



*Research article***A topological-lattice similarity measure for information systems****Mateb O. Aljabri¹, Abdelfattah A. El-Atik^{1,*} and Ahmed Zedan²**¹ Department of Mathematics, Faculty of Science, Tanta University, Tanta 31527, Egypt² Department of Mathematics, Faculty of Science, Aswan University, Aswan 81528, Egypt*** Correspondence:** Email: aelatik@science.tanta.edu.eg.

Abstract: In this paper, a topological framework induced by information systems is developed and investigated. The associated topological spaces are constructed through attribute-based relations, and their fundamental topological properties are examined. Based on these topological representations, a similarity-based framework is developed to quantify and analyze the degrees of similarity among objects in an information system. The proposed measure captures structural characteristics through topological concepts such as neighborhoods, closures, and open sets. To further enhance the analysis, lattice structures associated with the generated topologies are employed to reveal hierarchical relationships and provide an additional perspective for similarity assessment. Furthermore, an attribute reduction procedure is proposed to identify and eliminate redundant attributes while preserving the essential information contained in the system. The proposed method provides a theoretical framework for attribute reduction based on topological and lattice structures and is illustrated through a representative information-system example. The obtained results show that the integration of topological spaces and lattice structures provides a robust and efficient framework for measuring similarity between information systems, supporting data reduction, information analysis, and decision-making applications.

Keywords: topological spaces; similarity matrices; information system; lattice theory; rough set**Mathematics Subject Classification:** 03E75, 06A06, 06B23, 54A05

1. Introduction

The topological character of a set serves as the structure from which knowledge is extracted from data [1]. Topological structures provide a mathematical framework for representing connections among data objects and for extracting structural information from datasets (which are closed under finite intersections and arbitrary unions). The family β of subsets serves as a base for the topological space associated with the universe set U if every open set in U represents a finite number of unions of some

members drawn from β [2]. The collection of open sets (called σ) that is drawn from universe set U is established by that collection having an empty subset of the set of all members drawn from σ , an open subset from U , and remaining closed under finite intersections for this topological space. The information system [3–5] consists of a finite set of objects described by a finite set of attributes, each associated with a corresponding set of different values, which can be represented in a table called an information table. In fact, there are two types of information tables [6, 7]; the first is a complete information table, and when the information table contains missing attribute values, it is referred to as an incomplete information system. In an information table [8, 9], the neighborhood of an object is determined by the chosen relation among objects. Such neighborhoods can be used to define approximation operators (lower and upper approximations) and accuracy measures. An approximation space [10] is a mathematical concept used in rough set theory to model vagueness and uncertainty and depend on equivalence relation, which it is a base for quasi-discrete topology to approximate a larger set. Pawlak introduced rough set theory [11] as a mathematical approach used to deal with incomplete and uncertain information in intelligent systems by defining two exact sets, lower and upper approximation, with clear and possible members. Lattice theory [12] proves valuable for analyzing structures in theoretical computer science. Formal concept analysis [12] provides insight into a graphical knowledge representation method, specifically the diagrams of lattices. The method of formal concept analysis that Ganter and Wille [13] are developed serves multiple fields, including psychology, sociology, anthropology, medicine, biology, linguistics, computer sciences, mathematics and industrial engineering. Neighborhoods [14–16] play a key role in mathematics and may have different meanings depending on the information table. Generally, a neighborhood of an object o in an information table is the set of all objects that have the same attribute value. In other words, topologically, a neighborhood of an object o is a set that contains an open set containing o . So, a neighborhood of an object in a topological space is important concepts, is closely defined the concepts of open set that is important to obtain on interior and closure in topology. Open neighborhood in approximation space is defined by equivalence class, which consists of all objects that have the attribute values and information data. The German mathematician Felix Hausdorff published the importance of the concept of a neighborhood and its usefulness in the formulation of a topological space. Also, neighborhood objects [17] in an information table are the set that provides data about it, so it helps to know the information it, and we can compute the degree of accuracy measurement [18, 19] with respect to our knowledge base. We use the neighborhood of an object via attributes from the information table to indicate the complete vision of objects. The neighborhood of a vertex x in a graph is the set of vertices adjacent to x . Attribute reduction in the information table identifies superfluous attributes that can be removed without losing essential classification information [20, 21]. We can see attribute reduction as the process of discovering a more compact group of features with the same or near classification as the initial set. Throughout this study, we establish similarity-based neighborhood systems from an information system and use them to generate quasi-discrete, pretopological, and lattice-based structures. These structures are then used to evaluate the effectiveness of attribute reduction and analyze the information system.

2. Preliminaries

The idea of pretopology [22–25] stems from the concept of topology, as topological notions can use open sets, neighborhoods, and a Kuratowski closure operator, which are basic notions in topology.

Definition 2.1. A function $CL : P(U) \rightarrow P(U)$ is called closure operator if, for all $X, Y \subset U$,

- (1) $X \subset CL(X)$; inflation,
- (2) if $X \subset Y$, then $CL(X) \subset CL(Y)$; monotone,
- (3) $CL(CL(X)) = CL(X)$; idempotency.

Definition 2.2. Let $cl : \mathcal{P}(U) \rightarrow \mathcal{P}(U)$ be a closure operator. The collection $C = \{A \subseteq U : cl(A) = A\}$ of all fixed points of cl is called the closure system induced by cl . The family C is closed under arbitrary intersections.

Definition 2.3. The pair (X, a) , is called a pretopological space if $a : P(X) \rightarrow P(X)$ satisfies the following:

- (1) $a(\phi) = \phi$.
- (2) $A \subset a(A), \forall A \subset X$.

