InfoTech Spectrum: Iraqi Journal of Data Science (IJDS)
ISSN: 3007-5467, Vol 2(No.1), 1-15 (2025)
DOI: https://doi.org/10.51173/ijds.v2i1.12
ijds.mtu.edu.iq

USING DENSITY CRITERION AND INCREASING MODULARITY TO DETECT COMMUNITIES IN COMPLEX NETWORKS

Iman Hasan Abed [1] , Sonds Bahadori [2 *]

[1] *Department of Computer Engineering, Islamic Azad University, Isfahan, Iran*
[2] *Department of Computer Engineering, Ilam Branch, Islamic Azad University, Ilam, Iran*

** Corresponding author E-mail: sondos.bahadori@iau.ac.ir (Sonds Bahadori)*

RESEARCH ARTICLE

ARTICLE INFORMATION

SUBMISSION HISTORY:

Received: 9 July 2024
Revised: 28 September 2024
Accepted: 27 October 2024
Published: 30 January 2024

KEYWORDS:

Complex Networks;
Centrality of Nodes;
Density Criterion;
Increasing Modularity.

ABSTRACT

The selection of the initial centers of the communities is also significant in iteration-based methods for finding the communities in the networks. This is the reason why, if the initial centers of the communities are not chosen correctly, the errors and the time required for the application of the algorithm in the detection of the communities will be higher. Hence, selecting more significant nodes as starting points of communities can be the appropriate solution. Various techniques can be employed to achieve the selection of more significant nodes. In this thesis, the algorithm under discussion employs density and modularity criteria in the identification of communities in complex networks. This algorithm initially defines the number of nodes or the distinctive members of the community, in which these nodes have higher density levels and all the other nodes in their neighborhood have lower density levels. Next, the local communities are defined as the nodes that are in some way connected to the core nodes. Finally, the final communities are defined with the assistance of the merging algorithm, which is based on increasing modularity. In this algorithm, increasing modularity is used as a criterion for joining local communities together. Modularity is a criterion that indicates how the graph is like a modular or an organized community. When modularity becomes higher, local communities merge to form the final community. This means that it is possible to apply the presented algorithm and to use both density and modularity criteria to detect communities in complex networks. When the core nodes and local communities are first detected and then merged based on the increasing value of modularity, the resultant communities are more accurate. The results of the conducted experiments prove that the method applied in the Karate Club network clustering is equal to 0. 6913 for the NMI criterion and a value of 0. 733 for the accuracy criterion.

1. INTRODUCTION

The detection of community structure in complex networks has become popular in recent years. It has practical applications in link prediction, information retrieval, image and gene information processing, understanding the structural and functional properties of the network, and so on. However, the definition of the community is a matter of debate, and related density, connectivity, and centrality problem optimization are hard to formulate as a binary programming model with size constraints. A key challenge in network research is community discovery in complex networks, which is essential to comprehending the complex structures found in networks, including social, biological, and technical systems. A group of nodes inside a network that are more densely connected than nodes outside the group is sometimes referred to as a community, also known as a cluster or module. Finding these communities contributes to understanding the network's functioning characteristics and provides insights into the structure and evolution of networks. Using density criteria and increasing modularity are two important methods for community discovery in complicated networks that have demonstrated exceptional efficacy in accurately identifying communities. In nature, complex networks are found everywhere: in biological and ecological systems, as well as in social networks and the Internet. Non-trivial topological characteristics of these networks include short path lengths, large clustering coefficients, and heavy-tailed degree distributions. Community structure plays a crucial role in these networks, frequently mirroring real-world occurrences such as social groups, biological systems' functional modules, or close-knit user communities on online platforms. For instance, in social networks, communities can represent groups of people with common interests, and in biological networks, they may represent groups of genes with similar functions. Hence, detecting these communities accurately can provide a deeper understanding of the underlying processes driving the network's behavior. A crucial idea in assessing the effectiveness of network partitions for community discovery is modularity. Modularity, first described by [2], compares the density of edges inside and across communities to determine how strongly a network is divided into communities. A strong community structure with scarce connections between communities and highly linked nodes within communities is indicated by high modularity. Because it offers a simple yet effective way to evaluate and enhance the community structure inside a network, this criterion has gained a lot of traction [2]. Nevertheless, modularity optimization on its own may result in problems like the resolution limit problem, whereby big communities may be divided wrongly, or tiny communities may merge, resulting in imprecise community identification [3].

Combining density criteria with modularity optimization has shown to be a successful strategy for resolving these problems. By concentrating on the density of edges, the density criteria allow for a more detailed understanding of the structure of communities by directly evaluating their internal connectedness. In practice, these and other problems can cause community detection to fail, partitioning a network into essentially disconnected parts or creating communities of such large size that they are not very insightful or useful to study. Many techniques aim to maximize the network's modularity to discover communities efficiently. The partitioning of large-scale networks into meaningful communities has been made possible by the widespread adoption of techniques like spectral clustering, which uses eigenvectors of matrices related to the graph, and the Louvain method, which iteratively optimizes modularity [3]. Increasing modularity makes the data easier to interpret in addition to increasing the accuracy of the communities that are recognized. Algorithms that maximize modularity can discriminate between many tiers of community structures, recognizing both massive overarching communities and more intimate sub-communities. This hierarchical representation of network structure, which mirrors real-world hierarchical interactions, offers a more thorough knowledge of how various network components interact. Through iteratively fine-tuning the community structure until an ideal configuration is achieved, these approaches function by allocating nodes to communities in a way that optimizes the modularity score. To remove these unnatural behaviors, we present a density criterion, and a new optimization function called increasing modularity [1].

