Yugoslav Journal of Operations Research
31 (2021), Number 1, 65–74
DOI: https://doi.org/10.2298/YJOR191215031D

SOME PROPERTIES OF E-QUALITY FUNCTION FOR NETWORK CLUSTERING

Dušan DŽAMIĆ
Faculty of Organizational Sciences, University of Belgrade, Serbia
dzamic@fon.bg.ac.rs

Received: December 2019 / Accepted: October 2020

Abstract: One of the most important properties of graphs that represents real complex systems is community structure, or clustering, i.e., organizing vertices in cohesive groups with high concentration of edges within individual groups and low concentration of edges between vertices in different groups. In this paper, we analyze Exponential Quality function for network clustering. We consider different classes of artificial networks from literature and analyze whether the maximization of Exponential Quality function tends to merge or split clusters in optimal partition even if they are unambiguously defined. Our theoretical results show that Exponential Quality function detects the expected and reasonable clusters in all classes of instances and the Modularity function does not.

Keywords: Clustering, Equality Function, Complex Networks.

MSC: 90B85, 90C26.

1. INTRODUCTION

A complex system is the system composed of numerous elements with nonlinear interactions. Examples of some complex systems are different infrastructures such as energy network, transportation or communication systems, social and economic organizations, ecosystems, living cells, human brains, and even the entire universe. If neglecting the specificity of the components, the structure of a complex system can be represented by a network (graph) such that system elements are mapped to vertices in the network, and interactions between elements are represented as edges between vertices. The structure of such networks is neither regular and nor completely random but is hierarchical arranged like community structure, high clustering coefficient, or has some other characteristic features. One of the features is a high concentration of edges within individual groups of vertices and,

at the same time, low concentration of edges between vertices in different groups. Such groups of vertices are called clusters (modules, communities) and very often have common features and roles in a complex system. The problem of finding such groups in complex networks is called community detection, or clustering on networks. For example, clusters on the World Wide Web represent pages with similar themes, while in protein networks, clusters correspond to proteins with similar functions in a cell. The development of methods for solving this problem and their applications are of great importance for understanding the dynamics and evolution of complex systems. In addition, they can provide better visualization and necessary information about individual vertices and their roles in a network. For example, some vertices in a cluster may play a role in connecting the cluster to the rest of the network, and others in controlling and stabilizing the cluster.

It is important to emphasize that there is no strict definition of the cluster but there are multiple approaches to formalize it. The most common approach for clustering on complex networks is to define so called *quality function*, a function that measures the quality of partitioning a network, and to construct methods for finding optimal partition with respect to the defined function. In this approach, the problem of clustering is reduced to the problem of combinatorial optimization and a wide range of existing optimization methods are available. The most popular quality measure for network clustering is the modularity defined in 2004 by Newman and Girvan [11]. The modularity of cluster is difference between fraction of edges within the cluster and fraction of expected number of edges within the cluster in a network with vertices of the same degree and randomly positioned edges. Network clustering by modularity maximization within an arbitrary network is NP-hard [3]. There are many heuristic methods proposed to solve it, such as Greedy Randomized Adaptive Search Procedure [10], Memetic Algorithm [1], Ascent-Descent Variable Neighborhood Decomposition Search [5] and others [2, 8]. Fortunato et al. [7] found, through several examples on artificial networks, that optimizing modularity in large networks can fail to resolve small clusters even in cases where clusters are unambiguously defined. In the literature, this problem is broadly addressed as the *resolution limit problem*, and it initiated research for a new quality function that focuses on a local definition of claster rather than definitions relying on a global null model [4, 9, 6].

In this paper, we focus on theoretical analysis of Exponential Quality function (E-quality) for network clustering, proposed in [6]. We considered different classes of artificial networks introduced in [7] and analyzed whether E-quality function tends to merge or split clusters that are unambiguously defined. Our theoretical results confirmed experimental results presented in [6], and showed that E-quality function detects expected and reasonable clusters, which is not the case with the Modularity function.