Definition 2.4. [26] Let $\mathcal{I}_{ST} = (U, C, D, f, v_i \in V)$ be an information system, where $U \neq \phi$ is called the set of objects of approximation space, $\{C, D\}$ is the set of attributes, $f : U \times U \rightarrow V$ is the function from $U \times U$ to V , and V is the set values between objects and attributes.

In other words, an information system is an ordered $\mathcal{I}_{ST} = (U, C, D, f, v_i \in V)$ such that $U = \{x_i : i \in I\}$ is a nonempty finite set of items (pupils, playthings, $\dots$) termed universe, $AT = \{C, D\}$ comprises a nonempty group of attributes (colors, characteristic, $\dots$), and $V = \{v_i : i \in I\}$ is an attribute values collection.

Definition 2.5. [27] Let $\mathcal{I}_{ST} = (U, C, D, f, v_i \in V)$ be an information system, then the neighborhood class of every object is $[y] = \{x \in O : f(y, a) = f(x, a) = v : v \in V_a, \forall a \in A\}$.

Definition 2.6. [28] Let $\mathcal{I}_{ST} = (U, C, D, f, v_i \in V)$ be an information system, then the subbase σ for a topology $(O; \tau_U)$ is generated by the neighborhood class of every object of a given information system.

Definition 2.7. [29] Let $\mathcal{I}_{ST} = (U, C, D, f, v_i \in V)$ be a complete information system and the base β for a topology of the approximation space be the set of all equivalence classes in U , then for $X \subset U$, the union of equivalence classes contained in X is a lower approximation (the union of equivalence classes intersected with X is an upper approximation), and the ratio between the lower and upper approximation is the accuracy measure of X .

3. Assessing precision of information systems via similarity matrix

In this section, we use numerical measurements to obtain the equivalence classes for the information system immediately, which is the main method to derive different topologies with respect to similarity measurement between objects without the lengthy procedures as in previous work, that is, for every

similarity-separated value generating different topologies. We consider only two cases: unit similarity and partial similarity, focusing on the condition attributes. We also investigate some topological operators with respect to the decision attribute. In an information system and its corresponding similarity matrix, attribute reduction and the core of the attribute set are obtained by deleting attributes, repeating the preceding steps, and finally selecting the best numerical results. Complete and partial similarities between objects are determined based on the proportion of shared attribute values in the information system. These methods provide a new approach to studying the same processes by constructing a lattice among individual objects in the universe and clarifying the extent to which objects share attribute values within the information system. Finally, we define a similarity matrix based on the condition attributes. By applying different thresholds to this matrix, we obtain neighborhood systems that are subsequently used to construct the topological structures studied in this paper.

Let $\mathcal{I}_{ST} = (U, C, D, f, v_i \in V)$ be an information table, where $U \neq \emptyset$ is a universe set, $\{C, D\}$ is a nonempty finite set of conditions and decision attributes, and V is a set of attribute values. Our research focuses on developing new pretopological domains for information systems by implementing similarity methods.

Lemma 3.1. *Let $\mathcal{I}_{ST} = (U, C, D, f, v_i \in V)$ be an information system and q denote the value of conditional attributes. Then, the matrix of similarity values is defined as*

$$v_{C_k}(x_i, x_j) = \begin{cases} 1 & \text{if } C_k(x_i, q) = C_k(x_j, q), \\ 0 & \text{if } C_k(x_i, q) \neq C_k(x_j, q), \end{cases}$$

where $i, j = 1, 2, \dots, |U|$ and $k = 1, 2, 3, \dots, |C|$.

Definition 3.2. *For any information system $\mathcal{I}_{ST} = (U, C, D, f, v_i \in V)$, where U is a set of objects, $\{C, D\}$ is a set of attributes, and Q is the subset of the ordinal attribute scale, each entry in the similarity matrix is constructed by similarity degree $\mu_i(x_i, x_j) = \frac{\sum_{k=1}^{|C|} v_{C_k}(x_i, x_j)}{|C|}$, where $x_i, x_j \in U, i, j = 1, 2, \dots, |U|$. When choosing the values for μ , the conditions of a pretopology denoted in [22] must be considered.*

Definition 3.3. *For any information system $\mathcal{I}_{ST} = (U, C, D, f, v_i \in V)$ and $\xi : x_i \rightarrow P(U)$, where $P(U)$ is the power of U , for each object $x_i \in U$, define the neighborhood $\xi_\delta(x_i) = \{x_j \in U : \mu(x_i, x_j) \geq \delta\}$, $\delta \in [0, 1]$. The collection $\mathcal{S} = \{\xi(x_i) : x_i \in U\}$ forms a subbase for the pretopological structure generated by the similarity relation.*

Proposition 3.4. *Let $\mathcal{I}_{ST} = (U, C, D, f, v_i \in V)$ be an information system and μ be a similarity degree of a similarity matrix. Therefore, for $\delta = 1$, the relation $x_i \sim x_j \iff \mu(x_i, x_j) = 1$ is an equivalence relation, and the family of its equivalence classes forms a base for the corresponding quasi-discrete topology.*

Proposition 3.5. *Let $\mathcal{I}_{ST} = (U, C, D, f, v_i \in V)$ be an information system for each object $x_i \in U$. Then,*

- (1) *The subbase of the kernel pretopological space in the information system is $\psi = \{\xi_\delta(x_i) : x_i \in U, \mu(x_i, x_j) = 0\}$.*
- (2) *The base of the quasi-discrete space on information system is $\sigma = \{\xi_\delta(x_i) : x_i \in U, \mu(x_i, x_j) = 1\}$.*
- (3) *The family $\xi = \{\xi_\delta(x_i) : x_i \in U, 0 < \mu(x_i, x_j) < 1\}$ serves as a subbase of the associated pretopological structure.*

Proof. (1) For $\mu(x_i, x_j) = 0$, the relation $R = \{(x_i, x_j) \in U \times U : \mu(x_i, x_j) = 0\}$ is symmetric and not reflexive and transitive.