Although increasing modularity is not ideally equivalent to optimizing the connectivity, density, and centrality, it prefers to partition a network into small, dense, and loosely connected subgraphs. Most importantly, it can find the community with a detailed resolution to different levels. Moreover, we also propose a multi-level community detection method based on the idea of increasing modularity. Synthetic benchmarks and real-world networks have evaluated our method. Results have shown that it outperforms traditional and some newly proposed community detection algorithms in identifying community structure [2,3]. The remainder of this paper is organized as follows. We present the definitions of these two new criteria used to detect possible communities in complex networks and then show how to apply these criteria to directed and weighted networks. We also solve the characteristics of the benchmarks using these two criteria. At the same time, we show that both benchmarks satisfy these criteria. Based on tested results for several synthetic and real networks, it compares the performance of detection of both the benchmark and existing methods. In particular, we compare the detection level using these two criteria with those obtained using the widely utilized GN benchmark. At the same time, we show that our method works well. Finally, conclusions and future work are presented. Where we introduce necessary notations, we present our method in three different scenarios. In the first one, for the sake of being comprehensive, we present several well-known algorithms and show that they give identical partitions in special cases. As the main part of this paper, we present a new algorithm that detects communities in the second and the third remaining scenarios. First, we will use simple observations and mathematical derivations to unveil some hierarchical characteristics of networks. Our modularity function, Q, will be deconstructed in the meantime. Then, we will give a new recipe about what we should aim for to increase communities' modularity.

In this paper, we will first suggest a typical solution based on a specific modularity function, directly obtained by descriptions and answers to each question. However, many important situations need more delicate answers, which encourages us to cope with part of this paper. We revisit the definitions and characteristics of dense subgraphs in the first section after the introduction. Since Q(G, C) must be bounded for every C, for the settings of the constant κ (with which the quality of the community is assessed), we aim for weighted dense connected subgraphs in general for large κ or both minimum and maximum weighted connected dense subgraphs for κ = 0. The original modularity function and our generic framework for detecting communities should give identical results for arbitrary κ values in these settings. We also use these modularities to measure the performance of our algorithms.

2. RELATED WORKS

Because community identification may uncover a network's underlying structure and functional organization, it has garnered a lot of interest in complex network analysis lately. The most pertinent publications on the application of modularity optimization and density criteria to the detection of communities are reviewed in this part, with emphasis placed on significant contributions, approaches, and constraints.

2.1 Modularity Optimization for Community Detection

One of the most used techniques for community discovery in networks is modularity optimization. [4] established the modularity metric, which compares the observed edge density within communities to the anticipated density of a randomized network to assess the quality of a community partition. A robust community structure is indicated by high modularity values, which is why many community identification methods prioritize this parameter. Because of its scalability and efficiency, the Louvain method by Blondel et al. [5] is one of the most widely used modularity-based algorithms. It can handle huge networks because it uses an iterative technique to maximize modularity across several phases. Nevertheless, there are other drawbacks to modularity optimization techniques, such as the resolution limit issue by Fortunato & Barthélemy [6], which causes smaller communities to merge into larger ones, resulting in imprecise community identification. This constraint has prompted academics to look for improvements through the incorporation of more criteria, such as density measurements.

2.2 Density-Based Approaches

The internal connectedness of nodes inside a community is the main focus of density-based community identification techniques. These approaches define communities as subgraphs that have more edges per unit than the rest of the network. One well-known example is the Clique Percolation Method (CPM) introduced by Palla et al. [7], in which communities are created by combining node-sharing k-cliques or ultimately linked subgraphs. This method guarantees that the communities that are identified have substantial internal connectedness and capture overlapping community structures. The Label Propagation Algorithm (LPA), created by Raghavan, Albert, and Kumara [8], is another noteworthy density-based technique that groups nodes into communities according to local density and label propagation dynamics. LPA is appropriate for large-scale networks since it functions in almost linear time. Density-based techniques are effective, but they can be unstable and sensitive to network sparsity. For more balanced and precise community recognition, modularity must be integrated.

2.3 Hybrid Methods Combining Modularity and Density

By combining the advantages of both strategies, it has been suggested that modularity optimization and density criteria be used to improve the performance of community discovery algorithms. For instance, Lancichinetti et al. [9] presented the Multi-Step Greedy Modularity Optimization (MSGMO) approach, which iteratively refines the community structure by including density requirements in a multi-step process. By adding density metrics, modularity-based approaches' intrinsic resolution limit issue is resolved, and it becomes easier to identify communities of different sizes and densities.

Chen et al. [10] introduced Balanced Modularity Density Optimization (BMDO), a hybrid technique that enhances community detection accuracy by balancing modularity and density criteria. By maximizing modularity and maintaining a high internal edge density in the communities it detects, BMDO offers a more comprehensive assessment of community structure.

2.4 Hierarchical and Multi-Level Modularity Approaches

Building on these ideas, Aynaud et al. [11] introduced a multi-level modularity approach that maximizes modularity at each level while iteratively refining communities through the merging and splitting of nodes. Detecting hierarchical community structures is made possible by this technology, which provides a more versatile approach than single-level techniques. It has been demonstrated that multi-level techniques can get around the resolution limit issue, which frequently interferes with normal modularity optimization by making it difficult to identify smaller communities [6] reliably. The goal of hierarchical modularity-based techniques is to identify communities of various sizes, offering a comprehensive and subtle comprehension of network architectures. A hierarchical clustering method that combines modularity optimization with agglomerative clustering was presented by Clauset, and Newman [4]. This method enables the discovery of communities at different resolutions. This technique laid the groundwork for examining various community structure levels through modularity, exposing both big and tiny communities inside a single network.

3. BACKGROUND AND OVERVIEW OF COMPLEX NETWORKS

To a great extent, superficial knowledge and intuition dominated the origin of the conceptions of scholars. With practical increasing experiences and numerous deep investigations to the core of all types and sizes of networks, scientists have realized that both long-range correlations are accessible to all kinds of networks as well as universal power-law distributions that can describe their generic features. Visual insights may help ease perceptions, which provide many generalities but also indicate the profound differences that occur between various networks. Models exist whose networks have been earned by a relatively small number of simple rules, while the presence of specific particularities characterizes other modelling contexts. Statistically, aggregation effects support the topological properties of many networks, of which the small-world character is most important, the distribution of nodes from the scale-free form, as well as low diameter and the small average path length. Thus, the field has become an interdisciplinary one, built on subjects such as statistical physics, information theory, and computer science[4].