2. ANALYSIS OF E-QUALITY FUNCTION

2.1. Notations

In this subsection we briefly present the necessary notation introduced in [6]. Let $G = (V, E)$ be a simple graph defined by a set V of n vertices (nodes) and a set E of m edges that connect pairs of vertices. Let C be a subset of a vertex set V of a graph G. Then a subgraph $G_C = (C, E_C)$ is the vertex-induced subgraph of G on C if E_C is a subset of E such that edge (v_i, v_j) is in E_C, if and only if v_i and v_j are in C. The number of vertices from C and the number of edges from E_C are also called the number of vertices, and the number of edges inside the community C, and we denote them with n_C and m_C, respectively. The density of a graph G, is $D_G = \frac{2m}{n(n-1)}$ and represents the ratio of its number of edges and the maximum possible number of edges. Similarly, the density of a cluster C is $D_C = \frac{2m_C}{n_C(n_C-1)}$ and represents the density of vertex-induced subgraph G_C.

Let S and T be disjoint (nonempty) subsets of V for a graph $G = (V, E)$. Then, $\text{cut}(S, T)$ denote the set of edges (v_i, v_j) such that an endpoint $v_i \in S$, and an endpoint $v_j \in T$. Cut size of a vertex set C of a graph $G = (V, E)$, also called the number of external edges for the community C, is the number of edges from $\text{cut}(C, V \backslash C)$, and we denote it with l_C. Partition $\mathcal{P}$ of graph G represents disjoin subsets of vertices $C_1, C_2, \ldots, C_k$, called clusters, such that $\bigcup_{i=1}^{k} C_i = V$, and $C_i \cap C_j = \emptyset$, $\forall i, j,\ i \neq j$. Set of all possible partitions of graph G we denote with $\mathcal{P}_G$. For a given partition $\mathcal{P}$ of a network the modularity quality function MQ is defined as $MQ(\mathcal{P}) = \sum_{C \in \mathcal{P}} \left[\frac{m_C}{m} - \left(\frac{K_C}{2m} \right)^2 \right]$, where m_C is the number of edges inside the community C, and K_C is the total degree of vertices in the community C.

2.2. Exponential Quality function

Using the above notation, E-quality of a partition can be presented as follows:

$$EQ(\mathcal{P}) = \sum_{C \in \mathcal{P}} \left[e^{n_C (D_C - D_G)} - e^{\frac{2l_C}{n_C}} \right]$$

$$= \sum_{C \in \mathcal{P}} \left[e^{n_C \left(\frac{2m_C}{n_C(n_C-1)} - \frac{2m}{n \cdot (n-1)} \right)} - e^{\frac{2l_C}{n_C}} \right].$$

The first term in the sum refers to cluster benefit, based on its density and number of vertices. The second term refers to cluster separation or cost, based on its number of external edges and number of vertices. Network clustering problem can be defined as the optimization problem of E-quality function over the possible partitions of a graph:

$$\max_{\mathcal{P} \in \mathcal{P}_G} EQ(\mathcal{P}).$$

Below we show that maximizing E-quality function in a different class of graphs will not produce unwanted splits or merges of clearly defined clusters in the optimal partition. First, we prove the following lemma.

Lemma 1. *Let $x \geq 3$ and $0 < y \leq \frac{1}{2}$. Then the following inequality holds:*

$$e^{x(1-y)} - e^{\frac{4}{x}} > \frac{1}{2}\left(e^{x(1-2y)} - e^{\frac{2}{x}}\right).$$

Proof. Let $f(x,y)$ be a function with domain $D_f = [3,+\infty) \times \left(0, \frac{1}{2}\right]$ given by

$$f(x,y) = e^{x(1-y)} - \frac{1}{2}e^{x(1-2y)} - e^{\frac{4}{x}} + \frac{1}{2}e^{\frac{2}{x}}.$$

The partial derivatives of $f(x,y)$ with respect to x is

$$f'_x(x,y) = (1-y)\,e^{x(1-y)} - \left(\frac{1}{2} - y\right)e^{x(1-2y)} + \frac{4}{x^2}e^{\frac{4}{x}} - \frac{1}{x^2}e^{\frac{2}{x}},$$

and respect to y is

$$f'_y(x,y) = -xe^{x(1-2y)}(e^{xy} - 1).$$

Since

$$(1-y)\,e^{x(1-y)} > \left(\frac{1}{2} - y\right)e^{x(1-y)} \geq \left(\frac{1}{2} - y\right)e^{x(1-2y)}, \quad \text{for all } (x,y) \in D_f,$$