(2) For $\mu(x_i, x_j) = 1$, we prove that the relation $R = \{(x_i, x_j) \in U \times U : \mu(x_i, x_j) = 1\}$ is reflexive, symmetric, and transitive. Consequently, it induces a partition of the universe U and generates the corresponding quasi-discrete topology.

To prove this, it must be proved that all ordered pairs (x_i, x_j) are equivalence relations. First, let (x_i, x_i) for $x_i \in U$. Then, $\mu(x_i, x_i) = \frac{\sum_{j=1}^{|C|} v_k(x_i, x_i)}{|C|} = \frac{1+\dots+1}{|C|} = 1$. So, the relation is reflexive. Second, let $x_i, x_j \in U$ and

$$v_k(x_i, x_j) = \begin{cases} 1 & \text{if } C_k(x_i, q) = C_k(x_j, q) \\ 0 & \text{if } C_k(x_i, q) \neq C_k(x_j, q) \end{cases} = \begin{cases} 1 & \text{if } C_k(x_j, q) = C_k(x_i, q) \\ 0 & \text{if } C_k(x_j, q) \neq C_k(x_i, q) \end{cases} = v_k(x_j, x_i).$$

Then, the relation is symmetric. Lastly, let $x_i, x_j, x_l \in U$ and $\mu(x_i, x_j) = 1$, and $\mu(x_j, x_l) = 1$ means that

$$v_k(x_i, x_j) = \begin{cases} 1 & \text{if } C_k(x_i, q) = C_k(x_j, q) \\ 0 & \text{if } C_k(x_i, q) \neq C_k(x_j, q) \end{cases}, v_k(x_j, x_l) = \begin{cases} 1 & \text{if } C_k(x_j, q) = C_k(x_l, q) \\ 0 & \text{if } C_k(x_j, q) \neq C_k(x_l, q) \end{cases},$$

then

$$v_k(x_i, x_l) = \begin{cases} 1 & \text{if } C_k(x_i, q) = C_k(x_l, q) \\ 0 & \text{if } C_k(x_i, q) \neq C_k(x_l, q) \end{cases}.$$

Then, $\mu(x_i, x_l) = 1$. Therefore, the relation is transitive, and all ordered pairs satisfy equivalence relation conditions and form a partition of the universe U .

(3) However, for $0 < \mu(x_i, x_j) < 1$, the family $\xi(x_i) = \{x_j \in U : \mu(x_i, x_j) > \delta\}$, and $\delta \in (0, 1)$ is employed solely to construct a neighborhood-based pretopological structure. In general, the relation

$$R_\delta = \{(x_i, x_j) \in U \times U : \mu(x_i, x_j) > \delta\}$$

is not necessarily transitive and therefore does not, in general, define an equivalence relation or induce equivalence classes on U .

□

Definition 3.6. For any information system $\mathcal{I}_{ST} = (U, C, D, f, v_i \in V)$, where U is a set of objects and μ the similarity degrees, the accuracy measurement of the similarity density of an object in the information system with respect to the similarity matrix is $\mathcal{P}(x_i) = \frac{\sum_{j=1}^{|U|} \mu(x_i, x_j)}{|U|}$, $i, j = 1, 2, \dots, |U|$.

Remark 3.7. The similarity score $\mathcal{P}(x_i)$ is uniquely defined in Definition 3.6. Although, $\mathcal{P}(x_i)$ and $\mathcal{T}(x_i)$ may have the same numerical value in the present setting, they are obtained through different aggregation procedures and therefore have different computational interpretations. Hence, they are denoted by different symbols throughout the manuscript.

Definition 3.8. Let $\mathcal{I}_{ST} = (U, C, D, f, v_i \in V)$ be an information system and $P : U \rightarrow \mathbb{R}$ be the similarity-score function. Define a relation $\leq$ on U by $x_i \leq x_j \iff P(x_i) \leq P(x_j)$. Then $(U, \leq)$ is a preorder.

Definition 3.9. Let $\mathcal{I}_{ST} = (U, C, D, f, v_i \in V)$ be an information system. Define an equivalence relation $x_i \sim x_j \iff P(x_i) = P(x_j)$, where $U/\sim$ denote the set of equivalence classes. Then the induced order on the quotient set is defined by $[x_i] \leq [x_j] \iff P(x_i) \leq P(x_j)$.

Definition 3.10. Let $\mathcal{I}_{ST} = (U, C, D, f, v_i \in V)$ be an information system, $(U, \leq)$ be the lattice of $\mathcal{I}_{ST} = (U, C, D, f, v_i \in V)$ and $A \subset U$. Then,

- (a) An object $x_i \in U$ is the maximum node of A if $\mathcal{P}(x_j) < \mathcal{P}(x_i), \forall x_j \in A$.
- (b) An object $x_i \in U$ is the minimum node of A if $\mathcal{P}(x_j) > \mathcal{P}(x_i), \forall x_j \in A$.
- (c) An object $x_i \in A$ is the supremum node of A if x_i is the largest object with respect to $\mathcal{P}(x_i)$ denoted by $x_i = \text{Sup}\{A\} > x_j$, where $\mathcal{P}(x_j) < \mathcal{P}(x_i)$.
- (d) An object $x_i \in A$ is the infimum bound of A if x_i is the smallest object with respect to $\mathcal{P}(x_i)$ denoted by $\text{inf}\{A\}$.

