Mathematical Modelling of Engineering Problems
Vol. 11, No. 5, May, 2024, pp. 1383-1391
Journal homepage: http://iieta.org/journals/mmep

Comparing Energy in ζ-Labeling Similarity Measure with Alternative Similarity Metrics on Rough Graphs

R. Nithya, K. Anitha*

Department of Mathematics, SRM Institute of Science and Technology, Chennai 600089, India

Corresponding Author Email: anithak1@srmist.edu.in

Copyright: ©2024 The authors. This article is published by IIETA and is licensed under the CC BY 4.0 license (http://creativecommons.org/licenses/by/4.0/).

https://doi.org/10.18280/mmep.110530

Received: 31 October 2023
Revised: 22 January 2024
Accepted: 30 January 2024
Available online: 30 May 2024

Keywords:
rough membership value, rough graph, ζ-labeling similarity measure

ABSTRACT

This paper explores the effectiveness of different similarity measures for characterizing vertices and edges in rough graphs, which were introduced to handle imprecise and uncertain information. The authors examine traditional similarity measures like the Jaccard index, Dice coefficient, and overlap measure in this context. Additionally, a new ζ-labeling similarity measure for rough graphs is proposed. The main goal is to perform a comparative analysis evaluating the performance of these diverse similarity measures when applied to rough graphs. Furthermore, the paper computes the energy of rough graphs, defined as the sum of absolute eigenvalues, to demonstrate the superior potency of the proposed ζ-labeling measure compared to the other similarity measures considered. Overall, this work aims to advance techniques for assessing similarity in rough graphs, which have applications in dealing with vague and imprecise data.

1. INTRODUCTION

Rough set theory, a prominent mathematical framework developed by Professor Pawlak to address uncertain problems, was extended to the domain of graphs through the work of Tong in 2006 [1], which leveraged approximations and led to varied forms of representations of rough graphs, including weighted rough graphs and directed rough graphs [2, 3]. Mathew et al. [4] established the notion of a vertex rough graph, delving into precision at both vertex and edge levels. Their investigation involved comparing two rough graphs using the degree of Similarity Measure. In this paper, we propose labeling through similarity measures for the graph from an Information system.

For managing imprecise data, Rough sets address boundary cases, while Fuzzy sets handle graded membership. These sets have found widespread application by researchers in both academic studies and practical, real-world situations [5-10]. The realm of classical and fuzzy graph labeling is discussed [7-12], encompasses diverse variations. While classical graph labeling techniques have been applied in areas like network analysis, data compression, optimization, image processing, and cryptography, the labeling of rough and fuzzy graphs caters to data involving partial truths and uncertain knowledge bases. Many researchers have investigated the concept of rough graphs through the lens of approximation methods. Anitha and Arunadevi for instance, devised a rough graph by assigning fixed rough membership values to objects within an Information System. Their work included the computation of the metric dimension of the rough graph [13].

After extensive research on rough graphs, Anitha and Nithya pioneered and explored a labeling technique, as labeling has emerged as a rapidly growing area of study across various fields. Their novel approach introduced the concept of ζ-graceful labeling applied to different representations of rough graphs [14].

In graph theory, graph energy represents a significant concept that captures the structural characteristics of a graph in a numerical form. It is defined as the trace of the graph's adjacency matrix. This notion of graph energy was originally proposed by Ivan Gutman, who demonstrated its applicability to certain families of graphs [15-18]. Extending this line of inquiry, Nagarani et al. explored the notion of energy within the context of fuzzy labeling graphs [19]. Alexander et al. made noteworthy contributions by addressing four conjectures related to path energy in graphs. They also devised an efficient algorithm for computing the path matrix [20], while Pirzada and Ganie introduced the Laplacian matrix derived from the adjacency matrix [21]. In this study, we further advance the understanding of energy within the domain of ζ-labeling rough graphs, underpinned by similarity measures.

Pappis and Karacapilidis [22] put forth the subsequent trio of similarity measures applicable to fuzzy sets as,

$$M_{A,B} = \frac{\sum_i \min(a_i, b_i)}{\sum_i max(a_i, b_i)}$$

$$L_{A,B} = 1 - max_i(|a_i - b_i|)$$

$$S_{A,B} = 1 - \frac{\sum_i \min(a_i, b_i)}{\sum_i max(a_i, b_i)}$$

Zadeh pioneered the concept of using similarity measures for fuzzy sets, which proved to be a successful strategy in handling uncertainty [23]. Simultaneously, similarity indexes

emerged as a tool to gauge approximate equality among fuzzy sets within a specific universe of discourse. Following this, Wang [24] presented an overview of fuzzy set similarity measures and introduced two novel measures to quantify similarity between fuzzy sets and individual elements. Beyond Zadeh's groundwork in fuzzy set similarity measures, various researchers extended these notions to encompass multiple sets and vague sets [25-30]. In a parallel vein, other researchers ventured into the domain of similarity measures, tackling applications like text comparison using notions such as soft cardinality, similarity-based ranking, and query processing in multimedia databases and text mining. These authors proposed a novel similarity measure for fuzzy graphs, which seamlessly extends to the realm of fuzzy signatures, finding utility in analyzing workforce behavioral data. Furthermore, they generalized a similarity measure from trapezoidal fuzzy numbers to interval-valued trapezoidal fuzzy numbers, ensuring the preservation of its original properties [31, 32].

Extending previous similarity measures, we introduced a new similarity metric tailored to a novel labeling approach, which we explored for various representations of rough graphs constructed from high-dimensional data.

The primary aim of this paper is to propose a novel ζ-labeling similarity measure and evaluate its performance in comparison with existing similarity measures in the context of rough graphs. Additionally, this paper highlights the application of ζ-labeling similarity measures to rough graphs, with a focus on analyzing their energy. The synergy of rough ζ-labeling and the modified similarity measure yields the rough ζ-labeling similarity measure.

The paper commences with Section 2, which lays out the fundamental concepts of rough sets and rough graphs. Subsequently, Section 3 explores the notion of similarity relations in depth. Section 4 elucidates the methodology for labeling vertices and edges using similarity measures. Finally, the same section examines the connection between the energies of similarity measures, culminating in the concluding remarks presented in Section 5.

1.1 Exploring similarity measures

Similarity measures are mathematical techniques used to quantify the degree of similarity or dissimilarity between two objects, entities, or data points. They play a crucial role in various fields, including data analysis, pattern recognition, machine learning, and information retrieval. Similarity measures are used to compare objects based on their features or attributes and determine how closely they resemble each other.