and

$$\frac{4}{x^2}e^{\frac{4}{x}} > \frac{1}{x^2}e^{\frac{4}{x}} > \frac{1}{x^2}e^{\frac{2}{x}}, \quad \text{for all } (x,y) \in D_f,$$

we have that $f'_x(x,y) > 0$, which implies that for fixed y, f is strictly increasing on $x \in [3,+\infty)$. So $f(x,y) \geq f(3,y)$ for all $(x,y) \in D_f$. On the other hand, it is clear that $f'_y(x,y) < 0$, which implies that for fixed x, f is strictly decreasing on $y \in \left(0, \frac{1}{2}\right]$. So $f(3,y) \geq f\left(3, \frac{1}{2}\right)$ for all $y \in \left(0, \frac{1}{2}\right]$ Thereby,

$$f(x,y) \geq f\left(3, \frac{1}{2}\right) = \frac{1}{2}\left(e^{\frac{2}{3}} - 2e^{\frac{4}{3}} + 2e^{\frac{3}{2}} - 1\right) \approx 1.16 > 0, \quad \forall (x,y) \in D_f$$

Since $f(x,y) > 0$ on domain D_f, we have that
$e^{x(1-y)} - e^{\frac{4}{x}} > \frac{1}{2}\left(e^{x(1-2y)} - e^{\frac{2}{x}}\right), \quad$ for $x \geq 3$ and $0 < y \leq \frac{1}{2}$. □

Theorem 2. *Maximizing E-quality function does not divide a complete graph with n vertices ($n \geq 4$) into k clusters $\left(2 \leq k \leq \left\lfloor \frac{n}{2} \right\rfloor\right)$.*

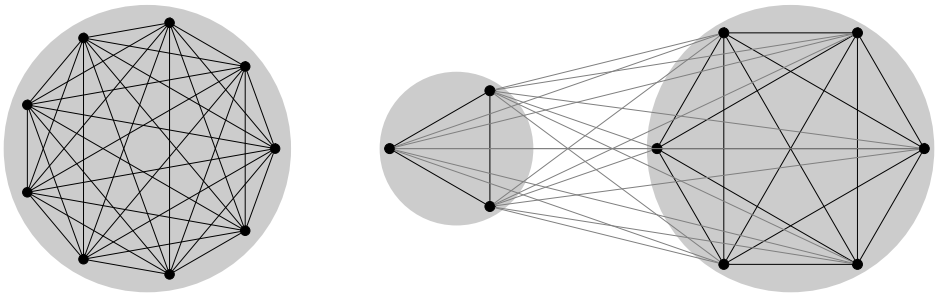
Proof. Let $G = (V,E)$ be a complete graph with n ($n \geq 4$) vertices and $m = \frac{n(n-1)}{2}$ edges. Let $\mathcal{P}_1^n$ be a partition formed by a single cluster C_1^1 containing all n nodes of the graph. Then

$$EQ(\mathcal{P}_1^n) = e^{n\left(\frac{2\frac{n(n-1)}{2}}{n(n-1)} - \frac{2\frac{n(n-1)}{2}}{n(n-1)}\right)} - e^{\frac{2\cdot 0}{n}} = 1 - 1 = 0.$$

Let $\mathcal{P}_2^{n_1,n_2}$ be a partition that divides the graph into two clusters C_1^2 and C_2^2 with n_1 and n_2 vertices, respectively. Then, the number of edges between clusters C_1^2 and C_2^2 is $n_1 \cdot n_2$, and the quality of partition $\mathcal{P}_2^{n_1,n_2}$ is

$$\begin{aligned} EQ(\mathcal{P}_2^{n_1,n_2}) &= e^{n_1\left(\frac{2\frac{n_1(n_1-1)}{2}}{n_1(n_1-1)}-\frac{2\frac{n(n-1)}{2}}{n(n-1)}\right)} - e^{\frac{2n_1n_2}{n_1}} + e^{n_2\left(\frac{2\frac{n_2(n_2-1)}{2}}{n_2(n_2-1)}-\frac{2\frac{n(n-1)}{2}}{n(n-1)}\right)} - e^{\frac{2n_1n_2}{n_2}} \\ &= 1 - e^{2n_1} + 1 - e^{2n_2} \\ &= 2 - (e^{2n_1} + e^{2n_2}). \end{aligned} \tag{1}$$

Example of two partitions in complete graph with 9 vertices is given in Figure 1.



