IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 13, No. 2, June 2024, pp. 1794~1804
ISSN: 2252-8938, DOI: 10.11591/ijai.v13.i2.pp1794-1804

Strategies for improving the quality of community detection based on modularity optimization

Tedy Setiadi[1,2], Mohd Ridzwan Yaakub[2], Azuraliza Abu Bakar[2]
[1]Department of Informatics, Faculty of Industrial Technology, Universitas Ahmad Dahlan, Yogyakarta, Indonesia
[2]Center for Artificial Intelligence Technology, Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, Bangi, Malaysia

Article Info

Article history:

Received Jan 6, 2022
Revised Dec19, 2023
Accepted Jan 16, 2024

Keywords:

Community detection
Modularity
Optimization
Quality
Resolution limit
Strategies

ABSTRACT

Community detection is a field of interest in social networks. Many new methods have emerged for community detection solution, however the modularity optimization method is the most prominent. Community detection based on modularity optimization (CDMO) has fundamental problems in the form of solution degeneration and resolution limits. From the two problems, the resolution limit is more concerned because it affects the resulting community's quality. During the last decade, many studies have attempted to address the problems, but so far they have been carried out partially, no one has thoroughly discussed efforts to improve the quality of CDMO. In this paper, we aim to investigate works in handling resolution limit and improving the quality of CDMO, along with their strengths and limitations. We derive seven categories of strategies to improve the quality of CDMO, namely developing multi-resolution modularity, creating local modularity, creating modularity density, creating new metrics as an alternative to modularity, improving the louvain algorithm, involving node attributes in determining community detection, and extending the single objective function into a multi-objective function. By considering network size, network type, and community distribution, we can choose the appropriate strategy in improving the quality of community detection.

Corresponding Author:

Tedy Setiadi
Department of Informatics, Faculty of Industrial Technology, Universitas Ahmad Dahlan
Yogyakarta, Indonesia
Email: tedy.setiadi@tif.uad.ac.id

1. INTRODUCTION

Recently, the realm of social networking has developed and attracted the attention of various fields of research [1]. Community detection becomes the initial task and main task of social network analysis because of the vital role of community in social network analysis. Research to develop community detection methods and algorithms is growing rapidly, in line with the needs of applications that are increasingly broad and complex in the real world. In addition to social networks, several examples of the application of community detection are in various fields, including criminal [2], public health [3], politics [4], library [5], and prediction [6], [7]. Many interdisciplinary researchers have attempted to solve this problem. However, there is no adequate solution [8]–[10].

Several researchers proposed a classification of community detection methods. Fortunato and Hric [11] classified the community detection method into four methods, namely spectral methods, statistical inference-based methods, optimization-based methods, and dynamic-based methods. George *et al.* [10] added Fartunato classification into nine methods namely Bayesian and regularized likelihood, diffusion, spin-dinamics, synchronization, greedy, and devide-conquer. Mittal and Bhatia [12] classified into four methods, namely modularity algorithm, information theoritic agorithm, network algorithm, and hierarchical algorithm. Dao *et al.* [13] classified into five method, namely edge removal, modularity optimization, spectral methods, and statistical inference. In the literature, from these various classifications, although they are classified as traditional methods, until now the modularity optimization method is the most popular [8], [11], [14]–[16].

Althought Community detection based on modularity optimization (CDMO) is prominent, it has several drawbacks. The first drawback is the fundamental problem in CDMO, namely the tendency to choose small communities over large networks, while others prefer large communities over small ones [17]. On the CDMO problem, has two cases. The first case is called solution degeneration, where a large number of community structures may be found from a network whose topological structures are very different from one another but produce a value of modularity that is very close to the optimal [18], [19]. The second case is called the resolution limit, which is a limitation in finding communities that are smaller than a certain scale [20].

Moreover, the problem of community detection arises along with the community's own ill-definition. Although there is already a qualitative definition that is a community is a node group that has a closer relationship than with other groups, but the definition of group boundaries can be called community does not yet exist. Therefore, quantitatively it is still an open debate and issue so that it allows many developing methods to measure community detection quality [1], [2].

The main issue of CDMO is finding a network partition that maximizes modularity. Modularity optimization is non-polinomial (NP)-hard problem [21] so it is difficult to directly find the optimal solution. Various methods have been developed to overcome the weaknesses of CDMO, but so far the discussion is still partial, it has not been discussed more thoroughly. Therefore, this paper focuses on discussing various method for improving the quality of CDMO in relation to the modularity and objective functions used.

The rest of this paper is organized as follows. Section 2 will describe the related work on metrics quality of community detection and community detection based on modularity, the detail of strategy to improve the quality of CDMO are described in section 3, section 4 presents analysis of strategy to improve the quality of CDMO. Finally, the conclusion is given in section 5.

2. RELATED WORK

2.1. Metrics quality of community detection

The ill-definition about community as mentioned above, creates complexity in analyzing the quality of the resulting community. The complexity of analyzing community quality is marked by the multitude of metrics for measuring the proposed community's quality. Communty analysis according to Chakraborty *et al.* [19] consists of two sequential phases: first is community detection, i.e., the process of finding a network community structure using a specific community detection algorithm, and second is community evaluation, i.e, the process of evaluating the feasibility of the structure of community detection results.

Evaluation metrics are used as an indication of community quality, then the quality of the community detection algorithm is estimated based on the output value of the metric. In addition, based on data availability, the analysis of community quality is classified into two categories, first, analysis of a community with ground truth. The quality of the community is compared to its ground truth; the second is an analysis that does not have ground truth [22]. Several of the community detection metrics are:

- Modularity [23] for measuring the strength of the community structure, where a good community has a larger number of internal edges and a smaller number of inter-community edges than expected when compared to a random graph
- Conductance [24] for measure the ratio to the total amount community edges c that are connected to other external communities by the total number of edges connected to the community c
- Separability [25] for measures the ratio of the number of internal edges to the total number of external edges in a community
- Density [25] for measures the density of the internal edges of the community, the ratio of the number of internal edges to the total maximum possible sides on the network.
- Transitivity/clustering coefficient [26] for measure the tendency of a group of nodes to form a community
- Surprise [27] for calculates the probability (minus the logarithms) of observing the side in the community against the possible population
- Significance [19] for comparing each community density with the average density
- Permanence [19] for measures the probability that a vertex remains in the community to which it is assigned and the extent to which it is "drawn" by neighboring communities

Meanwhile, several community evaluation metrics, including:

- Purity [28] for measure the number of matches between the detected community and its ground-truth
- F-measure [29] for measure the ratio to the total amount community edges c that are connected to other external communities by the total number of edges connected to the community c

- Adjusted rand index [30] for measures the ratio of the number of internal edges to the total number of external edges in a community
- Normalized mutual information (NMI) [31] for measures the probability that a vertex remains in the community to which it is assigned and the extent to which it is "drawn" by neighboring communities.