Definition 3.11. Define a relation $\leq$ on U by $x_i \leq x_j \iff P(x_i) \leq P(x_j)$. Since $P(x_i) \in [0, 1]$, the relation $\leq$ is reflexive, antisymmetric, and transitive. Hence $(U, \leq)$ is preorder.

Definition 3.12. Let $A \subseteq U$. The supremum of A , denoted by $\text{sup}(A)$, is the least upper bound of A with respect to $\leq$. The infimum of A , denoted by $\text{inf}(A)$, is the greatest lower bound of A with respect to $\leq$.

Proposition 3.13. Let $\mathcal{I}_{ST} = (U, C, D, g, v_i \in V)$ be an information system, then for any equivalence classes $[x_i], [x_j] \in U/\sim$,

$$[x_i] \wedge [x_j] = \min\{[x_i], [x_j]\}$$

and

$$[x_i] \vee [x_j] = \max\{[x_i], [x_j]\}$$

exist with respect to the induced order. Hence $(U/\sim, \leq)$ is a chain lattice.

Proof. The quotient classes are totally ordered by their score values. Every pair of classes therefore admits a unique minimum and maximum element. These correspond respectively to the meet and join operations. Hence, the quotient ordered forms a chain lattice. $\square$

Proposition 3.14. Let $\mathcal{I}_{ST} = (U, C, D, f, v_i \in V)$ be an information system, and let $P : U \longrightarrow [0, 1]$ be a function. Define a binary relation R_P on U by $R_P = \{(x_i, x_j) \in U \times U : P(x_i) = P(x_j)\}$. Equivalently, $x_i \sim x_j \iff P(x_i) = P(x_j)$, $x_i, x_j \in U$. Then, R_P is an equivalence relation on U .

Proof. We show that R_P is reflexive, symmetric, and transitive.

- (1) Reflexivity: Let $x_i \in U$. Since $P(x_i) = P(x_i)$, it follows that $(x_i, x_i) \in R_P$. Therefore, R_P is reflexive.
- (2) Symmetry: Let $x_i, x_j \in U$, and suppose that $(x_i, x_j) \in R_P$. Then $P(x_i) = P(x_j)$. By the symmetry of equality, $P(x_j) = P(x_i)$, and hence $(x_j, x_i) \in R_P$. Therefore, R_P is symmetric.
- (3) Transitivity: Let $x_i, x_j, x_k \in U$, and suppose that $(x_i, x_j) \in R_P$ and $(x_j, x_k) \in R_P$. Then, $P(x_i) = P(x_j)$ and $P(x_j) = P(x_k)$. Hence, $P(x_i) = P(x_k)$, which implies that $(x_i, x_k) \in R_P$. Therefore, R_P is transitive.

Thus, R_P is reflexive, symmetric, and transitive. Therefore, R_P is an equivalence relation on U . Note That the equality relation is both an equivalence and ordered relation. $\square$

Proposition 3.15. Let $\mathcal{I}_{ST} = (U, C, D, f, V)$ be an information system, and let $R_P = \{(x_i, x_j) \in U \times U : P(x_i) = P(x_j)\}$ be the equivalence relation on U induced by P . Then the quotient set $U/R_P = \{[x_i]_{R_P} : x_i \in U\}$, where $[x_i]_{R_P} = \{x_j \in U : x_i, x_j \in R_P\} = \{x_j \in U : P(x_i) = P(x_j)\}$, satisfies the following:

- (1) The quotient set U/R_P is a partition of U .
- (2) The topology induced by U/R_P is $\tau_{R_P} = \{\bigcup \mathcal{A} : \mathcal{A} \subseteq U/R_P\}$. Hence, (U, τ_{R_P}) is a quasi-discrete topological space, and U/R_P forms a base for τ_{R_P} .

Definition 3.16. Let $\mathcal{I}_{ST} = (U, C, D, f, v_i \in V)$ be an information system and $(U/\sim, \leq)$ be the lattice of an information system. For any $x_i, x_j \in U$,

$$x_i \wedge x_j = \min_{\leq} \{x_i, x_j\},$$

and

$$x_i \vee x_j = \max_{\leq} \{x_i, x_j\}.$$

These correspond respectively to the meet and join operations of the chain lattice $(U/\sim, \leq)$.

Definition 3.17. Let $\mathcal{I}_{ST} = (U, C, D, f, v_i \in V)$ be an information system and $P : x_i \rightarrow [0, 1]$ with $P(x_i) = \frac{\sum_{j=1}^{|U|} \mu(x_i, x_j)}{|U|}$. Then, $(x_i, P(x_i)), x_i \in U$ is the set of vertices of the lattice $(U, \leq)$, where $\leq$ is the ordered relation.

Proposition 3.18. Let $(U, \leq)$ be the lattice of information system $\mathcal{I}_{ST} = (U, C, D, f, v_i \in V)$ for any objects $x_i, x_j \in U$. Then,

- (1) $x_i = x_j \iff P(x_i) = P(x_j)$ are equal and parallel objects on the lattice.
- (2) $x_i \leq x_j \iff P(x_i) \leq P(x_j)$.

$(x_i, P(x_i)), x_i \in U$ is the set of vertices of the lattice $(U, \leq)$ induced by $\mathcal{I}_{ST}$.