(a) Partition $\mathcal{P}_1^9$ (b) Partition $\mathcal{P}_2^{3,6}$

Figure 1: An example of complete graph, $n = 9$

Since $\min\{e^{2n_1} + e^{2n_2} \,|\, n_1 + n_2 = 4\,, n_1 \geq 2, n_2 \geq 2\} = 2e^4$ at $(n_1, n_2) = (2, 2)$ we have that

$$EQ(\mathcal{P}_2^{n_1,n_2}) = 2 - (e^{2n_1} + e^{2n_2}) < 2 - 2e^4 < EQ(\mathcal{P}_1^n)$$

which implies that maximizing E-quality function does not divide G into two clusters.

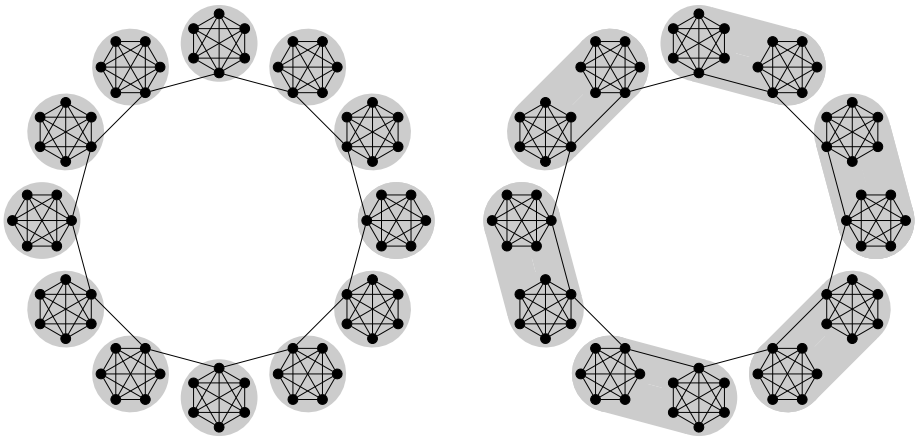
As each cluster (induced subgraph) of a complete graph is also complete, each division of a complete graph G into k clusters is obtained by successive divisions into two clusters. Therefore, it is clear that through iterations, the quality of the partitions will decrease. Thus,

$$EQ(\mathcal{P}_k^{n_1,\ldots,n_k}) < EQ(\mathcal{P}_1^n), \quad \text{for all} \quad 2 \leq k \leq \left\lfloor \frac{n}{2} \right\rfloor,$$

where $\mathcal{P}_k^{n_1,\ldots,n_k}$ is a partition that divides the graph into k clusters $C_1^k, \ldots, C_k^k$, with $n_1, \ldots n_k$ vertices respectively, $n_i \geq 2$, $(i = 1, \ldots k)$ and $\sum_{i=1}^k n_i = n$. Hence, maximization of E-quality function will not divide complete graph into k clusters. □

Theorem 3. *Maximizing E-quality function does not merge two cliques in a clique structure ring graph, with l cliques ($l \geq 3$), each contains s vertices ($s \geq 3$), and two consecutive cliques are connected by a single edge.*

Proof. Let $G = (V, E)$ be a ring network with l cliques ($l \geq 3$), each with s vertices ($s \geq 3$). The total number of vertices is $n = sl$. As each two consecutive cliques are connected by a single edge, the total number of edges is $\frac{1}{2}sl(s-1) + l$. Let $\mathcal{P}_1$ be a partition that divides the graph into l clusters, each corresponding to a single clique. Without loss of generality, suppose l is an even number and let $\mathcal{P}_2$ be partition that divides the graph into $l/2$ clusters, each corresponding to two consecutive cliques. Example of $\mathcal{P}_1$ and $\mathcal{P}_2$ partitions in clique structure ring graph with 12 cliques is given in Figure 2. We prove that quality of partition $\mathcal{P}_1$ is greater than quality of partition $\mathcal{P}_2$ in graph G, regarding EQ function.