There are several types of similarity measures and they can be broadly categorized into the following:

(a) Jaccard similarity coefficient

The Jaccard coefficient, also referred to as the Jaccard similarity coefficient or Jaccard index, is a metric used to quantify the similarity between two sets. It is defined as the size of the intersection of the sets divided by the size of their union. Mathematically, the Jaccard coefficient (J) is calculated using the following formula:

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|}$$

where, A and B are two sets to measure similarity. The Jaccard coefficient produces a value between 0 and 1. A value of 0 signifies that there is no similarity or common elements between the sets, while a value of 1 signifies complete similarity, indicating that the sets are identical.

(b) Dice similarity measure

The Dice coefficient is an alternative method for gauging the similarity between two sets. Its calculation employs the subsequent formula:

$$Dice\ (A,B) = \frac{2|A \cap B|}{|A| + |B|}$$

The Dice coefficient furnishes an output ranging between 0 and 1. A value of 0 denotes the absence of overlap or similarity amid the sets, while a value of 1 signifies an impeccable overlap or total similarity. A heightened Dice coefficient implies a superior overlap or concordance between the segmented regions and the established ground truth.

(c) Overlap coefficient similarity measure

The Overlap coefficient measures the similarity between two sets by expressing the fraction of their overlap. It's alternatively referred to as the Overlap index or Overlap coefficient of Tversky. The calculation of the overlap coefficient involves the use of the following formula:

$$Overlap(A,B) = \frac{|A \cap B|}{min(|A|,|B|)}$$

The outcome of the overlap coefficient ranges between 0 and 1. A value of 0 denotes the absence of overlap or similarity between the sets, while a value of 1 signifies full overlap or complete similarity.

2. PRELIMINARIES

This section covers fundamental concepts related to Rough sets and Rough graphs.

2.1 Information system

Let $\mathcal{U}$ be a non-empty finite set, referred to as the universe of discourse, and $\mathcal{A}$ be a set of attributes. An information system I_s is defined as a pair $(\mathcal{U}, \mathcal{A})$, where for every $k \in \mathcal{A}$, there exists a function $k: \mathcal{U} \to \mathcal{V}_k$, with $\mathcal{V}_k$ being the value set of attribute a. If there exists a decision attribute $d \notin \mathcal{A}$, called the decision attribute, and the elements of A are termed condition attributes, then the triplet $(\mathcal{U}, \mathcal{A}, d)$ is called a decision system.

2.2 Rough membership function (RM_f)

The RM_f value provides a way to represent and handle the uncertainty and imprecision associated with the membership of elements in a rough set, allowing for a more flexible and realistic representation of real-world data. RM_f is characterized by $\varphi_{\mathcal{R}}: \mathfrak{T} \to [0,1]$ and defined by

$$\omega_{\mathfrak{T}}^{\mathcal{R}}(y) = \frac{|[y]_{\mathcal{R}} \cap \mathfrak{T}|}{|[y]_{\mathcal{R}}|}, \forall y \in \mathcal{U}$$

2.3 Rough graph

Let $\mathcal{U}$ be a non-empty set called the universe, and let $\mathfrak{E}$ be a set of unordered pairs of distinct elements from $\mathcal{U}$. A graph $\mathbb{G}$ can be constructed from these elements with the following

considerations:

An edge between two vertices in the graph exists if and only if the maximum of their associated membership values is greater than zero.

2.4 Energy measure of rough labeling graph

Definition 2.4.1 The matrix representing the rough labeling relation is coined as $M^{\varphi} = [m_{ij}^{\varphi}]$ where $m_{ij}^{\varphi} = \sigma^{\varphi}(v_i v_j)$.

Definition 2.4.3 The sum of the absolute values of the eigenvalues of the rough labeling matrix is known as the Energy of the rough graph $\mathcal{R}_{\mathcal{L}}^{\varphi}$ which is denoted by $\mathfrak{E}(\mathcal{R}_{\mathcal{L}}^{\varphi}) = \sum_{i=1}^{n}|\psi_i|$ and also it should satisfies the following criteria:

i) $\mathfrak{E}(\mathcal{R}_{\mathcal{L}}^{\varphi}) = \sum_{i=1}^{n}|\psi_i|$
ii) $0 \le \omega(v_i) \le 1$
iii) If $\mathcal{R}_{\mathcal{L}}^{\varphi} = \max(\omega(v_i^{\varphi}), \omega(v_j^{\varphi})) > 0$ then edge exists for $v_i, v_j \in V$.

3. MEASURING SIMILARITY: A QUANTITATIVE APPROACH

Let $\mathbb{S}$: $\mathcal{R}_{\mathcal{L}}^{\varphi}(v_i, v_j) \to [0,1]$ be a function mapping pairs of elements from the universe $\mathcal{U}$ to the closed interval [0, 1]. Then, $\mathbb{S}$ is said to be a similarity measure between v_i, v_j in $\mathcal{U}$ if $\mathbb{S}$ satisfies the following properties:

i) $0 \le Sim(v_i, v_j) \le 1$
ii) $Sim(v_i, v_j) = Sim(v_j, v_i)$
iii) For any $v_i, v_j, v_k \in \mathcal{R}_{\mathcal{L}}^{\varphi}$, $Sim(v_i, v_k) \le Sim(v_i, v_j)$ and $Sim(v_i, v_k) \le Sim(v_j, v_k)$
iv) $[v_i]_{\mathcal{S}_r} = \{v_j : v_i \mathcal{S}_r v_j\}$

3.1 Characteristics of similarity metric

Theorem 3.1.1 $S(v_i, v_j) = 1$ iff $[v_i]_{\mathcal{S}_r} = [v_j]_{\mathcal{S}_r}$

Proof: $S(v_i, v_j) = 1$ iff $S([v_i]_{\mathcal{S}_r}, [v_j]_{\mathcal{S}_r}) = 1$ which is equivalent to $[v_i]_{\mathcal{S}_r} = [v_j]_{\mathcal{S}_r}$ taking into account the property $[v_i]_{\mathcal{S}_r} \subseteq X$ and $[v_j]_{\mathcal{S}_r} \subseteq X \iff \omega_X^R(v_i) = 1$ and $\omega_X^R(v_j) = 1$. So $S(v_i, v_j) = 1$ where $v_i, v_j \in V$.

Theorem 3.1.2 $S(v_i, v_j) = 0$ iff $[v_i]_{\mathcal{S}_r} \ne [v_j]_{\mathcal{S}_r}$

Proof: $S(v_i, v_j) = 0$ iff $S([v_i]_{\mathcal{S}_r}, [v_j]_{\mathcal{S}_r}) = 0$ which is equivalent to $[v_i]_{\mathcal{S}_r} \cap X = 0$ iff $\omega_X^R(v_i) = 0$ and $[v_j]_{\mathcal{S}_r} \cap X = 0$ iff $\omega_X^R(v_j) = 0$.