Remark 3.19. Let $\mathcal{I}_{ST} = (U, C, D, f, v_i \in V)$ be an information system, and let $(x_i, P(x_i)), x_i \in U$ be the set of vertices of the lattice $(U, \leq)$ induced by $\mathcal{I}_{ST}$. Then, the equivalent is

- (1)
$$N_{\leq}(x_i) = \left\{ x_j \in U : \begin{array}{l} x_i < x_j \text{ and there is no } x_k \in U \text{ such that } x_i < x_k < x_j, \\ \text{or} \\ x_j < x_i \text{ and there is no } x_k \in U \text{ such that } x_j < x_k < x_i \end{array} \right\}.$$

- (2) For a subset $A \subseteq U$, the interior of A , denoted by $\text{Int}(A)$, consists of all objects $x_i \in A$ whose $N_{\leq}(x_i)$ -neighborhood is completely contained in A . That is,

$$\text{Int}(A) = \{x_i \in A : N_{\leq}(x_i) \subseteq A\}.$$

- (3) The closure of A , denoted by $\text{cl}(A)$, consists of all objects in U whose $N_{\leq}(x_i)$ -neighborhoods intersect A . That is,

$$\text{cl}(A) = \{x_i \in U : N_{\leq}(x_i) \cap A \neq \emptyset\}.$$

(4) In the graphical representation of the lattice, the red edges represent the $N_{\leq}(x_i)$ -neighborhood relation among the objects. Now, for a subset $A \subseteq U$, the interior of A , denoted by $\text{Int}(A)$, consists of the objects of A that are connected by the red edges in the lattice. The closure of A , denoted by $\text{cl}(A)$, consists of all objects in U that belong to edges intersecting A .

Remark 3.20. Given an information system $\mathcal{I}_{\mathcal{ST}} = (U, C, D, g, v_i \in V)$ and $A \subset U$, there are three methods to obtain the interior and closure with the respective pretopology:

- (1) Topological space on information system $\text{int}_Q(A) = \cup\{B \in \mathcal{S}(x_i), B \subset A\}$, $\text{cl}_Q(A) = \cup\{B \in \mathcal{S}(x_i), B \cap A \neq \emptyset\}$.
- (2) Pretopological space on information system $\text{cl}_R(A) = A \cup \{x_i \in U, \xi(x_i) \cap A \neq \emptyset\}$, $\text{int}_R(A) = \text{co}(\text{cl}_R(\text{co}A))$, where $\text{co}(A)$ is a complement of A , $A \subset X$.
- (3) Analytical pretopology lattice (lattice of information system) $\text{cl}(A) = \{x_i \in U, \inf(A) \leq x_i \leq \sup(A)\}$ and $\text{int}(A) = \{x_j, x_i \wedge x_j \in A \text{ and } x_i \vee x_j \in A, x_i \in A\}$.

Remark 3.21. In the analytical lattice representation, the closure of a set A collects the objects whose similarity scores lie within the order interval determined by the lower and upper score levels associated with A . Thus, the closure represents the extent to which the similarity structure connects A with surrounding objects in the induced order. Conversely, the interior identifies objects of A that remain structurally supported within A under the lattice ordering. Hence, the lattice-based interior and closure provide an order-theoretic interpretation of local similarity and structural containment among information-system objects.

Remark 3.22. In a general pretopological space, the interior operator and the closure operator need not satisfy the complement-duality relation $\text{int}(A) = \text{co}(\text{cl}(\text{co}(A)))$. This identity holds for topological spaces and, more generally, for pretopological structures in which the interior and closure operators are defined to satisfy the closure–interior duality property. Hence, in the pretopological setting, this identity will be used only when the corresponding duality condition is explicitly assumed.

In the following computations, the complement-based interior operator is used under the closure–interior duality condition stated in Remark 3.21.

Example 3.23. (Effectiveness of attributes) For any information system $\mathcal{I}_{\mathcal{ST}} = (U, C, D, g, v_i \in V)$, where $U = \{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}$, $|U| = 8$, and $|C| = 4$ (Trans. Line 1 = $\mathcal{A}$, Trans. Line 2 = $\mathcal{B}$, Self-Gem = $\mathcal{C}$ and Self-Load = $\mathcal{L}$) as in Table 1. We begin with the complete information system, construct the similarity matrix, and then compute the corresponding interior, closure, and accuracy measures of the decision attributes. We repeat the above process after deleting one of the condition attributes and compare the results to obtain the important attributes. The information system is a table form that contains all the data needed by the user to give a decision.

Table 1. Complete information system.

	$\mathcal{A}$	$\mathcal{B}$	$\mathcal{C}$	$\mathcal{L}$	Decision
$\mathcal{E}_1$	Medium	Medium	Medium	Medium	Normal
$\mathcal{E}_2$	Medium	Low	Low	Medium	Rest
$\mathcal{E}_3$	Medium	Low	Low	Medium	Normal
$\mathcal{E}_4$	Low	Low	Low	High	Rest
$\mathcal{E}_5$	Low	Medium	Medium	Medium	Rest
$\mathcal{E}_6$	Medium	Medium	Low	High	Normal
$\mathcal{E}_7$	Medium	Medium	Low	Low	Rest
$\mathcal{E}_8$	Medium	Medium	Low	High	Normal

In the following, the comparison between the given results and Pawlak's approximation spaces of the information system with respect to rough sets concepts will be studied. The approximation space for every condition attributes of the information system is given by $U/\mathcal{A} = \{\{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}, \{\mathcal{E}_4, \mathcal{E}_5\}\}$, $U/\mathcal{B} = \{\{\mathcal{E}_1, \mathcal{E}_5, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}, \{\mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4\}\}$, $U/\mathcal{C} = \{\{\mathcal{E}_1, \mathcal{E}_5\}, \{\mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}\}$, $U/\mathcal{L} = \{\{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_5\}, \{\mathcal{E}_4, \mathcal{E}_6, \mathcal{E}_8\}, \{\mathcal{E}_7\}\}$.