Although many metrics have been proposed, none is the best and universally accepted metric [32]–[34]. Community detection can be formulated as an optimization problem, where the objective function assigned is to maximize the number of links in each network partition [33]. There are several evaluation metrics that can measure the quality of community outcomes as well as be used as an objective function, such as modularity, NMI, and purity. However, modularity is the most widely used metric and objective function of optimization problems [13], [19], [35].

2.2. Community detection based on modularity optimization

A community is a group of nodes connected more strongly to each other than the rest of the network. A community consists of nodes that share something, such as affiliations (friends, clubs, and colleagues), shared interests, and shared content (books, movies, web pages, and products). In the real world, there are various types of communities, including non-overlapping communtiy (e.g., each worker is only in one department), overlapping community (e.g., each student can have several hobbies), hierarchical community (e.g., a network composed of body cells, which in turn form organs), and local community (e.g., someone's unequal friendships on Facebook) [36].

Newman and Girvan [37] introduced modularity as a metric of community structure strength found in undirected and unweight graph as well as an objective function of the proposed algorithm. The author used the iterative hierarchy of edges betwenes deletion method. The result is a dendogram representing the partitioned community. Then, the Newman and Girvan modularity metric is widely used as an objective function in various CDMO algorithms [38]–[45].

Given a simple graph G(V, E), where V is the set of vertices and E is the set of (undirected) edges. A community or cluster $C \subseteq V$ is a subset from vertices, and clustering C={C_1, C_2,..., C_n} from G is a partition V into the clusters such that every vertex is in exactly one cluster. Modularity is defined in (1) [21]:

$$Q_c = \sum_{c=1}^{n_c} \left[\frac{L_c}{\mathrm{L}} - \left(\frac{d_c}{2\mathrm{L}} \right)^2 \right] \tag{1}$$

with n_c number of communities, L_c is total number of edges in community c, d_c total number of nodes in community c, and L is total number of edges on graph. The higher the value of modularity, the stronger the resulting community structure. The value of modularity in the interval [-1,1]. As an illustration, the community structure and its modularity values are shown in Figure 1. The community structure with non-optimal modularity is seen in Figure 1(a)-(c), while the community structure with optimal modularity is seen in Figure 1(d).

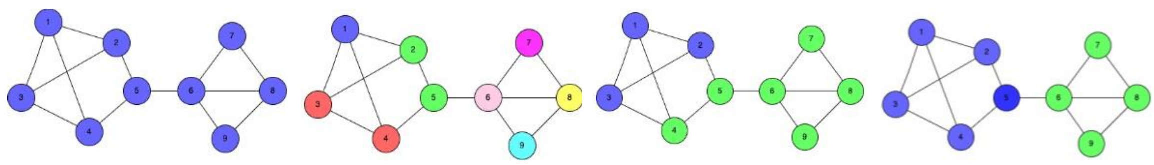

Figure 1. Structure communities by partition: (a) single community, (b) negative modularity, (c) sub optimal partition, and (d) optimal partition

For example, the process of calculating the total score of modularity in Figure 1(a) is:

$$Q = \sum_1^2 \left[\frac{k_c^{in}}{2M} - \left(\frac{k_c}{2M} \right)^2 \right] = Q_{1,2,3,4,5} + Q_{6,7,8}$$
$$= \left[\frac{14}{2.13} - \left(\frac{15}{2.13} \right)^2 \right] + \left[\frac{10}{2.13} - \left(\frac{11}{2.13} \right)^2 \right] = 0.44$$

Along with the variety of network types and communities the Newman and Girvan modularity was expanded to fulfill this. Modifications of modularity include modularity for weighted graphs, modularity for directed graphs, modularity for overlapping community, similarity-based modularity, motif modularity, max-min modularity, influenced-based modularity, diffusion-based modularity, dist-modularity [19]. All these variations of modularity have drawbacks regarding the resolution limit.

3. STRATEGY TO IMPROVE THE QUALITY OF COMMUNITY DETECTION BASED ON MODULARITY OPTIMIZATION

We investigate various methods developed to improve CDMO quality. We grouped them into six CDMO quality improvement strategies. We also describe the analysis of limitations and the appropriate types of data sets.

3.1. Developing multi-resolution modularity

As mentioned earlier, Newman-Girvan modularity has the disadvantage of resolution limits, namely that it can only find community structures at a certain characteristic scale, whereas on the other hand, many complex networks may have community structures at multi- scales [18], [46]. Therefore, various multi-resolution methods have been proposed, by directly adding parameter γ to the definition of modularity, resulting in multi-resolution modularity as in (2) [47]-[49]:

$$Q_g(\gamma) = \frac{1}{2M}\sum_c (k_c^{in} - \gamma \frac{k_c^2}{2M}) \tag{2}$$

Xiang *et al.* [50] developed a self-loop strategy in which the multi-resolution modularity equivalent of the parameter is derived indirectly from the Newman-Girvan modularity. One of the advantages of the self-loop strategy is that the resulting multi-resolution modularity can be optimized with existing modularity optimization algorithms, thus enriching the application of the modularity optimization algorithm in community detection. The drawback of this multiresolution method has an intrinsic limitation in that it increases the parameter value, the large community may have split when the small community becomes gradually visible in some cases where the community even becomes a fully connected sub-graph. In other words, increased resolution of modularity is obtained at the expense of community stability [48]. In addition, based on the data set used, this strategy is only suitable for small network [51], [52], and the distribution of community size is not wide [48].

3.2. Creating local modularity

As defined, modularity is the fraction of links in the community minus the expected value in the null (random network) model. The null model needs attention because it affects the value of modularity. Modularity using the null model is called global modularity because it is assumed to be the global connectivity of the community on the network. However, in many real-world networks the community is only connected to a small number of neighboring communities. This is known as the local connectivity community on the network. Based on this, some researchers propose local modularity as a modification of (global) modularity by changing the null model components with local components, with the hope of increasing the value of modularity so that the quality of detection of commutas increases. Several studies in making local modularits include.

Muff *et al.* [53] modified global modularity metrics to local ones for use in biological network detection. The modularity of each community i is calculated based on its subnet and the relationship with neighboring communities only. With all the links to its neighbors from the community i is L_{iN}, then the sum modularity of all k communities is local (L_Q) obtained:

$$L_Q = \sum_{i=1}^{k} \left[\frac{L_i}{L_{iN}} - \frac{(L_i)_{in}(L_i)_{out}}{(L_{iN})^2}\right] \tag{3}$$

In contrast to Q, the value of LQ is not limited to 1, but can be any value. The more local connected communities, the greater the LQ value. If all communities are connected to each other then the value of LQ will be the same as Q. The author claims local modularity detected more cohesive communities, and can complement each other with global modularity with higher detail. Xiang *et al.* [54] modify local modularity with a loop strategy. Local modularity is used as an objective function and self-consistency method for optimizing the local modularity.