Theorem 3.1.3 $0 \le S(v_i, v_j) \le 1$

Proof: For $X \subset U, [v_i]_{\mathcal{S}_r} \subseteq X$ and $[v_j]_{\mathcal{S}_r} \subseteq X$ where $v_i, v_j \in V$. Then $[v_i]_{\mathcal{S}_r} \subseteq X$ iff $[v_j]_{\mathcal{S}_r} \cap X \ne \emptyset$ iff $\omega_X^R(v_i) > 0$ and $[v_j]_{\mathcal{S}_r} \subseteq X$ iff $[v_j]_{\mathcal{S}_r} \cap U - X \ne \emptyset$ iff $\omega_X^R(v_j) < 1$. So, $0 \le \omega_X^R(v_i) \le 1$ and $0 \le \omega_X^R(v_j) \le 1$, $0 \le max(\omega_X^R(v_i), \omega_X^R(v_j)) \le 1$

Hence, we proved that $0 \le S(v_i, v_j) \le 1$.

Theorem 3.1.4 $S(v_i, v_j) = S(v_j, v_i)$

Proof: i) $S(v_i, v_j) = 1$ iff $S([v_i]_{\mathcal{S}_r}, [v_j]_{\mathcal{S}_r}) = 1$ which is equivalent to $[v_i]_{\mathcal{S}_r} = [v_j]_{\mathcal{S}_r}$ iff $\omega_X^R(v_i) = 1$ and $\omega_X^R(v_j) = 1$.

ii) $S(v_j, v_i) = 1$ iff $S([v_j]_{\mathcal{S}_r}, [v_i]_{\mathcal{S}_r}) = 1$ which is equivalent to $[v_j]_{\mathcal{S}_r} = [v_i]_{\mathcal{S}_r}$ iff $\omega_X^R(v_j) = 1$ and $\omega_X^R(v_i) = 1$.

From (i) and (ii) $\Longrightarrow S(v_i, v_j) = S(v_j, v_i)$

Theorem 3.1.5 $Sim(v_i, v_k) \le Sim(v_i, v_j)$ and $Sim(v_i, v_k) \le Sim(v_j, v_k)$

Proof: For any $v_i, v_j, v_k \in \mathcal{R}_{\mathcal{L}}^{\varphi}$,

$S(v_i, v_j) = 1$ if $S([v_i]_{\mathcal{S}_r}, [v_j]_{\mathcal{S}_r}) = 1$ which is equivalent to $[v_i]_{\mathcal{S}_r} = [v_j]_{\mathcal{S}_r}$ if $\omega_X^R(v_i) = 1$ and $\omega_X^R(v_j) = 1$

And $S(v_j, v_k) = 1 \iff S([v_j]_{\mathcal{S}_r}, [v_k]_{\mathcal{S}_r}) = 1$ which is equivalent to $[v_j]_{\mathcal{S}_r} = [v_k]_{\mathcal{S}_r}$ iff $\omega_X^R(v_j) = 1$ and $\omega_X^R(v_k) = 1$. Therefore, it is proved that $S(v_i, v_k) \le S(v_i, v_j)$. Similarly, $S(v_i, v_k) \le S(v_j, v_k)$

3.2 Operations of similarity measures

1. Let v_i is a subset of v_j (i.e) $v_i \subseteq v_j$ iff $[v_i]_{\mathcal{S}_r} \subseteq [v_j]_{\mathcal{S}_r}$ $\forall\, v_i, v_j \in V$.
2. Let v_i is equal to v_j (i.e.) $v_i = v_j$ iff $[v_i]_{\mathcal{S}_r} \subseteq [v_j]_{\mathcal{S}_r}$ and $[v_j]_{\mathcal{S}_r} \subseteq [v_i]_{\mathcal{S}_r}$,(i.e) iff $[v_i]_{\mathcal{S}_r} = [v_j]_{\mathcal{S}_r}$ $\forall\, v_i, v_j \in V$.
3. The union of v_i *and* v_j is denoted by $\mathcal{S}_r(v_i \cup v_j)$ whose similarity classes are defined as $\mathcal{S}_r(v_i \cup v_j) = [v_j]_{\mathcal{S}_r} \vee [v_i]_{\mathcal{S}_r}$.
4. The intersection of v_i and v_j is denoted by $\mathcal{S}_r(v_i \cap v_j)$ whose similarity classes are defined as $\mathcal{S}_r(v_i \cup v_j) = [v_j]_{\mathcal{S}_r} \wedge [v_i]_{\mathcal{S}_r}$.
5. The complement of v_i is denoted as $\overline{v_i}$ whose similarity classes for each $v_i \in V$ where the vertices $v_i * v_j \in V$.

4. PROPOSED WORK

The combination of $\zeta -$ labeling formula and similarity measure in rough graph is termed as rough $\zeta -$ labeling similarity graph.

4.1 Rough ζ -labeling similarity graph

A rough graph $\mathcal{R}_{\mathcal{L}}^{\varphi} = (V, E, \rho^{\varphi}, \sigma^{\varphi}, \omega)$ is said to be rough ζ -Labeling Similarity Graph if V=$\{\rho^{\varphi}(v_i)\}$ for $i = 1,2, \dots, n$ and E=$\{\sigma^{\varphi}(v_i, v_j)\}$ for $i = 1,2, \dots, n$ and $\omega: V * V \to [0,1]$ is bijection such that edges and vertices can be labeled using similarity classes and measures if it satisfies the following requirements:

i. If $\mathcal{R}_{\mathcal{L}}^{\varphi} = \max(\omega(v_i^{\varphi}), \omega(v_j^{\varphi})) > 0$ then edge exists for $v_i, v_j \in V$.
ii. Vertex labeling: $\rho^{\varphi}(v_i) = \frac{[v_i]_{\mathcal{S}_r}}{n}$, where $[v_i]_{\mathcal{S}_r} = \{v_j / v_i \mathcal{S}_r v_j\}$
iii. Edge labeling: $\sigma^{\varphi}(v_i, v_j) = Sim_{\zeta}(v_i, v_j) = \frac{\zeta}{\zeta + \eta}$, where $\eta = \frac{|[v_i]_{\mathcal{S}_r}| * |[v_j]_{\mathcal{S}_r}|}{|[v_i]_{\mathcal{S}_r}| + |[v_j]_{\mathcal{S}_r}|}$ and $\zeta = \rho^{\varphi}(v_i) + \rho^{\varphi}(v_j) + m$,

where, $\rho^{\varphi}(v_i)$ represents vertex labeling of v_i , $\rho^{\varphi}(v_j)$ represents the vertex labeling of v_j, m represents the total no. of edges in rough graph.

Here ζ gives a labeling formula and η mentions modified similarity measure based on similarity class.