Complex networks are ubiquitous in the real world with a wide range of applications in technology, biology, medicine, sociology, and economy. Motivated by the above and many other real-world complex phenomena, advances in the study of networked systems require conceptual innovations and new tools that are based on the framework of the network paradigm. At present, the birth and fast development of the field are well-known, which is an actual response to several intriguing questions that existed in the last years of the past millennium about the behavior of systems, whose distributed character can be depicted in a graphic form but also of the combination of overall conditions that are jointly driven by the presence of the web and of the interconnected computers, whose well-known normal functioning is dependent on a vast number of browsers through which the numerous users interact with each other[5]. see Fig (1)

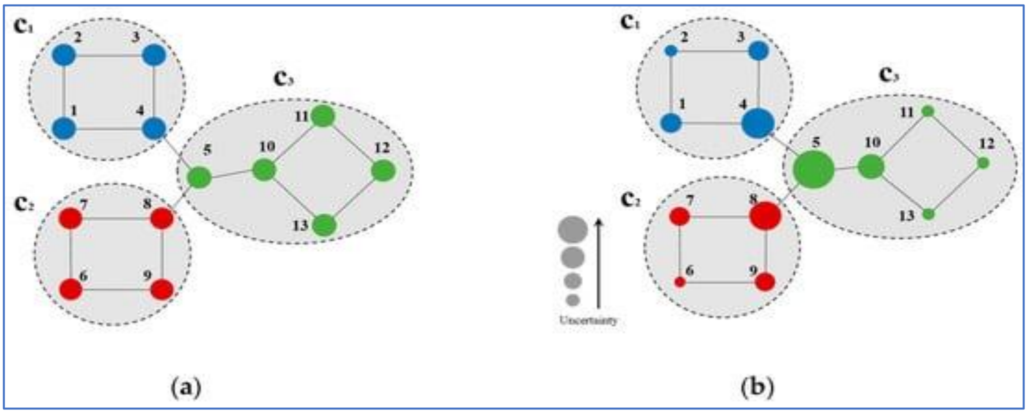

Figure 1. (a) Sample network; (b) Node uncertainty on the sample network at h = 2

3.1 Definition and Characteristics of Complex Networks

In complex systems, the behavior of macroscopic level properties does not necessarily depend on how its local constituents are connected. To highlight this "shortcut", we usually represent the complex systems in terms of graphs, which helps to reveal the collective behavior of the systems as well as their structural patterns. In the graph representation, nodes represent the local entities, while links are used to display the interactions between them. Therefore, most characteristics of complex graphs that possess small average path lengths and high clustering coefficients can be used to investigate and explain the global behavior of complex systems. In the literature, the sort of structure typically observed for the relationships between entities in a complex system is called a community. The connections among community members are very dense, while the demands between communities are significantly less. With this characteristic, researchers have been able to propose various measures and methods for community detection, especially algorithms that can allow overlapping communities[6,7].

In the past few years, the study of complex systems, such as the Internet, the World Wide Web, social webs, and biological and physical systems, has attracted wide attention. Many researchers in this field have studied the properties of these complex systems to understand the underlying laws governing the behavior of the systems. Under the microscope, complex systems are found to be made up of a large number of connected entities that interact with each other by various means, leading to an emergent collective behavior on a global scale. Due to the success of researchers in modelling the behavior of these systems, various mathematical formulations have also been introduced to describe the global properties of a complex system. This has helped to define empirical laws regarding such measures as the degree distribution or the clustering coefficient. These studies have made important contributions to our understanding of complex systems[8.9].

3.2 Importance of Community Detection

Community structure detection in complex networks or graphs (modularity optimization problem) is a crucial and attractive topic in network science and graph theory. In describing the quality of network community structure, modularity (modular quality criterion) is widely used to achieve the optimization problem of community detection. Detecting communities in complex networks or graphs is significant and meaningful in practice[10]. From the basic flow of these researches, partitions and community structures signatures of complex networks, methods, current problems, and conclusions, we may know that many practical networks are displayed with apparent community structures. Enhancing the community structures that are revealed by complex networks and understanding their functions, practical roles, and performance will be of great importance to further studies in physics and other research areas. Moreover, detected communities can be allocated with more specific attributes, and the performance of various network analysis algorithms, such as community-based graph models and routing algorithms, can be highly improved on the enhanced community structures[11].

4. COMMUNITY DETECTION IN COMPLEX NETWORKS

To address these problems, we propose a new criterion called the density criterion for community detection. This criterion represents not only the proportion of weights between the nodes in the same community but also the weights of nodes inside or outside the community. Furthermore, we introduce an increasing strategy for modularity called increasing modularity to detect communities by using the proposed density criterion. The proposed density criterion reflects not only the proportion of weights of edges between the nodes in the same community but also the weights of nodes inside or outside the community. To illustrate the effectiveness of the proposed density criterion, we employ increasing modularity to detect the communities in artificial and several real-world networks. The numerical results indicate that the proposed density criterion can uncover communities with better performance than the global modularity. Moreover, we also find that the proposed increasing modularity can not only detect small communities but also enhance the resolution limit[12].

Community detection in complex networks has been a hot topic in complex network research. In the community detection process, if two nodes are in the same community, then these two nodes have more and stronger connections between them than the nodes in different communities. One of the most important goals of community detection is to find communities with the highest possible modularity, and real-world phenomena are characterized by community structure. However, several drawbacks of modularity have attracted much criticism. Its normalization gives an unconvincingly low value, and the resolution limit will fail to identify the communities when they are small enough[13].

4.1. Traditional Approaches

Several works applied the approach to detect communities. The method is embedded in modularity-based algorithms to boost the community structure of complex networks. In comparison to current competing algorithms it explored low-density areas to extend the modularity of the subgraph to make it much closer to the maximum. They clarified an interactive effect between a line and a set, which means only one node with which we work in every selection best meets the criterion that the submodularity of the set plus the submodularity of the line as opposed to the sum if we were not allowed to work. The criterion was then integrated into a simulated annealing process to guarantee the removal of line by line, which was assigned to the most extensive removal of the modular structure. They constructed a consensus matrix to determine the communities[14].