A different approach is taken by Ronhovde and Nussinov [55], they was inspired by a physics approach to propose the potts spin glass model for community detection. The community is represented by the state of the potts spin glass model, while the partition quality is represented by the associated energy system. All of these local modularity proposals have the disadvantage of requiring higher computational costs because they have to perform parameter tuning to overcome the resolution limit [56]. In addition, based on the data set used, this strategy yields significant results for large scale network, and large community size distributions [57].

3.3. Creating modularity density

Apart from making local modularity, an effort to improve the modularity function is to create a density modularity metric. First introduced by Li *et al.* [58], he proposed the modularity density (Qd) as average modularity. The author mathematically proves that the modularity density does not divide a cliq into two parts, the maximum modularity is equivalent to the objective function of the k-means kernel. The author formulates the problem of community detection to a nonlinear integer programming model with the objective function of maximizing (Qd). Costa [59] performed modification of density modularity Li to overcome the possible optimal solutions, but there are communities that have negative density modularity. In addition, Costa reformulates the non-linear integer model from Li into mixed integer linear programming (MILP) [60].

Holmström *et al.* [61] taked a different approach using statistics and matrix algebra in finding density modularity. The distribution of the modularity value is a function of the number of communities, while the modularity density value is obtained from the normalized frequency of the modularity value. Chen [17] proposed modification of density modularity by adding split penalty component. Modularity density is the result of modularity reduced by a split penalty. Modularity calculates a positive effect on grouping a knot together in terms of considering the sides that are between the nodes, while the split penalty calculates the negative effect on ignoring the sides that join different community members. Furthermore, Chen *et al.* [62] also added a variety of modularity density by changing the community link density.

Botta and Genio [63] inspired by the proposed modularity maximization algorithm in [64], added new functionality to the proposed density of modularity by Chen. The associated algorithm for community detection is claimed to have quadratic computational complexity. Costa *et al.* [65] introduced the exact modularity density solution by proposing and comparing several MILP reformulations. Furthermore, Sato and Izunaga [66] developed a solution model with subproblem approach to column creation.

Izunaga *et al.* [67] proposed density modularity as a variant of semidefinite programming, and shows that its relaxation problem provides an upper limit to the optimal density of modularity. They also propose a lower bound algorithm based on a combination of spectral heuristics and dynamic programming. However, although better than modularity, density modularity does not completely overcome the resolution limit. Modularity density is not optimally used for ring laticces network and tree structures [63].

3.4. Creating new quality metrics as a substitute for modularity

Another strategy for overcoming resolution limits is to offer new metrics as an alternative to modularity. Biswas and Biswas [68] proposed a metric based on the nature of its social community formation, in contrast to the modularity which is developed based on the density of its connectivity. The idea is that people who have strong relationships (have the same personality) tend to experience unification. On the other hand, people who have weak relationships (different personalities) tend to experience isolation. The author proposes three quality metrics for this, namely average unifiablity (AVU), average isolability (AVI), and average unifiablity dan isobality (ANUI) which has the ability besides measuring quality as well as maintaining accuracy. Unifiability is a measure of the tendency for multiple clusters to become a single cluster, whereas isobality is a measure of a community to isolate itself from other communities.

Shang *et al.* [34] proposed predictable metrics related to prediction links as an alternative to modularity. This is due to the definition of a weak community, with the consideration that links are more predictable within a community than between communities. Predictability reveals high linkage forecasts for communities, whereas modularity reveals high link density for communities. The author claimed that predictive metrics are more robust than modularity.

Gharaghooshi *et al.* [56] proposed an approach based on the definition of weak community links and strong community links. The author proposes a new objective function, which is called strong inside, weak outside (SIWO) which encourages adding strong links to the community while avoiding weak links. This process is intended so that finding a community can avoid resolution limitations. The time complexity of this new method is linear in the number of edges. The author claimed to be an effective approach for various real and artificial data sets with large and small communities. The drawbacks of this method do not have a standard because it must have its alternative definition of community or objective function. This method also does not explain the extent to which the resolution limit can be overcome.

3.5. Improving the louvain algorithm

Since the modularity-based algorithm was introduced by Newman, many researchers have tried to improve computational time and community quality. Clauset *et al.* [39] have been developed algorithms based on modularity optimization, including clauset-newman-moore (CNM) algorithm [39] with a greedy strategy approach in the optimization process. Blondel *et al.* [69] proposed the Louvain algorithm. Louvain algorithm has two phases, the first phase is the modularity optimization process, then the second phase is community aggregation.

In literature, Louvain's algorithm is the most robust modularity optimization algorithm, so it is often referred to for further research [20], [70]-[73]. Forster [70] developed the Louvain algorithm to support parallel computing. Gach and Hao [71] proposes Louvain+ with improvements to Louvain's second step. The researchers pruning the node in the Louvain's first step [45], [72]. Waltman and Van Eck [73] proposed the smart local moving (SML) lgorithm by modifying the Louvain's first step. Traag *et al.* [74] proposed the Leiden algorithm by modifying the Louvain's first step with a random neighbor approach.

In addition, efforts to speed up computing time, some researchers combining the Louvain algorithm with the Label Propagation algorithm [8], [75], [76], and Zhu *et al.* [77] combining the Louvain algorithm with the concept of k-plex on the network. Li *et al.* [44] proposed an iterative greedy algorithm using the Louvain algorithm as the initial step. This method has good computational time and is suitable for large data sizes. The weakness of this method is that it only focuses on the structure of the graph, ignoring the information content so that it does not represent the real world [78].

3.6. Involving node attributes in determining community detection

Community detection problems usually only involve network structure but ignore node attributes/features, although the majority of real-world social networks provide additional information about actors such as gender, occupation, and interests. Community detection termonilogy focuses on the strength of the network structure, while clustering focuses on the similarity of node attributes [79]. By involving the attributes in the node, it is believed to be able to clarify and enrich the knowledge of the actors and provide a deeper understanding to the community of detection results. In other words, methods that involve network structures and attributes are expected to produce a more informative and qualified community [80], [81]. The general formula for improving the quality of community detection on a network with its node attributes is how to partition the network into communities in such a way that nodes in the same community are not only strongly connected to each other but also show a high degree of attribute homogeneity [82].

Chunaev [80] classified the method of combining structures and attributes on a network CD with attribute nodes into 3, namely, combining structures and attributes at the beginning before the CD process, simultaneous merging, namely combining structures and attributes simultaneously with the CD process, merging at the end, namely first break down the structure and attributes separately, then combine the results obtained. The method of modifying the objective function is part of the simultaneous merging method. The basic idea is that the objective function of modularity is initially applied to structures only, then modified to become structures and attributes are used together to optimize the process. For example, research modifies Louvain's algorithm, in which the initial objective function is modularity which only considers structure added entropy which considers attribute information, then applies the Louvain algorithm simultaneously so that the optimal solution is obtained, namely maximizing modularity and minimizing entropy.