When considering two similarity classes of objects, u and v, the computation of similarity measures is performed using Eqs. (1)-(4).

$$Sim_{jaccard}(v_i, v_j) = Sim_j(v_i, v_j) = \frac{\left|[v_i]_{\mathcal{S}_r} \cap [v_j]_{\mathcal{S}_r}\right|}{\left|[v_i]_{\mathcal{S}_r} \cup [v_j]_{\mathcal{S}_r}\right|} \quad (1)$$

$$Sim_{dice}(v_i, v_j) = Sim_d(v_i, v_j) = \frac{2\left|[v_i]_{\mathcal{S}_r} \cap [v_j]_{\mathcal{S}_r}\right|}{\left|[v_i]_{\mathcal{S}_r}\right| + \left|[v_j]_{\mathcal{S}_r}\right|} \quad (2)$$

$$Sim_{\min-overlap}(v_i, v_j) = Sim_O(v_i, v_j) = min\left(\frac{\left|[v_i]_{\mathcal{S}_r} \cap [v_j]_{\mathcal{S}_r}\right|}{\left|[v_i]_{\mathcal{S}_r}\right|}, \frac{\left|[v_i]_{\mathcal{S}_r} \cap [v_j]_{\mathcal{S}_r}\right|}{\left|[v_j]_{\mathcal{S}_r}\right|}\right) \quad (3)$$

$$Sim_{\zeta-\text{label}}(v_i, v_j) = Sim_{\zeta}(v_i, v_j) = \frac{\zeta}{\zeta + \eta} \quad (4)$$

where, $\eta = \frac{|[v_i]_{\mathcal{S}_r}| * \left|[v_j]_{\mathcal{S}_r}\right|}{|[v_i]_{\mathcal{S}_r}| + \left|[v_j]_{\mathcal{S}_r}\right|}$ and $\zeta = \rho^{\varphi}(v_i) + \rho^{\varphi}(v_j) + m$

$\rho^{\varphi}(v_i)$ and $\rho^{\varphi}(v_j)$ represent the vertex labeling of v_i, v_j and m represents the total no. of edges in rough graph.

All these similarity measures yield values within the range of 0 to 1. A value of 0 signifies a complete mismatch between two clusters, while a value of 1 signifies that the two clusters are identical.

4.2 Algorithm for computing the energy of rough graphs

The following algorithm outlines the steps to compute the Energy measure of a rough graph by employing rough vertex and edge labeling techniques.

(i) Construct the rough graph by assigning membership values to vertices and edges using a membership function.
(ii) Compute the similarity classes from the given information table or data.
(iii) Label the vertices using a vertex labeling formula based on the similarity classes.
(iv) Assign labels to the edges using edge labeling formulas for various similarity measures.
(v) Represent the labeled rough graph with similarity measures in a diagrammatic form.
(vi) Construct the adjacency matrix for rough graph, incorporating the various similarity measures assigned to the edges.
(vii) Calculate the eigenvalues of the adjacency matrix for each similarity measure.
(viii) Compute the Energy of the rough graph for each similarity measure by summing the absolute values of the corresponding eigenvalues.

Illustrative Case 1: Machine quality data

Consider a dataset consisting of information about five machines, with attributes: operation efficiency, number of machines, machine capacity, and a decision attribute representing quality.

Table 1. Decision system of machine quality data

Machines	Operation Efficiency	No. of Machine	Machine Capacity h/day	Quality
m_1	Advanced	35	20	Good
m_2	Advanced	29	12	Good
m_3	Advanced	35	12	Good
m_4	Moderate	18	15	Poor
m_5	Low	35	20	Poor

Equivalence classes of Table 1 are as follows:

$$R\{m_1\}=\{m_1\}, R\{m_2\}=\{m_2\}, R\{m_3\}=\{m_3\}, R\{m_4\}=\{m_4\}, R\{m_5\}=\{m_5\}.$$

Assuming that the outcome evaluation decision is good, we consider the target set as X=$\{a, b\}$. Rough Membership values are

$$\omega(m_1) = \frac{|[x]_R \cap X|}{|[x]_R|} = 1;\ \omega(m_2) = 1;\ \omega(m_3) = 0$$
$$\omega(m_4) = 0;\ \omega(m_5) = 0$$

4.3 Similarity class

4.3.1 Calculation of similarity class

Based on the decision table (Table 1), we have constructed a similarity matrix that captures the relationships between objects with respect to their attributes. As there are three attributes in the dataset, each entry in Table 1 can take values from the set {0, 1, 2, 3}. The value 0 indicates that the two objects being compared have no overlapping attribute values, while a value of 3 signifies that the two objects are identical across all attributes.

From Table 2, the following similarity classes have been identified,

$$[m_1]_{\mathcal{S}_r}=\{a,b,c,e\};\ \left|[m_1]_{\mathcal{S}_r}\right| = 4; [m_2]_{\mathcal{S}_r}=\{a, b, c\};\ \left|[m_2]_{\mathcal{S}_r}\right| = 3;\ [m_3]_{\mathcal{S}_r}=\{a,b,c,e\};\ \left|[m_3]_{\mathcal{S}_r}\right| = 4;\ [m_4]_{\mathcal{S}_r}=\{d\};\ \left|[m_4]_{\mathcal{S}_r}\right| = 1;\ [m_5]_{\mathcal{S}_r}=\{a,c,e\};\ \left|[m_5]_{\mathcal{S}_r}\right| = 3$$

Table 2. Similarity class of machine quality data

$\mathcal{S}_r$	m_1	m_2	m_3	m_4	m_5
m_1	0	0.75	1	0	0.75
m_2	0	0	0.75	0	0.5
m_3	1	0.75	0	0	0
m_4	0	0	0	0	0
m_5	0.75	0.5	0	0	0

4.4 Labeling methodology

$$\rho^{\varphi}(v_i) = \frac{\left|[v_i]_{\mathcal{S}_r}\right|}{n}$$

4.4.1 Vertex labeling

Figure 1 depicts the rough graph constructed from Table 1.

$$\rho^{\varphi}(a) = 0.8,\ \rho^{\varphi}(b) = 0.6,\ \rho^{\varphi}(c) = 0.8,$$
$$\rho^{\varphi}(d) = 0.2, \rho^{\varphi}(e) = 0.6.$$

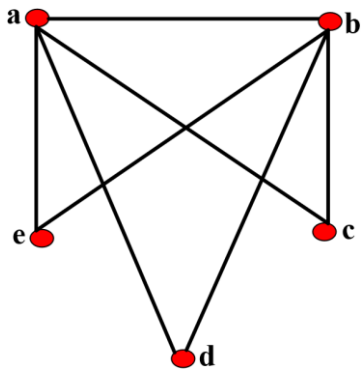

Figure 1. Rough graph of machine quality data

4.4.2 Edge labeling

$$E_G^{\varphi}(v_i, v_j) = Sim(v_i, v_j);\ Sim_{\zeta}(v_i, v_j) = \frac{\zeta}{\zeta+\eta}$$

where, $\eta = \frac{|[v_i]_{S_r}|*\left|[v_j]_{S_r}\right|}{|[v_i]_{S_r}|+\left|[v_j]_{S_r}\right|}$ and $\zeta_{ij} = \rho^{\varphi}(v_i)+\rho^{\varphi}(v_j) + m$

Table 3 presents the edge labeling values assigned to the edges of the rough graph derived from the machine quality data.