(a) Partition $\mathcal{P}_1$ (b) Partition $\mathcal{P}_2$

Figure 2: An example of clique structure ring graph, $l = 12$, $s = 6$.

First we consider density of graph G, and density of clusters in partitions $\mathcal{P}_1$ and $\mathcal{P}_2$. The density of graph G is $D_G = \frac{s^2-s+2}{s^2l-s}$. We prove that

$$0 < D_G \leq \frac{1}{3}.$$

Let $l \geq 3$ be fixed and consider the density of graph G as function $d : [3, +\infty) \longrightarrow (0, 1]$. Consider derivative of this function $d'(s) = \frac{s^2(l-1)-4ls+2}{s^2(sl-1)^2}$. As $l - 1 > 0$, it is clear that $d'(s) < 0$ for $s \in (s_1, s_2)$ and $d'(s) > 0$ for $s \in (-\infty, s_1) \cup (s_2, +\infty)$, where s_1 and s_2 are solutions of quadratic equation $(l-1)s^2-4ls+2 = 0$. Moreover, detailed analysis of this quadratic function shows that if $l \geq 3$ then $s_1 \in (0, 3-2\sqrt{2}]$ and $s_2 \in (4, 3 + 2\sqrt{2}]$. So function $d(s)$ is a monotonically decreasing function on $(3, s_2)$, and therefore we have

$$d(s) \leq d(3) = \frac{8}{9l-3} \leq \frac{1}{l}, \quad \text{for all } l \geq 3.$$

On the other hand, function $d(s)$ is a monotonically increasing on $(s_2, +\infty)$, and therefore we have

$$d(s) < \lim_{s\to\infty} d(s) = \lim_{s\to\infty} \frac{s^2 - s + 2}{s^2 l - s} = \frac{1}{l}$$

Thereby,

$$0 < d(s) \leq \frac{1}{l}, \quad \text{for all } \ l \geq 3,$$

from which it follows that

$$0 < D_G \leq \frac{1}{3}.$$

Since all clusters C_i^1 $(i = 1, \ldots, l)$ in partition $\mathcal{P}_1$ are complete subgraphs, we have that

$$D_{C_i^1} = 1, \quad \text{for all } i = 1, \ldots, l.$$

On the other hand, the density for all clusters C_j^2 $(j = 1, \ldots, l/2)$ in partition $\mathcal{P}_2$ is equal to $D_{C_j^2} = \frac{s^2-s+1}{2s^2-s}$, and it is easy to prove that

$$0 < D_{C_j^2} < \frac{1}{2}, \quad \text{for all } j = 1, \ldots, l/2.$$

Using the above notation, E-quality of a partition can be presented as follows:

$$EQ(\mathcal{P}_1) = \sum_{i=1}^{l} \left[e^{s\left(D_{C_i^1} - D_G\right)} - e^{\frac{4}{s}} \right] = l \left(e^{s(1-D_G)} - e^{\frac{4}{s}} \right).$$

$$EQ(\mathcal{P}_2) = \sum_{j=1}^{l/2} \left[e^{2s\left(D_{C_j^2} - D_G\right)} - e^{\frac{2}{s}} \right] \leq \frac{l}{2} \left(e^{2s(\frac{1}{2}-D_G)} - e^{\frac{2}{s}} \right)$$

Applying Lemma 1 for $x = s$ and $y = D_G$, we have

$$\left(e^{s(1-D_G)} - e^{\frac{4}{s}} \right) > \frac{1}{2} \left(e^{2s(\frac{1}{2}-D_G)} - e^{\frac{2}{s}} \right).$$