Several studies have modified Louvain's objective function with add a node attribute component. Some researchers formulated a new objective function, as a combination of the Louvain (structure) Qs modularity, with the similarity attribute with the attribute modularity (Qatr) [83], [84], [85]. Other researchers also combined structural modularity with the entropy attribute [82], [86]. Entropy is a measure of the regularity of a set of information content. The more irregular, the higher the entropy. Sets that have similar elements have low entropy. The optimum objective function is when the maximum modularity is achieved, while the entropy is minimum.

Combe *et al.* [81] proposed the I-louvain method with objective functions as a combination of modularity with inertia attributes. Inertia attributes are attributes that are not only categorical types but handle numeric types. Optimal objective function when maximum modularity and inertial are achieved. Singh *et al.* [87] proposed a new objective function as a combination of the Louvain-and-attribut and Louvain-or_attribut methods, which combines Louvain modularity with dependence on similarity attributes and considers irrelevant attributes (outlier). A different approach is taken in combining the structure with node attributes with mathematical programming [88], and with spectral clustering [89]. The drawback of this strategy is that it is difficult to achieve a trade-off between the similarity of node attributes and the connection density in finding communities, and often attributes that seem irrelevant actually reduce accuracy so it requires computation time [84]. Moreover, there is no generally accepted opinion on the effect of combining structures and attributes on whether or not it is useful, particularly on social networks linked to nodes [80].

3.7. Extending the single objective function into a multi-objective function

The CDMO problem has one objective function, namely maximizing modularity, called the single optimization objective function (SOP) [90]. SOP can be formulated by specifying a partition C*. Where Ω is the set of the fisible partitions, C is the community structure, and P is the size function to be optimized:

$P: \Omega \to R$, and $P(C^*) = \min_{c \epsilon \Omega} P(C)$

The weakness of SOP is only optimizing one criterion, so it can cause a fundamental difference that different algorithms can produce different solutions even on the same network. For instance, the objective function in optimization has a resolution limit problem. To overcome this problem, it was extended to multi objective optimization (MOP). Various researchers developed MOP by determining different P1, P2, .., Pn criteria. For instance, Shi taked the criteria of modularity (Q), cut-function, description length [90]. Huq *et al.* [91] compared the results of 2 MOPs with the K means kernel criteria, modularity with MOP with the criteria of community fitness, community score, and modularity. Pizzuti and Socievole [92] used MOP in the combined structure and node attributes where the criteria for the structure are modularity, community score, conductance, while attribute similarity is jaccard, euclidean-based similarity, Chen *et al.* [93] combined modularity density and NMI simultaneously with a local search approach.

The limitation of MOP is that because all data elements need to be explored to represent candidate solutions, outlier data often appears and with improper handling causes interference. Most methods with this strategy can find a better community structure. However, this strategy has high computational complexity and is not suitable for large-scale complex networks [44], [94].

4. ANALYSIS OF STRATEGIES TO IMPROVE THE QUALITY OF CDMO

From the description in section 3, we derive two approaches to six strategies to improve the quality of CDMO (see Figure 2). The first approach is to improve modularity metrics related to the resolution limit issue. This approach has four strategies, namely i) creating multi-resolution, ii) creating local-modularity, iii) creating modularity density, and iv) creating a new metrics quality. These four strategies are generally used independently and are not suitable to be combined because they have opposite properties. For example, making multiresolution suitable for small data sizes, as opposed to local modularity suitable for large data sizes.

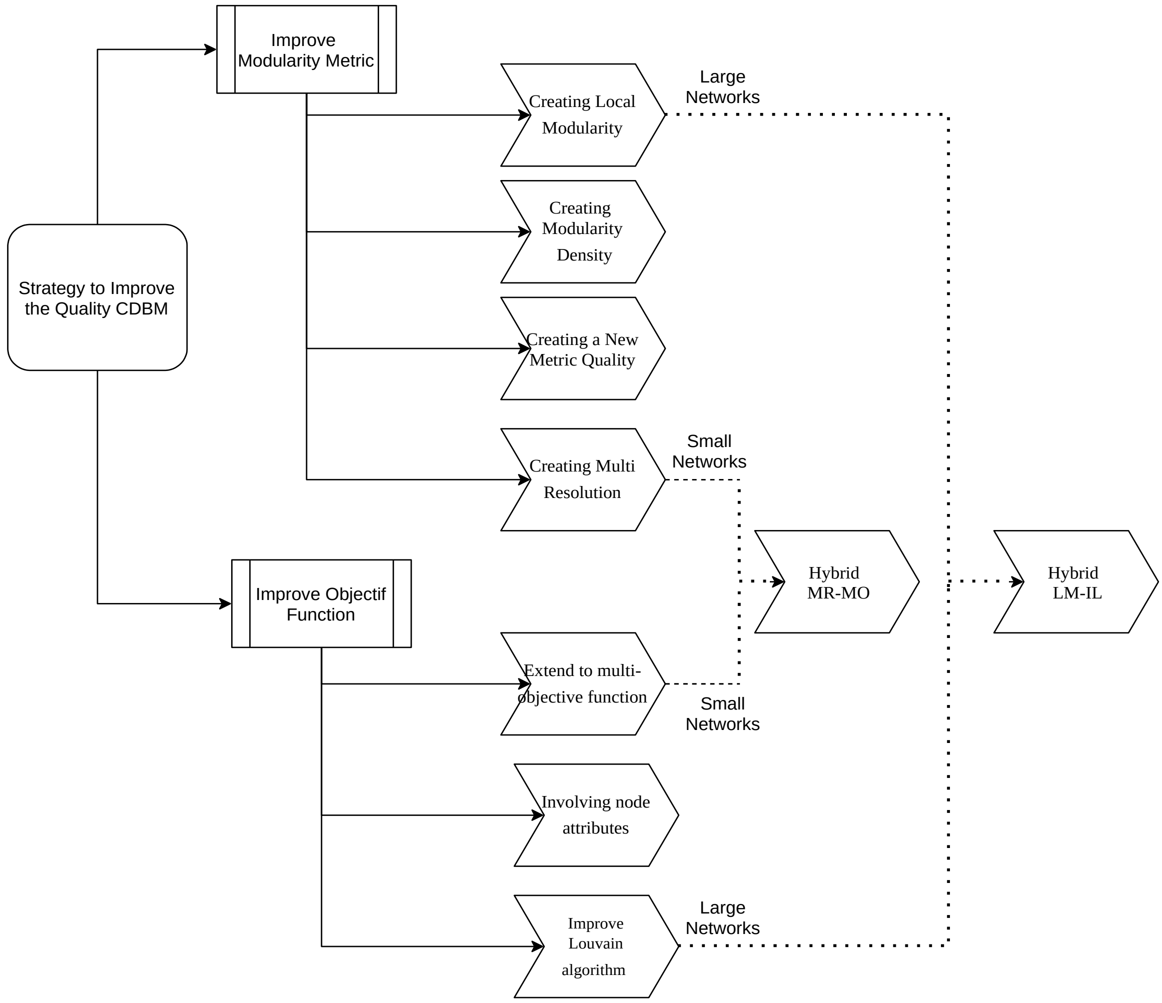

Figure 2. Strategies to improve the quality of CDMO

In the second approach, there are issues related to efforts to improve the objective function. This approach has three strategies that can be used, namely i) improving the Louvain algorithm, ii) involving node attributes, and iii) increasing the single objective to multi-objective. These three strategies are not mutually exclusive but can be combined so that a hybrid method emerges with the aim of producing better quality modularity or reducing computational time. As an example, Citraro and Rossetti [95] proposed the Eva algorithm as a hybrid strategy for applying the Louvain algorithm to attribute graphs. The author used the purity metric as a measure of the homogeneity of information. Li *et al.* [96] proposed an evolutionary algorithm as a combination of multi-objective development strategies on network attributes.