Table 3. Edge labeling of rough graph

Edges	ζ_{ij}	η	$\frac{\zeta}{\zeta+\eta}$
Sim(*a*,*b*)	$\zeta_{ab} = 8.4$	1.71	0.83
Sim(*a*,*c*)	$\zeta_{ac} = 8.6$	2	0.81
Sim(*a*,*d*)	$\zeta_{ad} = 8.0$	0.8	0.91
Sim(*a*,*e*)	$\zeta_{ae} = 8.4$	1.17	0.83
Sim(*b*,*c*)	$\zeta_{bc} = 8.4$	1.17	0.83
Sim(b,d)	$\zeta_{ab} = 7.8$	0.75	0.91
Sim(b,e)	$\zeta_{be} = 8.2$	1.5	0.85

4.5 Energy of rough graph using similarity measures

After assigning edge labels based on various similarity measures, such as Jaccard, Dice, Overlap, and ζ labeling, we construct the adjacency matrix for the rough graph. We then calculate the eigenvalues of the adjacency matrix. The Energy of the rough graph for each similarity measure is obtained by summing the absolute values of the corresponding eigenvalues.

4.5.1 Jaccard $Sim_j(v_i, v_j)$

Here edges are labeled using Jaccard Similarity Measure as $Sim_j(v_i, v_j) = \frac{\left|[v_i]_{S_r} \cap [v_j]_{S_r}\right|}{\left|[v_i]_{S_r} \cup [v_j]_{S_r}\right|}$ and the adjacency matrix is given as shown in Table 4:

Table 4. Adjacency matrix for Figure 2

S_r	m_1	m_2	m_3	m_4	m_5
m_1	3	1	2	0	2
m_2	3	3	2	0	0
m_3	2	2	3	0	1
m_4	0	0	0	3	0
m_5	2	0	1	0	3

Figure 2 illustrates the vertex, edge labeling values assigned to the rough graph.

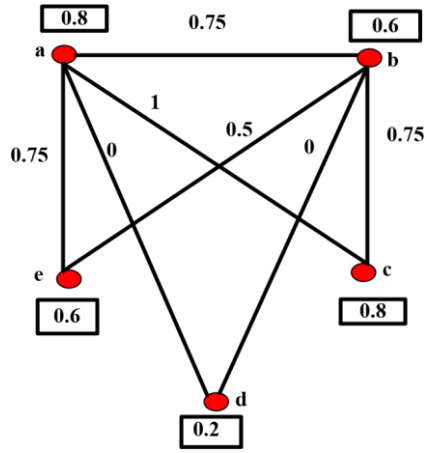

Figure 2. Jaccord similarity for Figure 1

Eigen values={0,-1.316,-0.383,0.005,1.695}
Energy of $Sim_j(v_i, v_j)$=3.399

4.5.2 Dice $Sim_d(v_i, v_j)$

Here, by using $Sim_d(v_i, v_j) = \frac{2\left|[v_i]_{S_r} \cap [v_j]_{S_r}\right|}{|[v_i]_{S_r}|+\left|[v_j]_{S_r}\right|}$, the edges are labeled and adjacency matrix is given as shown in Table 5:

Table 5. Adjacency matrix for Figure 3

S_r	m_1	m_2	m_3	m_4	m_5
m_1	0	0.85	1	0	0.85
m_2	0	0	0.85	0	0.66
m_3	1	0.85	0	0	0
m_4	0	0	0	0	0
m_5	0.85	0.66	0	0	0

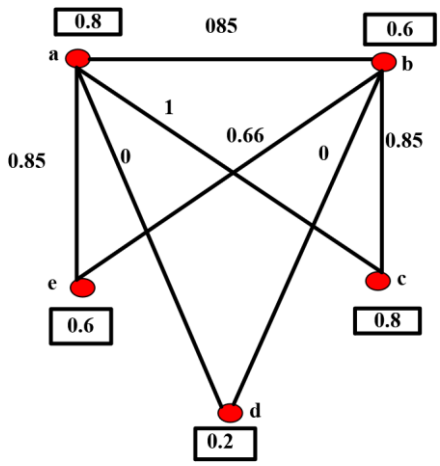

Figure 3. Labeling using Dice similarity for Figure 1

Eigen values = {0, -1.428, -0.451,0.003,1.876}
Energy = 3.758

Table 6. Adjacency matrix for Figure 4

S_r	m_1	m_2	m_3	m_4	m_5
m_1	0	0.75	1	0	0.75
m_2	0	0	0.7	0	0.6
m_3	1	0.75	0	0	0
m_4	0	0	0	0	0
m_5	0.75	0.6	0	0	0

4.5.3 Overlap $Sim_O(v_i, v_j)$

For overlap similarity measure $Sim_O(v_i, v_j) = min\left(\frac{|[v_i]_{S_r} \cap [v_j]_{S_r}|}{|[v_i]_{S_r}|}, \frac{|[v_i]_{S_r} \cap [v_j]_{S_r}|}{|[v_j]_{S_r}|}\right)$, the edges are labeled and its adjacency matrix is given in Table 6.

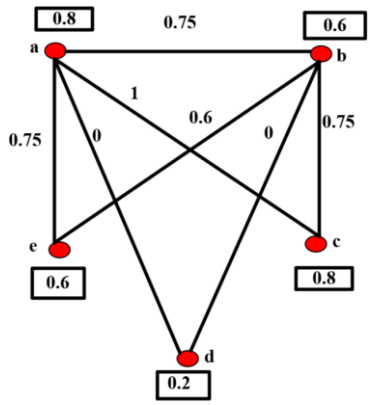

Figure 4. Overlap similarity measures for Figure 1

Eigen values: {0, -1.348, -0.387, 0.002, 1.733}
Energy: 3.47

4.5.4 Rough ζ $Sim_\zeta(v_i, v_j)$

By implementing $Sim_\zeta(v_i, v_j) = \frac{\zeta}{\zeta+\eta}$ where $\eta = \frac{|[v_i]_{S_r}| * |[v_j]_{S_r}|}{|[v_i]_{S_r}| + |[v_j]_{S_r}|}$ and $\zeta = \rho^{\varphi}(v_i) + \rho^{\varphi}(v_j) + m$ (Table 7).