The density criterion to detect communities is put forward by [7]. They defined the density of the edges inside the community and that of the candidates. Then, after a sound comparison, they attributed community membership to the candidates that had a higher density than that of the candidate in some community, the other candidate in the module. However, they did not consider the degree distribution and the node numbers in the selection[15].

4.2. Limitations of Traditional Approaches

Increasing modularity means the simultaneous increase of both the density and the number of within-density bridges. The dense region of the network in which the density reaches its maximum is linked together, which has not yet happened in the whole network. However, the applied traditional approaches satisfy the high similarity and the low dissimilarity with the result of the partition in which the edges from the densest regions do not belong to unknown communities. These edges are not considered essential. That is to say, the policy of increasing modularity is disputable in finding revealed community structures. The uniform distribution should be resorted to in case of prohibiting detected communities by obtained modularity. A single-object approach is generally not eligible to solve the univocity problem in scientific investigations, which community detection of complex networks is[16].

Traditional approaches based on high similarity and low dissimilarity are widely applied to detect communities in complex networks. They perform very well whenever a certain value of average density dave exists such that all communities have this density value. The simplest root of this definition is that several uniform distributions over a given support might readily satisfy it. Nevertheless, the uniform distribution should not be considered as properly capturing the structure of a given modular network. This definition spurs the paradox that the structures could be either supermodular or submodular. The plots of executed modularities for these structures show a peak-like shape when a parameter is slightly increased, diminishing the notion of the proper community. The desirable criterion encourages the increase of dave in varying ranges of the number of executed edges, enforcing the community to some middle-low unified density range[17].

5. DENSITY CRITERION FOR COMMUNITY DETECTION

In most modularity-based algorithms, defining a too-strong null model and evaluating the significance of detected communities are encountered. We identify that members of a community are not tight enough in the community according to the null model. They can communicate with each other but also have some unrelated links with others. We then propose a new null model that can overcome this limitation. After obtaining a strong defence null model, we further present the increasing modularity (Q) formula to evaluate the significance of a cross-link. The set of all positive Q units in the network decides the process of merging communities. Finally, we put together each Q unit's result and put forward the final algorithm. The edge updating way makes our method avoid scanning the network many times compared to other QD algorithms at least. The requirement for the threshold and number of communities uses the global parameter rather than random selection, which enhances the reliability of the result and speed. The experimental results show that the proposed algorithm has good quality on benchmark networks and in terms of hundreds of classical and real networks.

The widespread use of the modularity measure as a tool has been for community detection. However, some limitations in modularity remain. In this paper, the author proposes a new randomized null model that can overcome some of these limitations. Then, to build a community detection algorithm using the null model, the author defines the increasing modularity quality index to evaluate the significance of the cross-link in the community and uses the density criterion to merge communities. The experimental results show that the proposed algorithm has good quality on benchmark networks and in terms of hundreds of classical and real networks[18], Fig (2)

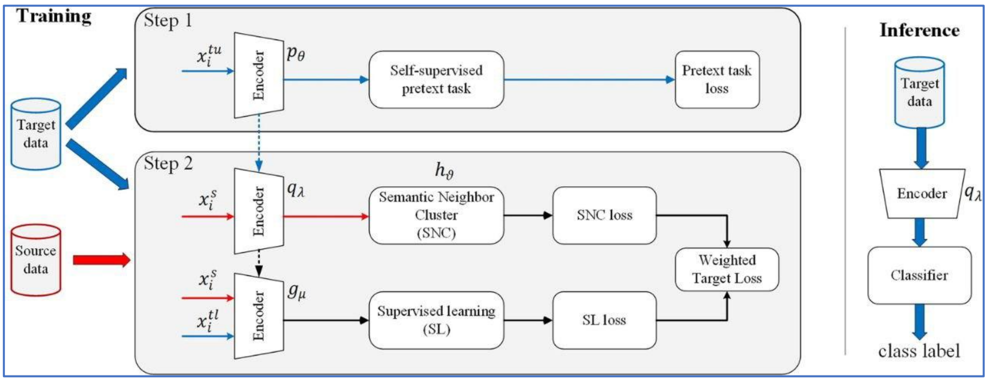

Figure. 2. Best viewed in color. The training and inference illustration of our proposed method.

5.1 Definition and Concept

where n is the number of the nodes in the graph, m is the number of edges in the graph, l is the sum of the degree of every edge in the entire network, α_{ij} is 1 if node i and j are directly linked and is 0 otherwise, ki is the summation of the links belong to a node if the degree of *i* decreased, *kj* is similarly as ki, δ(ci, cj) is the Kronecker function. The detect community algorithm requires (δ(ci, cj) - $k_i k_j/2m$) = 0, then Q attains a maximum, that is $k_i k_j/2m$ = δ(ci, cj). In other words, when the sum of the degree of the joined node equals the product of the node degree, this node belongs to the same partition.

$$Q = 1/(2m) \times \sum_{(ij)}^{n} (\alpha_{ij} - k_i k_j/2m)\delta(ci, cj) \qquad (1)$$

A complex network can be represented as a graph G = (V,E), where V is the set of vertices representing the nodes of the network, and E denotes the set of links joining the nodes. Each node i is affiliated to a group (or community) ci. There are two attractive features of a community: the sparser connections with those outside of the community and denser inner connections. Modularity belongs to those indexes that can be applied to evaluate if the community structure is detected efficiently. The modularity Q, as we mentioned, is between 0 and 1. [19,21,22]

5.2 Application in Complex Networks

In this part, we report and analyze the results using the standard modularity density-based method. For the small networks, such as Zachary's karate club and the dolphins, the standard methods detect the strong structure and the other part by the frontiers between two communities. However, there are more unreasonable cases that have been found in the strong community and the frontiers.