The approximation space of the decision attribute of the information system is given by

$$U/\mathcal{D} = \{\{\mathcal{E}_1, \mathcal{E}_3, \mathcal{E}_6, \mathcal{E}_8\}, \{\mathcal{E}_2, \mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_7\}\}.$$

The approximation space of the information system with respect to all condition attributes is given by

$$U/Ind(C) = \{\{\mathcal{E}_1\}, \{\mathcal{E}_2, \mathcal{E}_3\}, \{\mathcal{E}_4\}, \{\mathcal{E}_5\}, \{\mathcal{E}_6, \mathcal{E}_8\}, \{\mathcal{E}_7\}\}.$$

The approximation space for all attributes of the information system excluding Trans. Line attribute given by

$$U/Ind(C) - \mathcal{A} = \{\{\mathcal{E}_1, \mathcal{E}_5\}, \{\mathcal{E}_2, \mathcal{E}_3\}, \{\mathcal{E}_4\}, \{\mathcal{E}_6, \mathcal{E}_8\}, \{\mathcal{E}_7\}\}$$

is not reduct.

The approximation space for all attributes of the information system excluding Trans. Line attribute given by

$$U/Ind(C) - \mathcal{B} = \{\{\mathcal{E}_1\}, \{\mathcal{E}_2, \mathcal{E}_3\}, \{\mathcal{E}_4\}, \{\mathcal{E}_5\}, \{\mathcal{E}_6, \mathcal{E}_8\}, \{\mathcal{E}_7\}\}$$

is reduct.

The approximation space for all attributes of the information system excluding Self-Gem attribute given by

$$U/Ind(C) - \mathcal{C} = \{\{\mathcal{E}_1\}, \{\mathcal{E}_2, \mathcal{E}_3\}, \{\mathcal{E}_4\}, \{\mathcal{E}_5\}, \{\mathcal{E}_6, \mathcal{E}_8\}, \{\mathcal{E}_7\}\}$$

is reduct.

The approximation space for all attributes of the information system excluding Self-Load attribute given by

$$U/Ind(C) - \mathcal{L} = \{\{\mathcal{E}_1\}, \{\mathcal{E}_2, \mathcal{E}_3\}, \{\mathcal{E}_4\}, \{\mathcal{E}_5\}, \{\mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}\}$$

is not reduct.

The reducts of approximation space of the information system are $\{\mathcal{A}, \mathcal{C}, \mathcal{L}\}$ and $\{\mathcal{A}, \mathcal{B}, \mathcal{L}\}$, and the core is $\{\mathcal{A}, \mathcal{L}\}$

Example 3.24. (Continuing from Example 3.23) The similarity matrix needs us to position elements of U in both horizontal and vertical directions before we can find the value for each cell. When we compute $\mu(\mathcal{E}_i, \mathcal{E}_j)$, we summarize each attribute in $\mathcal{E}_i$ to $\mathcal{E}_j$ for each value of the attribute as shown in Table 2. When the attribute has the same order for the same attribute, it is valued 1; otherwise, it is given the value 0.

Table 2. Similarity matrix values of the information system.

	$\mathcal{E}_1$	$\mathcal{E}_2$	$\mathcal{E}_3$	$\mathcal{E}_4$	$\mathcal{E}_5$	$\mathcal{E}_6$	$\mathcal{E}_7$	$\mathcal{E}_8$	$\mathcal{P}(\mathcal{E}_i)$
$\mathcal{E}_1$	1	0.50	0.50	0	0.75	0.50	0.50	0.50	0.53
$\mathcal{E}_2$	0.50	1	1	0.50	0.25	0.50	0.50	0.50	0.59
$\mathcal{E}_3$	0.50	1	1	0.50	0.25	0.50	0.50	0.50	0.59
$\mathcal{E}_4$	0	0.50	0.50	1	0.25	0.50	0.25	0.50	0.44
$\mathcal{E}_5$	0.75	0.25	0.25	0.25	1	0.25	0.25	0.25	0.41
$\mathcal{E}_6$	0.50	0.50	0.50	0.50	0.25	1	0.75	1	0.63
$\mathcal{E}_7$	0.50	0.50	0.50	0.25	0.25	0.75	1	0.75	0.56
$\mathcal{E}_8$	0.50	0.50	0.50	0.50	0.25	1	0.75	1	0.63
Mean similarity score									0.55

Next, the comparison between pretopological spaces and the general topology of the information system is given.

Case 1. Pretopological spaces of information system

- (1) The neighborhoods $\sigma(x_i)$ for all $x_i \in U$ of the information system by using a similarity matrix at $\delta = 1$ are given by

$$\begin{aligned} \sigma\{\mathcal{E}_1\} &= \{\mathcal{E}_1\}, & \sigma\{\mathcal{E}_2\} &= \{\mathcal{E}_2, \mathcal{E}_3\}, & \sigma\{\mathcal{E}_3\} &= \{\mathcal{E}_2, \mathcal{E}_3\}, & \sigma\{\mathcal{E}_4\} &= \{\mathcal{E}_4\}, \\ \sigma\{\mathcal{E}_5\} &= \{\mathcal{E}_5\}, & \sigma\{\mathcal{E}_6\} &= \{\mathcal{E}_6, \mathcal{E}_8\}, & \sigma\{\mathcal{E}_7\} &= \{\mathcal{E}_7\}, & \sigma\{\mathcal{E}_8\} &= \{\mathcal{E}_6, \mathcal{E}_8\}. \end{aligned}$$

Then, the interior and closure of $N = \{\mathcal{E}_1, \mathcal{E}_3, \mathcal{E}_6, \mathcal{E}_8\}$ is given by $cl_Q(N) = \{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_6, \mathcal{E}_8\}$ and $int_Q(N) = \{\mathcal{E}_1, \mathcal{E}_6, \mathcal{E}_8\}$. Then, the degree of accuracy is $\mu(N) = \frac{||\{\mathcal{E}_1, \mathcal{E}_6, \mathcal{E}_8\}||}{||\{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_6, \mathcal{E}_8\}||} = 0.60$.