Table 7. Adjacency matrix for Figure 5

S_r	m_1	m_2	m_3	m_4	m_5
m_1	0	0.83	0.81	0.9	0.83
m_2	0.83	0	0.83	1	0.85
m_3	0.81	0.83	0	0.9	0
m_4	0.91	0.91	0	0	0
m_5	0.83	0.85	0	0	0

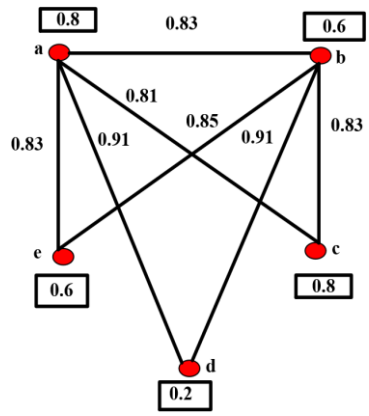

Figure 5. Overlap similarity measure for Figure 1

Eigen values: {0,-1.726,-0.830,0,2.556}
Energy: 5.112

The following Table 8 presents a comparative analysis of the Energy values for the machine quality dataset, computed using different similarity measures such as Jaccard, Dice, Overlap, and the Rough ζ Similarity Measure.

Table 8. Comparative analysis of energy of similarity measures for rough graph 1

Edges	Jaccard	Dice	Overlap	ζ
$Sim(a,b)$	0.75	0.85	0.75	0.83
$Sim(a,c)$	1	1	1	0.81
$Sim(a,d)$	0	0	0	0.91
$Sim(a,e)$	0.75	0.85	0.75	0.83
$Sim(b,c)$	0.75	0.85	0.75	0.83
$Sim(b,d)$	0	0	0	0.91
$Sim(b,e)$	0.5	0.66	0.6	0.85
Energy	3.399	3.758	3.47	5.112

Illustrative Case 2: In this case (Table 9), we have considered a dataset consisting of ten Iris flower samples, where the sepal length, sepal width, petal length, and petal width are treated as independent attributes or features. The decision attributes are the Iris species Setosa and versicolor.

Table 9. Iris flower data set

Iris Flower	Sepal Length (cm)	Sepal Width (cm)	Petal Length (cm)	Petal Width (cm)	Iris Class
f_1	5	4	3	1	Setosa
f_2	4	9	3	1	Setosa
f_3	4	8	3	4	Setosa
f_4	5	5	2	4	Versicolor
f_5	5	6	2	9	Versicolor
f_6	5	6	3	0	Versicolor
f_7	5	7	4	4	Setosa
f_8	5	8	2	7	Versicolor
f_9	6	0	2	9	Versicolor
f_{10}	5	7	3	8	Setosa

Let us consider the decision attribute as versicolor and the target set is X={f_4, f_5, f_6, f_8, f_9}.

Equivalence classes for Table 9:

$R\{f_1\}=\{f_1\}, R\{f_2\}=\{f_2\}, R\{f_3\}=\{f_3\}, R\{f_4\}=\{f_4\}, R\{f_5\}=\{f_5\}, R\{f_6\}=\{f_6\}, R\{f_7\}=\{f_7\}, R\{f_8\}=\{f_8\}, R\{f_9\}=\{f_9\}, R\{f_{10}\}=\{f_{10}\}$

Rough membership values are

$\omega(f_1) = 0; \omega(f_2) = 0; \omega(f_3) = 0; \omega(f_4) = 1; \omega(f_5) = 1; \omega(f_6) = 1; \omega(f_7) = 0; \omega(f_8) = 1; \omega(f_9) = 1; \omega(f_{10}) = 0$

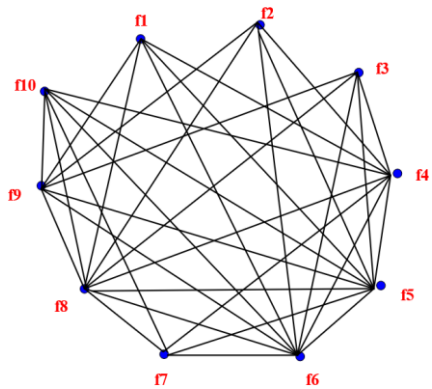

Figure 6. Rough graph of Iris flower data

Figure 6 illustrates the rough graph constructed based on the membership values assigned to each object in the dataset.

Table 10. Calculation of similarity class

S_r	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8	f_9	f_{10}
f_1	4	2	1	1	1	2	1	1	0	2
f_2	2	4	2	0	0	1	0	0	0	1
f_3	1	2	4	1	0	1	1	1	0	1
f_4	1	0	1	4	2	1	2	2	1	1
f_5	1	0	0	2	4	2	1	2	2	1
f_6	2	1	1	1	2	4	1	1	0	2
f_7	1	0	1	2	1	1	4	1	0	2
f_8	1	0	1	2	2	1	1	4	1	1
f_9	0	0	0	1	2	0	0	1	4	0
f_{10}	2	1	1	1	1	2	2	1	0	4

Table 11. Edge labels based on different similarity metrics

Edge Labeling	Jaccord	Dice	Overlap	ζ
Sim (f_1, f_4)	0.88	0.88	0.88	0.89
Sim (f_1, f_5)	0.7	0.82	0.77	0.90
Sim (f_1, f_6)	1	1	1	0.89
Sim (f_1, f_8)	0.8	0.88	0.8	0.89
Sim (f_1, f_9)	0.3	0.46	0.3	0.92
Sim (f_2, f_4)	0.4	0.57	0.44	0.92
Sim (f_2, f_5)	0.3	0.46	0.37	0.92
Sim (f_2, f_6)	0.5	0.71	0.55	0.92
Sim (f_2, f_8)	0.4	0.57	0.44	0.92
Sim (f_2, f_9)	0	0	0	0.94
Sim (f_3, f_4)	0.7	0.82	0.77	0.90
Sim (f_3, f_5)	0.6	0.75	0.75	0.90
Sim (f_3, f_6)	0.8	0.94	0.8	0.90
Sim (f_3, f_8)	0.7	0.82	0.77	0.90
Sim (f_3, f_9)	0.2	0.33	0.25	0.93
Sim (f_4, f_5)	0.88	0.94	0.8	0.90
Sim (f_4, f_6)	0.8	0.88	0.8	0.90
Sim (f_4, f_7)	0.88	0.94	0.8	0.90
Sim (f_4, f_8)	0.9	1	1	0.89
Sim (f_4, f_9)	0.44	0.61	0.44	0.92
Sim (f_4, f_{10})	0.9	1	1	0.89
Sim (f_5, f_6)	0.7	0.82	0.77	0.90
Sim (f_5, f_7)	0.77	0.87	0.87	0.90
Sim (f_5, f_8)	0.88	0.94	0.8	0.90
Sim (f_5, f_9)	0.5	0.66	0.5	0.93
Sim (f_5, f_{10})	0.7	0.82	0.77	0.90
Sim (f_6, f_7)	0.88	0.94	0.8	0.89
Sim(f_6, f_8)	0.8	0.88	0.8	0.89
Sim (f_6, f_9)	0.3	0.46	0.33	0.92
Sim (f_6, f_{10})	0.9	1	1	0.89
Sim(f_7, f_8)	0.8	0.94	0.8	0.90
Sim (f_7, f_9)	0.33	0.5	0.37	0.93
Sim (f_8, f_9)	0.44	0.61	0.44	0.92
Sim (f_8, f_{10})	0.8	0.88	0.88	0.89
Sim (f_9, f_{10})	0.3	0.46	0.33	0.92

Table 12. Energy values of Iris flower rough graph

S. No	Similarity Measure	Energy Value
1	Jaccord	10.31
2	Dice	10.41
3	Overlap	11.69
4	ζ	13.28

Table 10 presents the similarity classes derived from the rough graph shown in Figure 6, which was constructed based on the Iris flower dataset.