Hence,

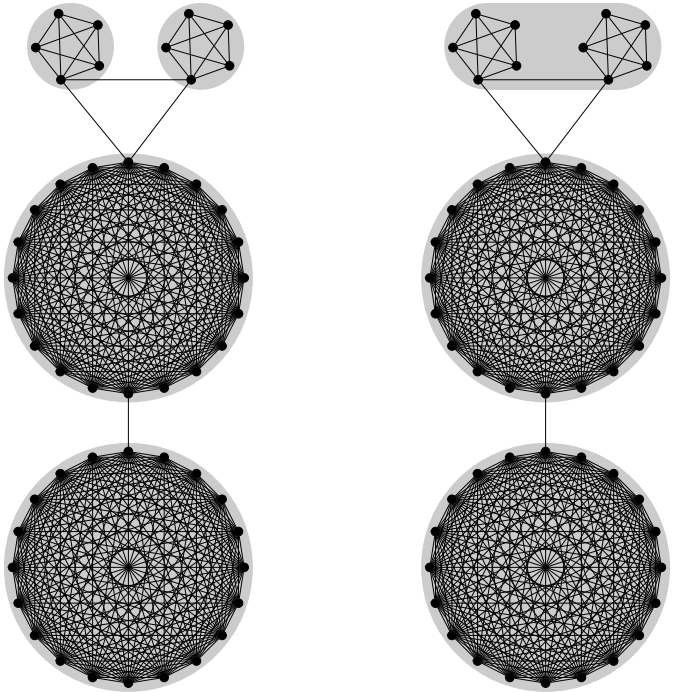
$$EQ(\mathcal{P}_1) = l \left(e^{s(1-D_G)} - e^{\frac{4}{s}} \right) \geq \frac{l}{2} \left(e^{2s(\frac{1}{2}-D_G)} - e^{\frac{2}{s}} \right) \geq EQ(\mathcal{P}_2).$$

We prove that maximization of E-quality function does not merge two cliques in a clique structure ring graph. □

Above discussion can be easily generalized in order to show that quality of partition $\mathcal{P}_1$ is greater than quality of partition $\mathcal{P}_k$, regarding E-quality function, where $\mathcal{P}_k$ is the partition that divides the graph into l/k clusters, each corresponding to k consecutive cliques.

Theorem 4. *Maximizing E-quality function could discover clusters of different sizes in a graph with two pairs of identical cliques, with q vertices ($q \geq 4$) and s vertices ($3 \leq s < q$), where cliques in the pair are connected by single edge and both s-cliques are connected with same q-clique.*

Proof. Let $G = (V, E)$ be a graph with two pairs of identical cliques with q vertices ($q \geq 4$) and s vertices ($3 \leq s < q$). Total number of vertices is $n = 2q + 2s$. As cliques in the pair are connected by a single edge and both s-cliques are connected with the q-clique, total number of edges is $m = q(q-1) + s(s-1) + 4$. Let $\mathcal{P}_1$ be a partition that divides the graph into four clusters, each corresponds to a single clique. Let $\mathcal{P}_2$ be a partition that divides the graph into three clusters, two corresponds to q-cliques and one to union of s-cliques. The example of $\mathcal{P}_1$ and $\mathcal{P}_2$ partitions in a graph with two pairs of identical cliques with 5 and 20 vertices is given in Figure 3. We prove that E-quality of partition $\mathcal{P}_1$ with four clusters, each corresponds to a single clique, is greater than the quality of partition $\mathcal{P}_2$ with three clusters, two corresponds to q-cliques and one to the union of s-cliques.



(a) Partition $\mathcal{P}_1$ (b) Partition $\mathcal{P}_2$

Figure 3: An example of graph with two pairs of identical cliques, $q = 20$, $s = 5$.

First, we consider density of graph G and density of clusters in partitions $\mathcal{P}_1$ and $\mathcal{P}_2$. The density of graph G is $D_G = \frac{q^2-q+s^2-s+4}{2q^2+4qs-q+2s^2-s}$. It is clear that $D_G > 0$ and

let us suppose that $D_G < \frac{1}{2}$. We have

$$\frac{q^2 - q + s^2 - s + 4}{2q^2 + 4qs - q + 2s^2 - s} < \frac{1}{2} \quad \Leftrightarrow \quad \frac{qs + \frac{q}{4} + \frac{s}{4} - 2}{(q + s - \frac{1}{2})(q + s)} > 0$$

which is true for all $q \geq 4$ and $3 \leq s < q$.
Hence,

$$0 < D_G \leq \frac{1}{2}.$$

Since all clusters in partition $\mathcal{P}_1$ and two clusters in partition $\mathcal{P}_2$ are complete subgraphs, we have that $D_{C_i^1} = 1$, for $i = 1, 2, 3, 4$ and $D_{C_j^2} = 1$, for $j = 1, 2$. Only density of the third cluster in partition $\mathcal{P}_2$ is $D_{C_3^2} = \frac{s^2 - s + 1}{2s^2 - s}$, and it is easy to prove that $0 < D_{C_3^2} < \frac{1}{2}$.
E-quality of partitions $\mathcal{P}_1$ and $\mathcal{P}_2$ can be presented as follows:

$$EQ(\mathcal{P}_1) = 2e^{q(1-D_G)} - e^{\frac{2}{q}} - e^{\frac{6}{q}} + 2e^{s(1-D_G)} - 2e^{\frac{4}{s}},$$

$$EQ(\mathcal{P}_2) = 2e^{q(1-D_G)} - e^{\frac{2}{q}} - e^{\frac{6}{q}} + e^{2s\left(D_{C_3^2} - D_G\right)} - e^{\frac{2}{s}}.$$

Let us consider the difference $EQ(\mathcal{P}_1) - EQ(\mathcal{P}_2)$. As $0 < D_{C_3^2} < \frac{1}{2}$, we have that

$$EQ(\mathcal{P}_1) - EQ(\mathcal{P}_2) \geq 2\left(e^{s(1-D_G)} - e^{\frac{4}{s}} - \frac{1}{2}\left(e^{2s(1-D_G)} - e^{\frac{2}{s}}\right)\right).$$

Applying Lemma 1 for $x = s$ and $y = D_G$, we have that

$$EQ(\mathcal{P}_1) > EQ(\mathcal{P}_2).$$

Hence, maximization of E-quality function could discover clusters of different sizes in graph G. □

3. CONCLUSION

In this paper, we analyzed Exponential Quality function for network clustering. We considered different classes of artificial networks from literature and analyzed whether the maximization of E-quality function tends to merge or split clusters in optimal partition even if they are unambiguously defined. Our theoretical results verified experimental results presented in [6], and showed that E-quality function detects the expected and reasonable clusters but the Modularity function does not.

Acknowledgement: This work is supported by Serbian Ministry of Education and Science under the grants No. 174010.

REFERENCES

[1] Biedermann, S., Henzinger, M., Schulz, C., and Schuster, B., "Memetic Graph Clustering", in: 17th International Symposium on Experimental Algorithms (SEA 2018), Gran Sasso Science Institute (GSSI) in L'Aquila (Italy), Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 103 (3) (2018) 1–15.

[2] Blondel, V. D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E., "Fast unfolding of communities in large networks", *Journal of statistical mechanics: theory and experiment*, 10 (2008).

[3] Brandes, U., Delling, D., Gaertler, M., Gorke, R., Hoefer, M., Nikoloski, Z., and Wagner, D. "On modularity clustering", *IEEE transactions on knowledge and data engineering*, 20 (2) (2007) 172–188.

[4] Chen, M., Nguyen, T., and Szymanski, B. K., "A new metric for quality of network community structure", *arXiv preprint arXiv:1507.04308* (2015).

[5] Džamić, D., Aloise, D., and Mladenović, N., "Ascent–descent variable neighborhood decomposition search for community detection by modularity maximization", *Annals of Operations Research*, 272 (1-2) (2019) 273–287.

[6] Džamić, D., Pei, J., Marić, M., Mladenović, N., and Pardalos, P. M. "Exponential quality function for community detection in complex networks", *International Transactions in Operational Research*, 27 (5) (2018) 245-266.

[7] Fortunato, S. and Barthelemy, M., "Resolution limit in community detection", *Proceedings of the national academy of sciences*, 104 (1) (2007) 36–41.

[8] Liu, X. and Murata, T., "Advanced modularity-specialized label propagation algorithm for detecting communities in networks", *Physica A: Statistical Mechanics and its Applications*, 389 (7) (2010) 1493–1500.

[9] Miyauchi, A. and Kawase, Y., "Z-score-based modularity for community detection in networks", *PloS one*, 11 (1) (2016) e0147805.

[10] Nascimento, M. C. and Pitsoulis, L., "Community detection by modularity maximization using grasp with path relinking", *Computers & Operations Research*, 40 (12) (2013) 3121–3131.

[11] Newman, M. E. and Girvan, M., "Finding and evaluating community structure in networks", *Physical review E*, 69 (2) (2004) 026113.