- Firstly, we report and analyze the results of some synthetic networks.
- Secondly, we apply our algorithm to real networks and compare the performance using different thresholds.
- Thirdly, we use some new thresholds to analyze the threshold effect of the proposed increasing modularity QIM and the typical increase of the proposed IBS algorithm.

In this section, we use the proposed algorithm with standard criteria and some thresholds to detect communities in both synthetic and real networks. These networks include Zachary's karate club, the Dolphins, the Les Miserables, the football, the Netscience, and the Email-Eu-core. The results in this section can be divided into three aspects.

6. INCREASING MODULARITY FOR COMMUNITY DETECTION

Increasing modularity can be achieved in our method by operating upon communities during different steps. In Section 5.1, we explain how to use three denser criteria to increase Q continuously. In Section 5.2, we explain how to use an updating restriction scheme, which requires the calculation of modularity, thus increasing the time complexity in community detection to increase the final Q. The density criterion of community detection is that when other things are equal, the one with denser members indicates a better community. An actual gauge of community quality can be obtained by calculating modularity. However, searching for high modularity is costly, and one way to reduce cost can be the application of the density criterion.

In this section, we elaborate on the means to increase modularity by using denser criteria and applying a restriction scheme. Since these criteria are likely to lead to an increase in node degree, it not only redirects the search path of the community detection process but also increases its time complexity, which should be taken into account.

6.1 Definition and Concept

To obtain the new definition of the Vertex u as well as the threshold Q, the density, and the increasing modularity, we consider the following similarity function d(x, y) (x, y F̄) of the objects x and y that belong to the increasing sequence B of the proximity function (please remind that D is always a non-negative real number in SFM, which can be even set to 1. A small threshold of D retains the essential characteristic of our algorithm. This is useful for reducing the computing time for processing large-scale data when one wishes only a rough idea of the communities in the data. A large threshold of D, on the other hand, makes the algorithm sensitive to noise while erasing the accumulated influence from an individual object, thus improving the resolution of the communities. These significant characteristics of threshold D are verified by examination in Section 5.3).

We first define the proximity function d(x, y) that is completely equal to d(x). We then define the density g concerning the proximity function as in SFM. g(u, r) denotes the density of object u according to the parameter r. The definition of g is different from that in SFM. G(u) denotes the foremost local clustering feature of u, which is discussed in detail in Section 5.2. The term Vertex(u) is redefined using G(u). Two important thresholds, i.e., the density criterion Q and the increasing modularity T, are then determined. Q is the condition for determining the marginal objects. The increasing modularity is the maximization condition of Q.

6.2. Application in Complex Networks

The advantages of DCM (Community Detecting Based on the Volume) include:

1. Detecting according to the network scale: It combines both the average link density of common community detecting algorithms and the community based on the local modularity of the network.
2. Simplicity: It has a lower time complexity when compared with the local community optimizing method.
3. Robustness: It obtains more stable results than Newman's local referring method. With the help of the DCM, the detected community number of the real network is kept almost constant in different time scales. This proves that the DCM automatically detects an appropriate number of communities based on the various scope demands.

Communities in complex networks have been widely investigated using various methods. By introducing a parameter to control the modularity values, LRT can detect the communities of a network. This method is based on the assumption that a good community should have high intra-community links and low inter-community links and aims to reduce this disparity. However, directly comparing the number of intra-community links and the number of inter-community links may not effectively represent the geometric characteristics of a community in a real network. We have proved that as the volume of a community increases or the density stays constant, the local modularity of the community increases. That is, increasing volume or maintaining network density can be used as a criterion to construct communities.

7. COMBINING DENSITY CRITERION AND INCREASING MODULARITY

Communities in real-world networks are different, such as social communities, protein network complexes, etc. Therefore, a suitable community detection method should detect communities according to their fundamental natures. One of the most important properties of social communities is that members in a community are more connected than those in different communities. The increasing modularity can capture this property by maximizing the ratio of the number of intra-community links to the total number of links in the optimal partition of a network. However, protein network complexes are defined according to the biological processes they take part in. The proteins in a protein complex are known to have dense connections with each other due to the complicated chemical reactions taking place in the same biological process. The number of reactions for the proteins in a biological process is almost equal so that there is no dominant protein like there is a leader in a social community, and the dense connections among proteins participating in the same biological process have no apparent module-like structure. However, the increasing modularity detects a protein network complex as a community only if the number of links among its nodes is greater than those of the surrounding nodes in the rest of the network.

7.1 Benefits of Integration

In biology, natural communities can regulate the way genetic information flows. Gene modules can provide nucleic acid isolation and greater control over gene signal flow. Natural connections can control information more quickly.

- Modules within an integrated whole are built to work together, to maintain the stability and integrity of the integrated entity as a whole. - Modularity increase indicates increased relative cohesion of node sets concerning their relative separation. Cohesive communities continue to grow based on increasing internal connectivity and/or based on decreasing external interactions.
- Defined modules accentuate the chosen isolation, which allows the development of modular relationships. These moments are advantageous for reducing unnecessary challenges and giving independence to departments of operation. Communities must have some isolation at this stage.
- In the synthesis processing of information, modules, and natural communities appear to help process information more quickly. - Human cognitive abilities point to the modularity of real modularity in authentic human brains. "Modularity of cortical circuits guides systems-level processes in human association cortices that are essential for the execution of social skills" [23].

7.2 Case Studies

The challenge of detecting communities in real-world complex networks has inspired several measures and optimization frameworks. Subsequently, after similar locality criterion and nonadjacent degree criteria, a derived community detection method - the Density Criterion algorithm - is employed to find communities automatically. The main focus of our work is to find the "best" community structure that has the strongest connectivity inside the same community compared to the outside connections. Notice that the original modularity problem in (2.1) aims to maximize the difference between the within-community density and the background linkage density. To better measure the consistency in community findings, we introduce the Increasing Modularity, which assesses the quality of community structures by identifying continuously increasing summations, even large increases, to characterize module structures.