Vertex labeling: $\rho(f_1) = 0.9$; $\rho(f_2) = 0.5$; $\rho(f_3) = 0.8$; $\rho(f_4) = 0.9$; $\rho(f_6) = 0.9$; $\rho(f_7) = 0.8$; $\rho(f_8) = 0.9$; $\rho(f_9) = 0.4$; $\rho(f_{10}) = 0.9$.

Table 11 presents the edge labeling values for the rough graph constructed from the Iris flower dataset, computed using the Jaccard, Overlap, Dice, and Rough ζ similarity measures.

Table 13. Iron pipes data

Iron Pipes	Coal	Sulfur	Phosphorus	Cracks
P_1	High	High	Low	Yes
P_2	Average	High	Low	No
P_3	High	High	Low	Yes
P_4	Low	Low	Low	No
P_5	Average	Low	High	Yes
P_6	High	Low	High	Yes

Table 12 presents the Energy values calculated for rough graph constructed from the Iris flower dataset, based on the adjacency matrix derived from the vertex and edge labeling using different similarity measures.

Table 14. Energy values of iron pipes data

S. No	Similarity Measure	Energy Value
1	Jaccord	6.345
2	Dice	7.626
3	Overlap	7.294
4	ζ	8.072

Illustrative Case 3: The dataset comprises information on six cast iron pipes (Table 13) subjected to a high-pressure endurance test. Let's designate the target set as $\mathcal{X} = \{P_1, P_5, P_6\}$. Rough membership values are as follows: $\omega(P_1) = 0.5; \omega(P_2) = 0; \omega(P_3) = 0.5, \omega(P_4) = 0; \omega(P_5) = 1, \omega(P_6) = 1.$

Table 14 presents the energy values computed using the Jaccard, Dice, Overlap, and ζ similarity measures for the given dataset.

Illustrative Case 4: The dataset comprises information on fifteen female patients who underwent multiple tests to assess their diabetic condition (Table 15).

Table 16 presents the Energy values computed using the Jaccard, Dice, Overlap, and ζ similarity measures for the female diabetic dataset.

Table 15. Female diabetic data

Patients	Thirst	Hunger	Frequent	Weight Loss	Tiredness	Diabetic
w_1	High	High	Low	Low	High	High
w_2	High	High	Low	Low	Low	High
w_3	High	High	High	Low	High	High
w_4	High	High	High	Low	Low	High
w_5	High	Low	High	High	High	High
w_6	High	High	High	High	High	High
w_7	High	Low	Low	Low	Low	High
w_8	High	High	High	High	High	High
w_9	High	High	Low	Low	High	Low
w_{10}	High	Low	High	Low	High	Low
w_{11}	High	High	High	Low	High	Low
w_{12}	High	Low	Low	Low	Low	Low
w_{13}	Low	High	Low	High	High	Low
w_{14}	Low	Low	Low	High	Low	Low
w_{15}	High	High	Low	Low	High	Low

Table 16. Energy values of female diabetic data

S. No	Similarity Measure	Energy Value
1	Jaccord	25.848
2	Dice	26.137
3	Overlap	25.848
4	ζ	27.4285

5. COMPARATIVE ANALYSIS OF EFFECTIVENESS OF ENERGY VALUES VARIOUS ROUGH GRAPHS

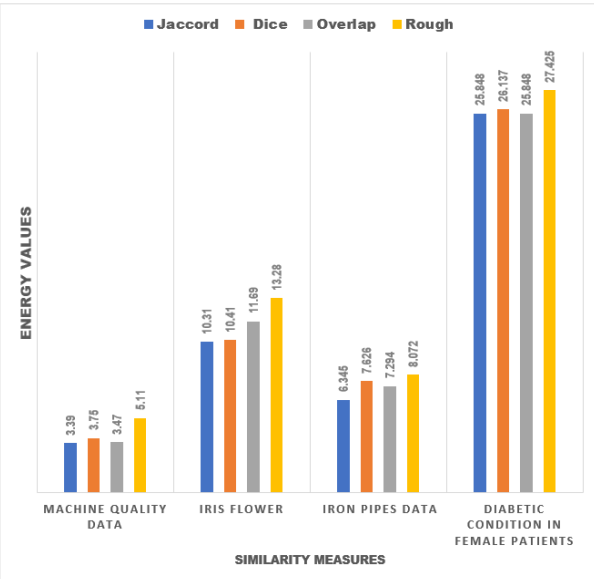

Figure 7. Graphical depiction of comparative study

We have analyzed four distinct datasets, calculating energy values using Jaccard, Dice, Overlap, and Rough ζ measures. Across all four measures, Rough ζ consistently yields the highest energy values. Rough graph energy can provide insights into the structure and properties of complex networks, such as social networks, biological networks, and communication network. Figure 7 illustrates the graphical depiction of the comparative analysis.

6. CONCLUSION

In this study, we conduct a comprehensive evaluation of various similarity measures for edge labeling in rough graphs. We present a unified view of these labeling similarity measures through the visualization of multiple bar charts. Additionally, we introduce a novel ζ labeling similarity metric designed to leverage similarity classes within rough graphs, utilizing its associated energy to characterize the graph's properties. This comprehensive overview of similarity metrics is facilitated through the representation of multiple bar charts. The energy of rough graphs can be exploited to identify significant nodes or links in a network, which has implications for network optimization, vulnerability analysis, and targeted interventions. This novel ζ labeling similarity metric, which capitalizes on similarity classes within rough graphs and quantifies the graph's characteristics through its energy, may find applications in cluster boundary region identification in Wireless Sensor Networks.

