaii International Conference on System Sciences, vol. 4, pp. 95–101, January 1993
12. Aiken, M., Vanjani, M.: An automated GDSS facilitator. In: Proceedings of the 29th Annual Conference of the Southwest Decision Sciences Institute, Dallas, TX, 3–7 March 1998, pp. 87–89 (1998)
13. Wong, Z., Aiken, M.: Automated facilitation of electronic meetings. Inf. Manage. **41**(2), 125–134 (2003)
14. Derrick, D.C., Read, A., Nguyen, C., Callens, A., de Vreede, G.: Automated group facilitation for gathering wide audience end-user requirements. In: 2013 46th Hawaii International Conference on System Sciences, pp. 195–204, January 2013
15. Gu, W., Moustafa, A., Ito, T., Zhang, M., Yang, C.: A case-based reasoning approach for automated facilitation in online discussion systems. In: The Proceedings of The 2018 International Conference on Knowledge, Information and Creativity Support Systems (KICSS 2018), Thailand, November 2018
16. Kunz, W., Rittel, H.W.J.: Issues as elements of information systems. Center for Planning and Development Research, Institute of Urban and Regional Development, Working Paper No. 131, University of California, Berkeley (1970)
17. Dasarathy, B.V. (ed.): Nearest Neighbor Pattern Classification Techniques. IEEE Computer Society Press, Los Alamitos (1991)
18. Champin, P.-A., Solnon, C.: Measuring the similarity of labeled graphs. In: Ashley, K.D., Bridge, D.G. (eds.) ICCBR 2003. LNCS (LNAI), vol. 2689, pp. 80–95. Springer, Heidelberg (2003). https://doi.org/10.1007/3-540-45006-8_9
19. Lin, D.: An information-theoretic definition of similarity. In: Proceedings of ICML 1998, Fifteenth International Conference on Machine Learning, pp. 296–304. Morgan Kaufmann (1998)