The interior and closure of $R = \{\mathcal{E}_2, \mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_7\}$ is given by $cl_Q(R) = \{\mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_7\}$ and $int_Q(R) = \{\mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_7\}$. Then, the degree of accuracy is $\mu(R) = \frac{||\{\mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_7\}||}{||\{\mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_7\}||} = 0.60$.

- (2) The neighborhoods $\xi(x_i)$ for all $x_i \in U$ of the information system by using a similarity matrix at $\delta > 0.50$ are given by

$$\begin{aligned} \xi\{\mathcal{E}_1\} &= \{\mathcal{E}_1, \mathcal{E}_5\}, & \xi\{\mathcal{E}_2\} &= \{\mathcal{E}_2, \mathcal{E}_3\}, & \xi\{\mathcal{E}_3\} &= \{\mathcal{E}_2, \mathcal{E}_3\}, & \xi\{\mathcal{E}_4\} &= \{\mathcal{E}_4\}, \\ \xi\{\mathcal{E}_5\} &= \{\mathcal{E}_1, \mathcal{E}_5\}, & \xi\{\mathcal{E}_6\} &= \{\mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}, & \xi\{\mathcal{E}_7\} &= \{\mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}, & \xi\{\mathcal{E}_8\} &= \{\mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}. \end{aligned}$$

The interior and closure of $N = \{\mathcal{E}_1, \mathcal{E}_3, \mathcal{E}_6, \mathcal{E}_8\}$ are given by $cl_R(N) = \{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_5, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}$ and $int_R(N) = U - cl[U - N] = U - cl(\{\mathcal{E}_2, \mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_7\}) = U - (\{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}) = \phi$.

Then, the degree of accuracy is $\mu(N) = \frac{||\phi||}{||\{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_5, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}||} = 0$.

The interior and closure of $R = \{\mathcal{E}_2, \mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_7\}$, $cl_R(R) = \{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}$ and $int_R(R) = U - cl[U - R] = U - cl(\{\mathcal{E}_1, \mathcal{E}_3, \mathcal{E}_6, \mathcal{E}_8\}) = U - (\{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_5, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}) = \{\mathcal{E}_4\}$.

Then, the degree of accuracy is $\mu(R) = \frac{||\{\mathcal{E}_4\}||}{||\{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}||} = 0.13$.

Case 2. Topological space on information system

- (1) At $\delta = 1$, $\sigma = \beta = \{\{\mathcal{E}_1\}, \{\mathcal{E}_2, \mathcal{E}_3\}, \{\mathcal{E}_4\}, \{\mathcal{E}_5\}, \{\mathcal{E}_6, \mathcal{E}_8\}, \{\mathcal{E}_7\}\}$ is the base for the quasi-discrete topological space of the information system, then the accuracy measure is $\mu(N) = \frac{|\{\mathcal{E}_1, \mathcal{E}_6, \mathcal{E}_8\}|}{|\{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_5, \mathcal{E}_7\}|} = 0.60$, and $\mu(R) = \frac{|\{\mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_7\}|}{|\{\mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_7\}|} = 0.60$.
- (2) At $\delta > 0.50$, $\xi = \beta = \{\{\mathcal{E}_1, \mathcal{E}_5\}, \{\mathcal{E}_2, \mathcal{E}_3\}, \{\mathcal{E}_4\}, \{\mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}\}$ is the base of the induced topological space of information system. Then, the accuracy measure is $\mu(N) = \frac{|\phi|}{|\{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_5, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}|} = 0$ and $\mu(R) = \frac{|\mathcal{E}_4|}{|\{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}|} = 0.13$.

Case 3. The similarity matrix lattice of information system (analytical method)

In the similarity lattice of the information system, objects with equal similarity scores belong to the same equivalence class under the equivalence relation, although they need not have identical attribute values. The highest object in the lattice has the greatest similarity with other objects for all attributes values of the information table, and the lowest object has the least shared values for the attribute. Each of interiors and closures for decision subsets of a universal set can be determined from the similarity lattice of the information system as follows:

$$\mathcal{P}(x_i) = \frac{\sum_{j=1}^{|U|} \mu(x_i, x_j)}{|U|}. \text{ Then,}$$

$$\mathcal{P}(\mathcal{E}_1) = \frac{\sum_{j=1}^{|U|} \{1+0.50+0.50+0+0.75+0.50+0.50+0.50\}}{8} = 0.53,$$

$$\mathcal{P}(\mathcal{E}_2) = \frac{\sum_{j=1}^{|U|} \{0.50+1+1+0.50+0.25+0.50+0.50+0.50\}}{8} = 0.59,$$

$$\mathcal{P}(\mathcal{E}_3) = \frac{\sum_{j=1}^{|U|} \{0.50+1+1+0.50+0.25+0.50+0.50+0.50\}}{8} = 0.59,$$

$$\mathcal{P}(\mathcal{E}_4) = \frac{\sum_{j=1}^{|U|} \{0+0.50+0.50+1+0.25+0.50+0.25+0.50\}}{8} = 0.44,$$

$$\mathcal{P}(\mathcal{E}_5) = \frac{\sum_{j=1}^{|U|} \{0.75+0.25+0.25+0.25+1+0.25+0.25+0.25\}}{8} = 0.41,$$

$$\mathcal{P}(\mathcal{E}_6) = \frac{\sum_{j=1}^{|U|} \{0.50+0.50+0.50+0.50+0.25+1+0.75+1\}}{8} = 0.63,$$

$$\mathcal{P}(\mathcal{E}_7) = \frac{\sum_{j=1}^{|U|} \{0.50+0.50+0.50+0.25+0.25+0.75+1+0.75\}}{8} = 0.56,$$

$$\mathcal{P}(\mathcal{E}_8) = \frac{\sum_{j=1}^{|U|} \{0.50+0.50+0.50+0.50+0.25+1+0.75+1\}}{8} = 0.63.$$