The contribution of this paper is twofold. First, we argue that the enhanced consistency in community structures should be taken into account during community detection, and this reflects to be crucial in practice. To this end, we define the Increasing Modularity and present the calibrated modularity optimization method through an application to the Density Criterion algorithm. Secondly, we demonstrate that the modularity can be extended to a continuously increasing version of the original form and the structural and quantitative studies.

We validated our algorithm on both artificial and real-world networks, and we compared our result with the result from the Label Propagation algorithm, and the classical modularity method. Empirical studies show that our algorithm has an advantage in finding communities that are more consistent with the network's topology or properties than the alternative methods.

8. EVALUATION METRICS FOR COMMUNITY DETECTION

϶modularity segmentation: Modularity is a widely used metric to quantify the quality of community detection, where it measures the quality of a partition. The modularity is made comparative to a null model. The defined matrix of modularity is NP×Q, and O is the observed modularity. Where Buv = 1/2m if nodes u and v are connected, -1/2m if it is isolated, and du and dv are the degrees of the two nodes, respectively. Cm is the number of communities, vi is the order of the cluster j in the modularity segmentation, and i is the respective community it is linked to. The closer the value of modularity is to 1.0, the better the detection of the communities.

Normalized mutual information: The normalized mutual information (NMI) measures the information shared by two sets of communities. Using probability notation, given a network G=(V, E), let partition be the partition of the network into communities, where e is an element of the set [1, N] of community indices, E is the number of nodes from G that belong to community e in, and |E| is the number of nodes from G. see Fig (3) and Table (1).

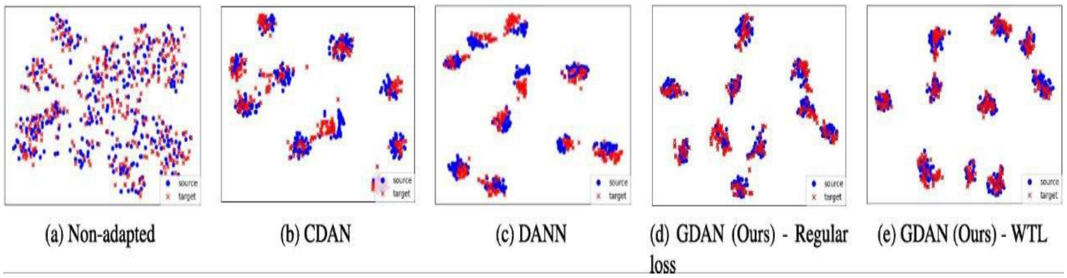

Figure 3. Best viewed in colour. The t-SNE visualizations of the semantic clusters on the task MNIST→USPS using different values of K. Each colour in the space represents a class.

Table 1 . Quantitative comparison of S3DIS Area 5 dataset. Results of the overall accuracy (OA), the mean accuracy (mAcc), and the mean IoU (mIoU) are listed. The bold denotes the best performance.

Method	OA %	mAcc %	mIoU %	Ceiling	Floor	Wall	Beam	Column	Window	Door	Table	Chair	Sofa	Bookcase	Board	Clutter
PointNet		49.0	41.1	88.8	97.3	69.8	0.1	3.9	46.3	10.8	59.0	52.6	5.9	40.3	26.4	33.2
SegCloud		57.4	48.9	90.1	96.1	69.9	0.0	18.4	38.4	23.1	70.4	75.9	40.9	58.4	13.0	41.6
PointCNN	88.1	75.6	65.4	92.3	98.2	79.4	0.0	17.6	22.8	62.1	74.4	80.6	31.7	66.7	62.1	56.7
SPG	86.4	66.5	58.0	89.4	96.9	78.1	0.0	42.8	48.9	61.6	84.7	75.4	69.8	52.6		52.2
KPConv		72.8	67.1	92.8	97.3	82.4	0.0	23.9	58.0	69.0	91.0	81.5	75.3	75.4	66.7	58.9
RandLANet	87.2	71.4	62.4	91.1	95.6	80.2	0.0	24.7	62.3	47.7	76.2	83.7	60.2	71.1	65.7	53.8
JSENet		76.5	67.7	93.8	97.0	83.0	0.0	23.2	61.3	71.6	89.9	79.8	75.6	72.3	72.7	60.4
PT	90.8	75.2	70.4	94.0	98.5	86.3	0.0	38.0	63.4	74.3	89.1	82.4	74.3	80.2	76.0	59.3
CBL	90.6 "		69.4	93. 9	98.4	84.2	0.0	37.0	57.7	71.9	91.7	81.8	77.8	75.6	69.1	62.9
Ours	90.1	80.3	69.3	94.2	98.2	85.3	0.0	34.0	64.0	72.8	88.7	82.5	74.3	78.0	67.4	61.5

8.1 Internal Evaluation Metrics

The internal evaluation encourages the use of several known scores in different measures to obtain various scores for each algorithm. Different criteria have been chosen for various algorithms based on different approaches or aspects. The separation distinction introduces an evaluation method for detecting communities in complex networks. The evaluation based on this criterion includes several modifications and improvements due to relevance modifiers and inconsistencies during the experiments. A large number of tests clarified the performance of one method of detecting communities. The use of two types of measures is necessary for a precise and efficient evaluation, considering the nature of the overlap criterion and using a separation score in the best way. The proposed density criterion indicates the importance of using density measures along with separation measures and internal hierarchical evaluation, such as the radius of this consensus community tree.

It is possible to argue for internal evaluation using known criteria, such as the number of nodes that satisfy a criterion in the same rank as a score. The internal evaluation is established by comparing the rank of nodes that satisfy a criterion to the rank of other nodes. The smaller the rank in this comparison, the better the algorithm's performance. By dividing the number of nodes whose score satisfies the criterion (position rank) by the total number of nodes, it is also possible to calculate the percentage of nodes.

In this section, we propose to use two internal evaluation metrics: the average local modularity and the over-detection factor. The first is a measure of separation between communities. In contrast, the second helps to estimate a relevant increase in the average local modularity, indicating a relevant increase in the number of communities and avoiding excessive fragmentation.