5. CONCLUSION

Based on the description of the strengths and limitations of the six strategies, conclusions can be drawn that there is no always best approach and strategy, because it depends on the size of the data, the size distribution of the community and the type of data. In the approach to increasing the modularity metric, it is found that the larger the network size and the size of the community distribution, the best ranking strategies are i) local-modularity, ii) modularity density, and iii) multiresolution. The development of modularity density will be better as long as the network type is not tree or not laticces. Research to create new community detection quality metrics is needed to implement community detection in other studies involving the community. In the approach of increasing the objective function, it is found that the larger the network size, the best ranking is i) modifying the Louvain algorithm, ii) involving an attribute graph, and iii) multi objective function. Multi-objective development strategy to be suitable for data that does not contain outliers.

ACKNOWLEDGEMENTS

The research is supported by Ministry of Higher Education Malaysia under the Fundamental Research Grant Scheme FRGS/1/2020/ICT02/UKM/01/2.

REFERENCES

[1] A. Mahmoudi, M. R. Yaakub, and A. A. B. U. Bakar, "The relationship between online social network ties," vol. 13, no. 3, pp. 1-15, 2019, doi: 10.1145/3314204.
[2] X. Li *et al.*, "Higher-order multiple-feature-based community evolution model with potential applications in criminal network investigation," *Future Generation Computer Systems*, vol. 125, pp. 364–375, 2021, doi: 10.1016/j.future.2021.06.028.
[3] H. Elgazzar, K. Spurlock, and T. Bogart, "Evolutionary clustering and community detection algorithms for social media health surveillance," *Machine Learning with Applications*, vol. 6, pp. 1-11, 2021, doi: 10.1016/j.mlwa.2021.100084.
[4] A. F. Al-Mukhtar and E. S. Al-Shamery, "Community detection of political blogs network based on structure-attribute graph clustering model," *International Journal of Electrical and Computer Engineering*, vol. 9, no. 3, pp. 2121–2130, 2019, doi: 10.11591/ijece.v9i3.pp2121-2130.
[5] T. Setiadi, M. R. Yaakub, and A. A. Bakar, "Community detection methods in library's books and borrowers social network segmentation," *Journal of Advances in Information Technology*, vol. 14, no. 6, pp. 1177–1185, 2023, doi: 10.12720/jait.14.6.1177-1185.
[6] H. Yuliansyah, Z. A. Othman, and A. A. Bakar, "Taxonomy of link prediction for social network analysis: a review," *IEEE Access*, vol. 8, pp. 183470-183487, 2020, doi: 10.1109/ACCESS.2020.3029122.
[7] I. S. Ahmad, A. A. Bakar, and M. R. Yaakub, "Movie revenue prediction based on purchase intention mining using YouTube trailer reviews," *Information Processing and Management*, vol. 57, no. 5, 2020, doi: 10.1016/j.ipm.2020.102278.
[8] S. Yazdanparast, M. Jamalabdollahi, and T. C. Havens, "Linear time community detection by a novel modularity gain acceleration in label propagation," *IEEE Transactions on Big Data*, vol. 7, no. 6, pp. 961–966, 2021, doi: 10.1109/TBDATA.2020.2995621.
[9] M. F. Shirjini, S. Farzi, and A. Nikanjam, "MDPCluster: a swarm-based community detection algorithm in large-scale graphs," *Computing*, vol. 102, no. 4, pp. 893–922, 2020, doi: 10.1007/s00607-019-00787-4.
[10] R. George, K. Shujaee, M. Kerwat, Z. Felfli, D. Gelenbe, and K. Ukuwu, "A comparative evaluation of community detection algorithms in social networks," *Procedia Computer Science*, vol. 171, pp. 1157–1165, 2020, doi: 10.1016/j.procs.2020.04.124.
[11] S. Fortunato and D. Hric, "Community detection in networks: A user guide," *Physics Reports*, vol. 659, pp. 1–44, 2016, doi: 10.1016/j.physrep.2016.09.002.
[12] R. Mittal and M. P. S. Bhatia, "Classification and comparative evaluation of community detection algorithms," *Archives of Computational Methods in Engineering*, vol. 28, no. 3, pp. 1417–1428, 2021, doi: 10.1007/s11831-020-09421-5.
[13] V. L. Dao, C. Bothorel, and P. Lenca, "Community structure: A comparative evaluation of community detection methods," *Network Science*, vol. 8, no. 1, pp. 1-41, 2020, doi: 10.1017/nws.2019.59.
[14] X. You, Y. Ma, and Z. Liu, "A three-stage algorithm on community detection in social networks," *Knowledge-Based Systems*, vol. 187, pp. 1-23, 2020, doi: 10.1016/j.knosys.2019.06.030.
[15] M. E. J. Newman, "Equivalence between modularity optimization and maximum likelihood methods for community detection," *Physical Review E*, vol. 94, no. 5, pp. 1-9, 2016, doi: 10.1103/PhysRevE.94.052315.
[16] S. Chattopadhyay, T. Basu, A. K. Das, K. Ghosh, and C. A. Murthy, "A similarity based generalized modularity measure towards effective community discovery in complex networks," *Physica A: Statistical Mechanics and its Applications*, vol. 527, 2019, doi: 10.1016/j.physa.2019.121338.
[17] M. Chen, T. Nguyen, and B. K. Szymanski, "A new metric for quality of network community structure," *ASE Human Journal*, vol. 2, no. 4, pp. 226-240, 2013.
[18] B. H. Good, Y. A. De Montjoye, and A. Clauset, "Performance of modularity maximization in practical contexts," *Physical Review*