REFERENCES

[1] Tong, H. (2006). Rough graph and its structure. Journal

of Shandong University. https://en.cnki.com.cn/Article_en/CJFDTotal-SDDX200606010.htm.
[2] He, T., Chen, Y., Shi, K. (2006). Weighted rough graph and its application. In Sixth International Conference on Intelligent Systems Design and Applications, Jian, China, pp. 486-491. https://doi.org/10.1109/ISDA.2006.279
[3] He, T. (2012). Representation form of rough graph. Applied Mechanics and Materials, 157: 874-877. https://doi.org/10.4028/www.scientific.net/amm.157-158.874
[4] Mathew, B., John, S.J., Garg, H. (2020). Vertex rough graphs. Complex & Intelligent Systems, 6: 347-353. https://doi.org/10.1007/s40747-020-00133-8
[5] Pawlak, Z. (1991). Rough Sets: Theoretical Aspects of Reasoning about Data. http://bcpw.bg.pw.edu.pl/Content/1845/ZP_ksiazka.pdf.
[6] Anitha, K. (2014). Rough set theory on topological spaces. In Rough Sets and Knowledge Technology: 9th International Conference, RSKT 2014, Shanghai, China, pp. 69-74. https://doi.org/10.1007/978-3-319-11740-9_7
[7] Anitha, K., Venkatesan, P. (2016). Properties of hesitant fuzzy sets. Global Journal of Pure and Applied Mathematics (GJPAM), 12(1): 114-116.
[8] Kausar, N., Munir, M., Kousar, S., Gulistan, M., Anitha, K. (2020). Characterization of non-associative ordered semigroups by the properties of F-ideals. Fuzzy Information and Engineering, 12(4): 490-508. https://doi.org/10.1080/16168658.2021.1924513
[9] Anitha, K., Thangeswari, M. (2020). Rough set based optimization techniques. In 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, pp. 1021-1024. https://doi.org/10.1109/ICACCS48705.2020.9074212.
[10] Chen, J., Li, J. (2012). An application of rough sets to graph theory. Information Sciences, 201: 114-127. https://doi.org/10.1016/j.ins.2012.03.009
[11] Borzooei, R.A., Rashmanlou, H., Samanta, S., Pal, M. (2016). A study on fuzzy labeling graphs. Journal of Intelligent & Fuzzy Systems, 30(6): 3349-3355. https://doi.org/10.3233/ifs-152082
[12] Gani, A.N., Akram, M., Subahashini, D.R. (2014). Novel properties of fuzzy labeling graphs. Journal of Mathematics, 1-6. https://doi.org/10.1155/2014/375135
[13] Ramsankar, A.D., Krishnamoorthy, A. (2024). Exploring metric dimensions for dimensionality reduction and navigation in rough graphs. Mathematical Modelling of Engineering Problems, 11(4): 1037-1043. https://doi.org/10.18280/mmep.110421
[14] Nithya, R., Anitha, K. (2023). Even vertex ζ-graceful labeling on rough graph. Mathematics and Statistics, 11(1): 71-77. https://doi.org/10.13189/ms.2023.110108
[15] Gutman, I. (2001). The energy of a graph: Old and new results. In Algebraic Combinatorics and Applications: Proceedings of the Euro conference, Algebraic Combinatorics and Applications (ALCOMA), Gößweinstein, Germany, pp. 196-211. https://doi.org/10.1007/978-3-642-59448-9_13
[16] Li, X., Li, Y., Shi, Y., Gutman, I. (2012). Graph Energy. Springer, New York. https://doi.org/10.1007/978-1-4614-4220-2
[17] Gutman, I., Ramane, H.S. (2020). Research on graph energies in/2019. MATCH Communications in Mathematical and in Computer Chemistry, 84: 277-292.
[18] Nikiforov, V. (2007). The energy of graphs and matrices. Journal of Mathematical Analysis and Applications, 326(2): 1472-1475. https://doi.org/10.1016/j.jmaa.2006.03.072
[19] Nagarani, A., Vimala, S. (2017). Energy of fuzzy regular and graceful graphs. Asian Research Journal of Mathematics, 4(2): 1-8. https://doi.org/10.9734/arjom/2017/33057
[20] Ilić, A., Bašić, M. (2019). Path matrix and path energy of graphs. Applied Mathematics and Computation, 355: 537-541. https://doi.org/10.1016/j.amc.2019.03.002
[21] Pirzada, S., Ganie, H.A. (2015). On the Laplacian eigenvalues of a graph and Laplacian energy. Linear Algebra and its Applications, 486: 454-468. https://doi.org/10.1016/j.laa.2015.08.032
[22] Pappis, C.P., Karacapilidis, N.I. (1993). A comparative assessment of measures of similarity of fuzzy values. Fuzzy Sets and Systems, 56(2): 171-174. https://doi.org/10. 1016/0165-0114(93)90141-4.
[23] Zadeh, L.A. (1971). Similarity relations and fuzzy orderings. Information Sciences, 3(2): 177-200. https://doi.org/10.1016/S0020-0255(71)80005-1
[24] Wang, W.J. (1997). New similarity measures on fuzzy sets and on elements. Fuzzy Sets and Systems, 85(3): 305-309. https://doi.org/10.1016/0165-0114(95)00365-7
[25] Candan, K.S., Li, W.S., Priya, M.L. (2000). Similarity-based ranking and query processing in multimedia databases. Data & Knowledge Engineering, 35(3): 259-298. https://doi.org/10.1016/S0169-023X(00)00025-2
[26] Shijina, V., Unni, A., John, S.J. (2020). Similarity measure of multiple sets and its application to pattern recognition. Informatica, 44(3): 335-347. https://doi.org/10.31449/inf.v44i3.2872
[27] Farhadinia, B., Ban, A.I. (2013). Developing new similarity measures of generalized intuitionistic fuzzy numbers and generalized interval-valued fuzzy numbers from similarity measures of generalized fuzzy numbers. Mathematical and Computer Modelling, 57(3-4): 812-825. https://doi.org/10.1016/j.mcm.2012.09.010
[28] Kóczy, L.T., Susniene, D., Purvinis, O., Konczosné Szombathelyi, M. (2022). A new similarity measure of Fuzzy signatures with a case study based on the statistical evaluation of questionnaires comparing the influential factors of Hungarian and lithuanian employee engagement. Mathematics, 10(16): 2923. https://doi.org/10.3390/ math10162923
[29] Nithya, R. (2024). Unveiling the potential of similarity measures in rough labeling of graphs. Asia Pacific Journal of Mathematics, 11(22). https://doi.org/10.28924/APJM/11-22
[30] Vijaymeena, M.K., Kavitha, K. (2016). A survey on similarity measures in text mining. Machine Learning and Applications: An International Journal, 3(2): 19-28. https://doi.org/10.5121/mlaij.2016.3103.
[31] Jimenez, S., Gonzalez, F., Gelbukh, A. (2010). Text comparison using soft cardinality. In 17th International Symposium, SPIRE 2010, Los Cabos, Mexico, pp. 297-302. https://doi.org/10.1007/978-3-642-16321-0_31
[32] Liu, X. (1992). Entropy, distance measure and similarity measure of fuzzy sets and their relations. Fuzzy Sets and Systems, 52(3): 305-318. https://doi.org/10.1016/0165-0114(92)90239-Z