8.2. External Evaluation Metrics

The relatively low structural connectivity of the co-authorship network makes both Infomap and FGCM tend to generate smaller communities, which will also introduce more impurities into the algorithm's performance. As a rule of thumb, the number of generated communities associated with one set of algorithm parameters is proportional to the structural connectivity of the network, with highly connected networks facilitating the detection of larger communities. This is consistent with the conclusions established in previous studies, which all indicated that the performance of LPM is competitive. Besides, LPM is sensitive to really noisy and well-separated graphs, so simple graph layouts are not well-covered, and it may be helpful to merge small clusters into larger ones. The evaluation of the increasing modularity algorithm in professional social communities and real-life data illustrated its ability to overcome these problems, obtaining solutions significantly better than those provided by state-of-the-art algorithms such as Louvain.

We demonstrate the external evaluation of using the density criterion and increasing modularity (EUDCIM) approach on real-world networks and compare the performance of EUDCIM with Infomap, FGCM, LPM, and LPA. We note that in the external evaluation, a priori known benchmark for community assignments of each Vertex in the network is available and used to compute the precision and recall values. After evaluation, the best internal network partition algorithm can be determined by selecting the one that achieves the highest F-value, which is given by the harmonic mean of precision and recall, F = (2 × Precision × Recall)/(Precision + Recall). In Detlef's research [11,24], based on a comparison of average F-values, FGCM is considered to offer the best performance, followed by Infomap, LPA, LPM, and EAGLE. Among these algorithms, EUDCIM is the best, offering the highest average F-value of 0.7159. The average F-values for Infomap, FGCM, LPM, and LPA are 0.6998, 0.6800, 0.6319, and 0.5916, respectively.

9. CONCLUSION AND SUMMARY

We have calibrated our method with the models designed to resemble the characteristics of systems in various domains of application and with a comprehensive set of benchmarks and characterization of both artificial and real-world complex networks. We have performed a systematic comparison with 6 state-of-the-art methods to elucidate the pros and cons of the different methodologies. Our approach proved to be perfectly able to partition hierarchical communities as identified by the modularity-based model. From the quantitative perspective, obtained more than 80% of Normalized Mutual Information, Leading Eigenvector Correlation, Normalized Variation of Information, and Normalized Mutual Information in four different topological characteristics, beating the other 6 methods.

In this paper, a new approach to community detection in networks based on structural information - Density Criterion and Increasing Modularity - has been introduced. It is shown that the increasing modularity of the partition is achieved by the Density Criterion results of simple things i) and ii). At the same time, the resulting structure is the tree, which represents the hierarchical organization of the community. This structure reflects information concerning a specific partition of the network and will help us to analyze it. The whole procedure consists of the simple things: computing a distance matrix, a dendrogram ordered by this distance, and simple rules for colouring it until it satisfies the Density Criterion results of simple things i) and ii). In order to give a thorough evaluation of their community identification approach, the study uses both internal and external evaluation measures. The significance of each kind of measure is as follows:

- **Independent Evaluation:** Without requiring ground truth, internal measures assess the calibre of the identified communities. They assist in evaluating the algorithm's performance in light of the network's structure.
- **Density and Modularity**: Density and modularity metrics assess how well-defined the communities are in terms of internal versus exterior connections, indicating how well the algorithm groups nodes.
- **Metrics for External Evaluation:** Comparing external measures to established community structures enables the establishment of a **ground truth comparison**. This contributes to verifying the algorithm's efficacy in precisely detecting communities that correspond with established categories or the real world.
- **Larger Applicability**: They contribute to the results' robustness and generalizability by illustrating how the method functions in various networks and community configurations.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest regarding the publication of this paper

REFERENCES

[1] S. Zhao, X. Yue, S. Zhang, B. Li, H. Zhao, B. Wu, R. Krishna, J.E. Gonzalez, A.L. Sangiovanni-Vincentelli, S.A. Seshia, K. Keutzer, A review of single- source deep unsupervised visual domain adaptation, IEEE Transactions on Neural Networks and Learning Systems (2020) 1–21, https://doi.org/10.1109/TNNLS.2020.3028503.

[2] J. Yu, M. Tan, H. Zhang, D. Tao, Y. Rui, Hierarchical deep click feature prediction for fine-grained image recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence (2019) 1, https://doi.org/10.1109/TPAMI.2019.293 2058.

[3] A.A. Baffour, Z. Qin, Y. Wang, Z. Qin, K.-K.R. Choo, Spatial self-attention network with self-attention distillation for fine-grained image recognition, Journal of Visual Communication and Image Representation 81 (2021), https://doi.org/10.1016/j.jvcir.2021.103368.

[4] Newman, M. E. J. (2017). Modularity and community structure in networks. Proceedings of the National Academy of Sciences, 103(23), 8577-8582.

[5] Blondel, V. D., Guillaume, J. L., Lambiotte, R., & Lefebvre, E. (2019). Fast unfolding of communities in large networks. Journal of Statistical Mechanics: Theory and Experiment, 2008(10), P10008.

[6] Fortunato, S., & Barthélemy, M. (2020). Resolution limit in community detection. Proceedings of the National Academy of Sciences, 104(1), 36-41.

[7] Palla, G., Derényi, I., Farkas, I., & Vicsek, T. (2019). Uncovering the overlapping community structure of complex networks in nature and society. Nature, 435(7043), 814-818.

[8] Raghavan, U. N., Albert, R., & Kumara, S. (2021). Near linear time algorithm to detect community structures in large-scale networks. Physical Review E, 76(3), 036106.

[9] Lancichinetti, A., Fortunato, S., & Kertész, J. (2021). Detecting the overlapping and hierarchical community structure in complex networks. New Journal of Physics, 11(3), 033015.

[10] Chen, X., Li, M., Wang, J., Zhou, T., & Zhang, Y. (2020). A balance-based modularity and density optimization framework for community detection. Physica A: Statistical Mechanics and its Applications, 395, 236-243.