E - Statistical, Nonlinear, and Soft Matter Physics, vol. 81, no. 4, pp. 1–19, 2010, doi: 10.1103/PhysRevE.81.046106.
[19] T. Chakraborty, A. Dalmia, A. Mukherjee, and N. Ganguly, "Metrics for community analysis: a survey," *ACM Computing Surveys*, vol. 50, no. 4, pp. 1-37, 2017, doi: 10.1145/3091106.
[20] A. Mahmoudi, A. A. Bakar, M. Sookhak, and M. R. Yaakub, "A temporal user attribute-based algorithm to detect communities in online social networks," *IEEE Access*, vol. 8, pp. 154363–154381, 2020, doi: 10.1109/ACCESS.2020.3018941.
[21] U. Brandes *et al.*, "Maximizing modularity is hard," *arXiv- Physics*, pp. 1-10, 2006.
[22] M. Jebabli, H. Cherifi, C. Cherifi, and A. Hamouda, "Community detection algorithm evaluation with ground-truth data," *Physica A: Statistical Mechanics and its Applications*, vol. 492, pp. 651–706, 2018, doi: 10.1016/j.physa.2017.10.018.
[23] M. E. J. Newman, "Detecting community structure in networks," *European Physical Journal B*, vol. 38, no. 2, pp. 321–330, 2004, doi: 10.1140/epjb/e2004-00124-y.
[24] J. Yang and J. Leskovec, "Defining and evaluating network communities based on ground-truth," *IEEE International Conference on Data Mining, ICDM*, pp. 745–754, 2012, doi: 10.1109/ICDM.2012.138.
[25] S. Fortunato, "Community detection in graphs," *Physics Reports*, vol. 486, no. 3–5, pp. 75–174, 2010, doi: 10.1016/j.physrep.2009.11.002.
[26] D. J. Watts, and S. H. Strogatz, "Collectivedynamics of 'small-world' networks," *Nature*, vol. 393, no. 4, pp. 440–442, 1998.
[27] R. Aldecoa and I. Marín, "Deciphering network community structure by surprise," *PLoS ONE*, vol. 6, no. 9, pp. 1-8, 2011, doi: 10.1371/journal.pone.0024195.
[28] A. Shiri, "Introduction to modern information retrieval (2nd edition)," *Library Review*, vol. 53, no. 9, pp. 462–463, 2004, doi: 10.1108/00242530410565256.
[29] J. Artiles, J. Gonzalo, and S. Sekine, "The SemEval-2007WePS evaluation: Establishing a benchmark for the web people search task," *ACL 2007 - SemEval 2007 - Proceedings of the 4th International Workshop on Semantic Evaluations*, pp. 64–69, 2007.
[30] L. Hubert and P. Arabie, "Comparing partitions," *Journal of Classification*, vol. 2, no. 1, pp. 193–218, 1985, doi: 10.1007/BF01908075.
[31] L. Danon, A. Diaz-Guilera, J. Duch, and A. Arenas, "Comparing community structure identification," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2005, no. 09, p. P09008, 2005, doi: 10.1088/1742-5468/2005/09/P09008
[32] H. Almeida, D. Guedes, W. Meira, and M. J. Zaki, "Is there a best quality metric for graph clusters?," *Machine Learning and Knowledge Discovery in Databases*, Springer: Berlin, Heidelberg, pp. 44–59, 2011, doi: 10.1007/978-3-642-23780-5_13.
[33] S. Kaur, S. Singh, S. Kaushal, and A. K. Sangaiah, "Comparative analysis of quality metrics for community detection in social networks using genetic algorithm," *Neural Network World*, vol. 26, no. 6, pp. 625–641, 2016, doi: 10.14311/NNW.2016.26.036.
[34] K. K. Shang, M. Small, Y. Wang, D. Yin, and S. Li, "A novel metric for community detection," *Europhysics Letters*, vol. 129, no. 6, pp. 1-7, 2020, doi: 10.1209/0295-5075/129/68002.
[35] N. Dakiche, F. Benbouzid-Si Tayeb, Y. Slimani, and K. Benatchba, "Tracking community evolution in social networks: A survey," *Information Processing & Management*, vol. 56, no. 3, pp. 1084–1102, 2019.
[36] J. Reichardt and S. Bornholdt, "Detecting fuzzy community structures in complex networks with a potts model," *Physical Review Letters*, vol. 93, no. 21, pp. 1-4, 2004, doi: 10.1103/PhysRevLett.93.218701.
[37] M. E. J. Newman and M. Girvan, "Finding and evaluating community structure in networks," *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, vol. 69, no. 2, pp. 1-15, 2004, doi: 10.1103/PhysRevE.69.026113.
[38] U. Brandes *et al.*, "On modularity clustering," *IEEE Transactions on Knowledge and Data Engineering*, vol. 20, no. 2, pp. 172–188, 2008, doi: 10.1109/TKDE.2007.190689.
[39] A. Clauset, M. E. J. Newman, and C. Moore, "Finding community structure in very large networks," *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*, vol. 70, no. 6, pp. 1-6, 2004, doi: 10.1103/PhysRevE.70.066111.
[40] P. D. Meo, E. Ferrara, G. Fiumara, and A. Provetti, "Generalized Louvain method for community detection in large networks," *International Conference on Intelligent Systems Design and Applications, ISDA*, pp. 88–93, 2011, doi: 10.1109/ISDA.2011.6121636.
[41] V. A. Traag, "Faster unfolding of communities: Speeding up the Louvain algorithm," *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, vol. 92, no. 3, pp. 1-10, 2015, doi: 10.1103/PhysRevE.92.032801.
[42] S. Ghosh *et al.*, "Distributed louvain algorithm for graph community detection," *2018 IEEE 32nd International Parallel and Distributed Processing Symposium, IPDPS 2018*, pp. 885–895, 2018, doi: 10.1109/IPDPS.2018.00098.
[43] L. Chaudhary and B. Singh, "Community detection using maximizing modularity and similarity measures in social networks," in *S mart Systems and IoT: Innovations in Computing*, Springer: Singapore, pp. 197–206, 2020, doi: 10.1007/978-981-13-8406-6_20.
[44] W. Li, Q. Kang, H. Kong, C. Liu, and Y. Kang, "A novel iterated greedy algorithm for detecting communities in complex network," *Social Network Analysis and Mining*, vol. 10, no. 1, 2020, doi: 10.1007/s13278-020-00641-y.
[45] G. S. Carnivali, A. B. Vieira, A. Ziviani, and P. A. A. Esquef, "CoVeC: Coarse-grained vertex clustering for efficient community detection in sparse complex networks," *Information Sciences*, vol. 522, pp. 180–192, 2020, doi: 10.1016/j.ins.2020.03.004.
[46] X. S. Zhang *et al.*, "Modularity optimization in community detection of complex networks," *Europhysics Letters*, vol. 87, no. 3, pp. 1-6, 2009, doi: 10.1209/0295-5075/87/38002.
[47] J. Zhang, K. Zhang, X. K. Xu, C. K. Tse, and M. Small, "Seeding the kernels in graphs: Toward multi-resolution community analysis," *New Journal of Physics*, vol. 11, pp. 1-12, 2009, doi: 10.1088/1367-2630/11/11/113003.
[48] J. Xiang and K. Hu, "Limitation of multi-resolution methods in community detection," *Physica A: Statistical Mechanics and its Applications*, vol. 391, no. 20, pp. 4995–5003, 2012, doi: 10.1016/j.physa.2012.05.006.
[49] A. Arenas, A. Fernández, and S. Gómez, "Analysis of the structure of complex networks at different resolution levels," *New Journal of Physics*, vol. 10, pp. 1-22, 2008, doi: 10.1088/1367-2630/10/5/053039.
[50] J. Xiang *et al.*, "Multi-resolution community detection based on generalized self-loop rescaling strategy," *Physica A: Statistical Mechanics and its Applications*, vol. 432, pp. 127–139, 2015, doi: 10.1016/j.physa.2015.03.006.
[51] K. R. Žalik and B. Žalik, "A local multiresolution algorithm for detecting communities of unbalanced structures," *Physica A: Statistical Mechanics and its Applications*, vol. 407, pp. 380–393, 2014, doi: 10.1016/j.physa.2014.03.059.
[52] J. M. Kumpula, J. Saramäki, K. Kaski, and J. Kertész, "Limited resolution and multiresolution methods in complex network community detection," *Noise and Stochastics in Complex Systems and Finance*, vol. 6601, pp. 1-9, 2007, doi: 10.1117/12.725560.
[53] S. Muff, F. Rao, and A. Caflisch, "Local modularity measure for network clusterizations," *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, vol. 72, no. 5, pp. 1-5, 2005, doi: 10.1103/PhysRevE.72.056107.
[54] J. Xiang *et al.*, "Local modularity for community detection in complex networks," *Physica A: Statistical Mechanics and its Applications*, vol. 443, pp. 451–459, 2015, doi: 10.1016/j.physa.2015.09.093.
[55] P. Ronhovde and Z. Nussinov, "Local resolution-limit-free Potts model for community detection," *Physica A: Statistical Mechanics and its Applications*, vol. 443, pp. 451–459, Mar. 2008, doi: 10.1103/PhysRevE.81.046114.