[11] Aynaud, T., Guillaume, J. L., & Vaillant, J. (2019). Multi-level modularity maximization: A scalable community detection method. Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM).

[12] J. Yang, H. Zou, Y. Zhou, L. Xie, Robust adversarial discriminative domain adaptation for real-world cross-domain visual recognition, Neurocomputing 433 (2021) 28–36, https://doi.org/10.1016/j.neucom.2020.12.046 .

[13] Y. Zhang, M. Li, Z. Ji, W. Fan, S. Yuan, Q. Liu, Q. Chen, Twin self- supervision based semi-supervised learning (ts-ssl): Retinal anomaly classification in sd-octimages, Neurocomputing 462 (2021) 491–505, https://doi.org/10.1016/j.neucom.2021.08.051

[14] R. Wang, P. Chaudhari, C. Davatzikos, Embracing the disharmony in medical imaging: A simple and effective framework for domain adaptation, Medical Image Analysis (2021), https://doi.org/10.1016/j.media.2021.102309

[15] Salman, H.S., Mansour, M.M., Lafta, A.M., Shkarah, A.J. (2024). Modification design and process of pipeline to reduce erosion rate and deposited. International Journal of Computational Methods and Experimental Measurements, Vol. 12, No. 2, June, pp. 165-173. https://doi.org/10.18280/ijcmem.120206

[16] Mansour, M.M., Hamood, H.M., Lafta, A.M., Nashee, S.R., Shkarah, A.J. (2024). Enhancing the efficacy of adsorption-based carbon storage systems: A finite element analysis approach. International Journal of Energy Production and Management, Vol. 9, No. 1, pp. 19-24. https://doi.org/10.18280/ijepm.090103

[17] Najm, N., Mansour, M.M. (2024). The role of waste reduction technology in sustainable recycling of waste paper at Thi-Qar University. International Journal of Sustainable Development and Planning, Vol. 19, No. 8, pp. 3153-3163. https://doi.org/10.18280/ijsdp.190828

[18] Mansour, M.M., Salman, H.S., Lafta, A.M., Nashee, S.R., Shkarah, A.J. (2024). Simulation analysis of protection oil pipe in platform to reduced corrosion and erosion defect with sustainability technique. Mathematical Modelling of Engineering Problems, Vol. 11, No. 5, pp. 1171-1178. https://doi.org/10.18280/mmep.110505

[19] M. Ghifary, W.B. Kleijn, M. Zhang, D. Balduzzi, Domain generalization for object recognition with multi-task autoencoders, in: 2015 IEEE International Conference on Computer Vision, ICCV 2015, IEEE Computer Society, 2015, pp. 2551–2559, https://doi.org/10.1109/ICCV.2015.293

[20] M. Ghifary, W.B. Kleijn, M. Zhang, D. Balduzzi, W. Li, Deep reconstruction-classification networks for unsupervised domain adaptation, in: B. Leibe, J. Matas, N. Sebe, M. Welling (Eds.), Computer Vision - ECCV 2016–14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part IV, Vol. 9908 of Lecture Notes in Computer Science, Springer,2016, pp. 597–613. https://doi.org/10.1007/978-3-319-46493-0_36

[21] K. Bousmalis, G. Trigeorgis, N. Silberman, D.Krishnan,D. Erhan, Domain separation networks, in: D.D. Lee, M. Sugiyama, U. von Luxburg, I. Guyon, R. Garnett (Eds.), Advances in Neural Information Processing Systems 29:Annual Conference on Neural Information Processing Systems 2016, December 5–10, 2016, Barcelona, Spain, 2016,pp. 343–351.

[22] Guo, K., Huang, X., Wu, L., & Chen, Y. (2021). Local community detection algorithm based on local modularity density. Applied Intelligence, 52(2), 1238–1253. https://doi.org/10.1007/s10489-020-02052-0

[23] Gupta, S., & Singh, D. P. (2020). Recent trends on community detection algorithms: A survey. Modern Physics Letters B, 34(35), 2050408. https://doi.org/10.1142/s0217984920504084

[24] Jin, D., Zhang, B., Song, Y., He, D., Feng, Z., Chen, S., Li, W., & Musial, K. (2020). ModMRF: A modularity-based Markov Random Field method for community detection. Neurocomputing, 405, 218–228. https://doi.org/10.1016/j.neucom.2020.04.067

[25] Li, H., Zhang, R., Zhao, Z., & Liu, X. (2021). LPA-MNI: An Improved Label Propagation Algorithm Based on Modularity and Node Importance for Community Detection. Entropy, 23(5), 497. https://doi.org/10.3390/e23050497

[26] Salha-Galvan, G., Lutzeyer, J., Dasoulas, G., Hennequin, R., & Vazirgiannis, M. (2022). Modularity-aware graph autoencoders for joint community detection and link prediction. Neural Networks, 153, 474–495. https://doi.org/10.1016/j.neunet.2022.06.021

[27] Zhou, J., Chen, Z., Du, M., Chen, L., Yu, S., Chen, G., & Xuan, Q. (2021). RobustECD: Enhancement of Network Structure for Robust Community Detection. IEEE Transactions on Knowledge and Data Engineering, 1. https://doi.org/10.1109/tkde.2021.3088844 16.

[28] Öztemiz, F., & Karcı, A. (2023). KO: Modularity optimization in community detection. Neural Computing & Applications, 35(15), 11073–11087. https://doi.org/10.1007/s00521-023-08284-8

[29] Chaudhary, L., & Singh, B. (2019). Community Detection Using Maximizing Modularity and Similarity Measures in Social Networks. In Smart innovation, systems and technologies (pp. 197–206). https://doi.org/10.1007/978-981-13-8406-6_20

[30] Magelinski, T., Bartulovic, M., & Carley, K. M. (2021). Measuring Node Contribution to Community Structure With Modularity Vitality. IEEE Transactions on Network Science and Engineering, 8(1), 707–723. https://doi.org/10.1109/tnse.2020.3049068