[56] S. Z. Gharaghooshi, O. R. Zaïane, C. Largeron, M. Zafarmand, and C. Liu, "Addressing the resolution limit and the field of view limit in community mining," in *Advances in Intelligent Data Analysis XVIII*, Springer: Cham, 2020, pp. 210–222, doi: 10.1007/978-3-030-44584-3_17.

[57] J. Xiang *et al.*, "Comparing local modularity optimization for detecting communities in networks," *International Journal of Modern Physics C*, vol. 28, no. 6, 2017, doi: 10.1142/S012918311750084X.

[58] Z. Li, S. Zhang, R. S. Wang, X. S. Zhang, and L. Chen, "Quantitative function for community detection," *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, vol. 77, no. 3, pp. 1-9, 2008, doi: 10.1103/PhysRevE.77.036109.

[59] A. Costa, "Some remarks on modularity density," *arXiv-Computer Science*, pp. 1-12, 2014.

[60] A. Costa, "MILP formulations for the modularity density maximization problem," *European Journal of Operational Research*, vol. 245, no. 1, pp. 14–21, 2015, doi: 10.1016/j.ejor.2015.03.012.

[61] E. Holmström, N. Bock, and J. Brännlund, "Modularity density of network community divisions," *Physica D: Nonlinear Phenomena*, vol. 238, no. 14, pp. 1161–1167, 2009, doi: 10.1016/j.physd.2009.03.015.

[62] T. Chen, P. Singh, and K. E. Bassler, "Network community detection using modularity density measures," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2018, no. 5, pp. 1-17, 2018, doi: 10.1088/1742-5468/aabfc8.

[63] F. Botta and C. I. D. Genio, "Finding network communities using modularity density," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2016, no. 12, pp. 1-33, 2016, doi: 10.1088/1742-5468/2016/12/123402.

[64] S. Treviño, A. Nyberg, C. I. Del Genio, and K. E. Bassler, "Fast and accurate determination of modularity and its effect size," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2015, no. 2, pp. 1-23, 2015, doi: 10.1088/1742-5468/2015/02/P02003.

[65] A. Costa, T. S. S. Ng, and L. X. X. Foo, "Complete mixed integer linear programming formulations for modularity density based clustering," *Discrete Optimization*, vol. 25, pp. 141–158, 2017, doi: 10.1016/j.disopt.2017.03.002.

[66] K. Sato and Y. Izunaga, "An enhanced MILP-based branch-and-price approach to modularity density maximization on graphs," *Computers and Operations Research*, vol. 106, pp. 236–245, 2019, doi: 10.1016/j.cor.2018.01.012.

[67] Y. Izunaga, T. Matsui, and Y. Yamamoto, "A doubly nonnegative relaxation for modularity density maximization," *Discrete Applied Mathematics*, vol. 275, pp. 69–78, 2020, doi: 10.1016/j.dam.2018.09.023.

[68] A. Biswas and B. Biswas, "Defining quality metrics for graph clustering evaluation," *Expert Systems with Applications*, vol. 71, pp. 1–17, 2017, doi: 10.1016/j.eswa.2016.11.011.

[69] V. D. Blondel, J. L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2008, no. 10, pp. 1-12, 2008, doi: 10.1088/1742-5468/2008/10/P10008.

[70] R. Forster, "Louvain community detection with parallel heuristics on GPUs," *INES 2016 - 20th Jubilee IEEE International Conference on Intelligent Engineering Systems, Proceedings*, pp. 227–232, 2016, doi: 10.1109/INES.2016.7555126.

[71] O. Gach and J. K. Hao, "Improving the Louvain algorithm for community detection with modularity maximization," in *Artificial Evolution*, Springer: Cham, pp. 145–156, 2014, doi: 10.1007/978-3-319-11683-9_12.

[72] N. Ozaki, H. Tezuka, and M. Inaba, "A simple acceleration method for the Louvain algorithm," *International Journal of Computer and Electrical Engineering*, vol. 8, no. 3, pp. 207–218, 2016, doi: 10.17706/ijcee.2016.8.3.207-218.

[73] L. Waltman and N. J. Van Eck, "A smart local moving algorithm for large-scale modularity-based community detection," *European Physical Journal B*, vol. 86, no. 11, pp. 1-33, 2013, doi: 10.1140/epjb/e2013-40829-0.

[74] V. A. Traag, L. Waltman, and N. J. van Eck, "From Louvain to Leiden: guaranteeing well-connected communities," *Scientific Reports*, vol. 9, no. 1, pp. 1–12, 2019, doi: 10.1038/s41598-019-41695-z.

[75] B. Hu, W. Li, X. Huo, Y. Liang, M. Gao, and P. Pei, "Improving Louvain algorithm for community detection," in *International Conference on Artificial Intelligence and Engineering Applications (AIEA 2016)*, vol. 63, pp. 110-115, 2016, doi: 10.2991/aiea-16.2016.20.

[76] Y. Ma, Y. Zhao, J. Wang, M. Liu, W. Shen, and Y. Ma, "Modularity-based incremental label propagation algorithm for community detection," *Applied Sciences*, vol. 10, no. 12, pp. 1-13, 2020, doi: 10.3390/app10124060.

[77] J. Zhu, B. Chen, and Y. Zeng, "Community detection based on modularity and k-plexes," *Information Sciences*, vol. 513, pp. 127–142, 2020, doi: 10.1016/j.ins.2019.10.076.

[78] X. Zhang, Z. Ma, Z. Zhang, Q. Sun, and J. Yan, "A review of community detection algorithms based on modularity optimization," *Journal of Physics: Conference Series*, vol. 1069, no. 1, pp. 1-5, 2018, doi: 10.1088/1742-6596/1069/1/012123.

[79] J. Yang, J. McAuley, and J. Leskovec, "Community detection in networks with node attributes," *IEEE International Conference on Data Mining, ICDM*, pp. 1151–1156, 2013, doi: 10.1109/ICDM.2013.167.

[80] P. Chunaev, "Community detection in node-attributed social networks: A survey," *Computer Science Review*, vol. 37, 2020, doi: 10.1016/j.cosrev.2020.100286.

[81] D. Combe, C. Largeron, M. Géry, and E. Egyed-Zsigmond, "I-Louvain: An attributed graph clustering method," *Advances in Intelligent Data Analysis XIV*, Springer: Cham, pp. 181–192, 2015, doi: 10.1007/978-3-319-24465-5_16.

[82] C. Jia, Y. Li, M. B. Carson, X. Wang, and J. Yu, "Node attribute-enhanced community detection in complex networks," *Scientific Reports*, vol. 7, no. 1, pp. 1-15, 2017, doi: 10.1038/s41598-017-02751-8.

[83] T. A. Dang and E. Viennet, "Community detection based on structural and attribute similarities," in *International conference on digital society (ICDS)*, pp. 7–14, Mar. 2012, doi: 10.1080/00949655.2019.1568435.

[84] A. Reihanian, M. R. F. -Derakhshi, and H. S. Aghdasi, "Community detection in social networks with node attributes based on multi-objective biogeography based optimization," *Engineering Applications of Artificial Intelligence*, vol. 62, pp. 51–67, 2017, doi: 10.1016/j.engappai.2017.03.007.

[85] Y. Zhang, E. Levina, and J. Zhu, "Community detection in networks with node features," *Electronic Journal of Statistics*, vol. 10, no. 2, pp. 668–685, Jan. 2016, doi: 10.1214/16-EJS1206.

[86] J. D. Cruz, C. Bothorel, and F. Poulet, "Entropy based community detection in augmented social networks," *Proceedings of the 2011 International Conference on Computational Aspects of Social Networks, CASoN'11*, pp. 163–168, 2011, doi: 10.1109/CASON.2011.6085937.

[87] B. C. Singh, M. M. Rahman, M. S. Miah, and M. K. Baowaly, "Community detection using node attributes and structural patterns in online social networks," *Computer and Information Science*, vol. 10, no. 4, pp. 50-59, 2017, doi: 10.5539/cis.v10n4p50.

[88] E. Alinezhad, B. Teimourpour, M. M. Sepehri, and M. Kargari, "Community detection in attributed networks considering both structural and attribute similarities: two mathematical programming approaches," *Neural Computing and Applications*, vol. 32, no. 8, pp. 3203–3220, 2020, doi: 10.1007/s00521-019-04064-5.

[89] F. Tang and W. Ding, "Community detection with structural and attribute similarities," *Journal of Statistical Computation and Simulation*, vol. 89, no. 4, pp. 668–685, 2019, doi: 10.1080/00949655.2019.1568435.

[90] C. Shi, Z. Yan, Y. Cai, and B. Wu, "Multi-objective community detection in complex networks," *Applied Soft Computing Journal*, vol. 12, no. 2, pp. 850–859, 2012, doi: 10.1016/j.asoc.2011.10.005.

[91] S. T. U. Huq, V. Ravi, and K. Deb, “Evolutionary multi objective optimization algorithm for community detection in complex social networks,” *arXiv-Computer Science*, pp. 1-33, 2020.
[92] C. Pizzuti and A. Socievole, “Multiobjective optimization and local merge for clustering attributed graphs,” *IEEE Transactions on Cybernetics*, vol. 50, no. 12, pp. 4997–5009, 2020, doi: 10.1109/TCYB.2018.2889413.
[93] G. Chen, Y. Wang, and J. Wei, “A new multiobjective evolutionary algorithm for community detection in dynamic complex networks,” *Mathematical Problems in Engineering*, vol. 2013, pp. 1-8, 2013, doi: 10.1155/2013/161670.
[94] L. Martí, J. Garca, A. Berlanga, C. A. Coello Coello, and J. M. Molina, “MB-GNG: Addressing drawbacks in multi-objective optimization estimation of distribution algorithms,” *Operations Research Letters*, vol. 39, no. 2, pp. 150–154, 2011, doi: 10.1016/j.orl.2011.01.002.
[95] S. Citraro and G. Rossetti, “Eva: Attribute-aware network segmentation,” *Studies in Computational Intelligence*, vol. 881, pp. 141–151, 2020, doi: 10.1007/978-3-030-36687-2_12.
[96] Z. Li, J. Liu, and K. Wu, “A multiobjective evolutionary algorithm based on structural and attribute similarities for community detection in attributed networks,” *IEEE Transactions on Cybernetics*, vol. 48, no. 7, pp. 1963–1976, 2018, doi: 10.1109/TCYB.2017.2720180.

BIOGRAPHIES OF AUTHORS

Tedy Setiadi recieved a bachelor's degree in computer science from Universitas Gadjahmada (UGM), Indonesia, in 1992 and a master's degree in informatics engineering from the Institut Teknologi Bandung (ITB), Indonesia in 1999. He is currently pursuing the Ph.D. degree with the Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia (UKM). He has been working as a lecturer with the Department of Informatics, Universitas Ahmad Dahlan (UAD), Indonesia, since 2003. His research includes data mining, graph theory, and social network analysis. He can be contacted at email: tedy.setiadii@tif.uad.ac.id.

Mohd Ridzwan Yaakub is currently a senior lecturer with the Center for Artificial Intelligence and Technology (CAIT), Faculty of Information Science and Technology, National University of Malaysia (FTSM, UKM). His expertise is sentiment analysis/opinion mining, feature selection, feature extraction, ontology, and data mining. A Ph.D. holder from the Queensland University of Technology (QUT), Australia. He is also the head researcher with the Sentiment Analysis Lab, CAIT. He can be contacted at email: ridzwanyaakub@ukm.edu.my.

Azuraliza Abu Bakar received the Ph.D. degree in artificial intelligence from Universiti Putra Malaysia, in 2002. She has been a professor of data mining with the Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, since 2011. She has served as an Advisor for the Data Mining and Optimization Lab and a member of the Sentiment Analysis Lab. She has 13 research projects (including three in progress) and a member of 42 research projects. She is a member of the ieee computational intelligence society. She has also served on roughly 30 conference and workshop program committees and served as the program chair. Her main research interests include data analytics and artificial intelligence specifically in rough set theory, feature selection, selection algorithms, nature inspired computing, and sentiment analysis. She can be contacted at email: azuraliza@ukm.edu.my.