The binary relation is given by the ordered pairs $(\mathcal{E}_1, 0.53)$, $(\mathcal{E}_2, 0.59)$, $(\mathcal{E}_3, 0.59)$, $(\mathcal{E}_4, 0.44)$, $(\mathcal{E}_5, 0.41)$, $(\mathcal{E}_6, 0.63)$, $(\mathcal{E}_7, 0.56)$, $(\mathcal{E}_8, 0.63)$. The preorder is $\mathcal{E}_5 \leq \mathcal{E}_4 \leq \mathcal{E}_1 \leq \mathcal{E}_7 \leq \mathcal{E}_2 \sim \mathcal{E}_3 \leq \mathcal{E}_6 \sim \mathcal{E}_8$. This can be represented in Table 3. The interior and closure subsets of decision equivalence classes attributes are

- (1) $cl(N) = \{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}$ and $int(N) = \{\mathcal{E}_3, \mathcal{E}_6, \mathcal{E}_8\}$. Then, $\mu(N) = 0.50$.
- (2) $cl(R) = \{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_7\}$ and $int(R) = \{\mathcal{E}_2, \mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_7\}$. Then, $\mu(R) = 0.67$.

Table 3. Lattice of similarity values of information system (a).

	$\mathcal{E}_1$	$\mathcal{E}_2$	$\mathcal{E}_3$	$\mathcal{E}_4$	$\mathcal{E}_5$	$\mathcal{E}_6$	$\mathcal{E}_7$	$\mathcal{E}_8$	ϵ_P
$\mathcal{P}(\mathcal{E}_i)$	0.53	0.59	0.59	0.44	0.41	0.63	0.56	0.63	0.55

In Figure 1, the similarity matrix can be represented by the lattice of information system.

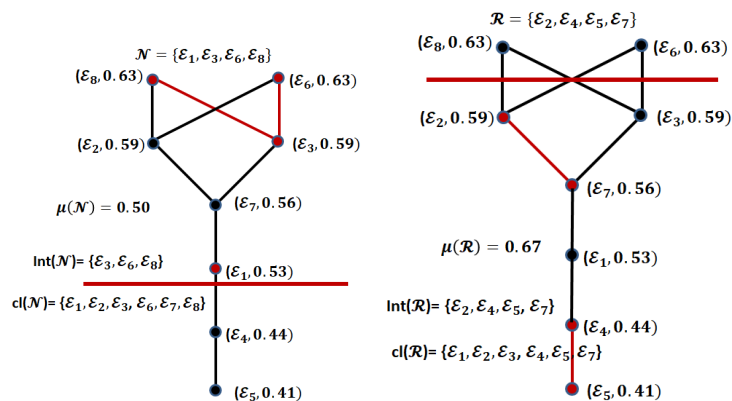


Figure 1. The lattice information system for all attributes.

Remark 3.25. Let $\mathcal{I}_{ST} = (U, C, D, f, v_i \in V)$ be an information system $P : x_i \rightarrow [0, 1]$. Then,

- (1) The accuracy measurement of the similarity density of an object with respect to the similarity matrix is given by $\mathcal{T}(x_i) = \frac{\sum_{j=1}^{|U|} \sum_{k=1}^{|C|} v_{C_k}(x_i, x_j)}{|U||C|} = \mathcal{P}(x_i)$.
- (2) The mean similarity score of an information system with respect to the attribute k is given by $\epsilon_P = \frac{\sum_{i=1}^{|U|} P(x_i)}{|U|}$.

Example 3.26. (Continuing from Example 3.23) From the matrix of similarity values, we have $\mathcal{T}(x_i) = \frac{\sum_{j=1}^{|U|} \sum_{k=1}^{|C|} v_{C_k}(x_i, x_j)}{|U||C|}$,

$$\mathcal{T}(\mathcal{E}_1) = \frac{4+2+2+0+3+2+2+2}{8.4} = 0.53 = \mathcal{P}(\mathcal{E}_1), \quad \mathcal{T}(\mathcal{E}_2) = \frac{2+4+4+2+1+2+2+2}{8.4} = 0.59 = \mathcal{P}(\mathcal{E}_2),$$

$$\mathcal{T}(\mathcal{E}_7) = \frac{2+2+2+1+1+3+4+3}{8.4} = 0.56 = \mathcal{P}(\mathcal{E}_7), \quad \mathcal{T}(\mathcal{E}_8) = \frac{2+2+2+2+1+4+3+4}{8.4} = 0.63 = \mathcal{P}(\mathcal{E}_8).$$

This can be represented in Table 4.

Table 4. Lattice of similarity values of information system (a).

	$\mathcal{E}_1$	$\mathcal{E}_2$	$\mathcal{E}_3$	$\mathcal{E}_4$	$\mathcal{E}_5$	$\mathcal{E}_6$	$\mathcal{E}_7$	$\mathcal{E}_8$	ϵ_P
$\mathcal{T}(\mathcal{E}_i)$	0.53	0.59	0.59	0.44	0.41	0.63	0.56	0.63	0.55

Remark 3.27. Although $T(x_i)$ and $P(x_i)$ may have the same numerical value, they are obtained through different aggregation procedures. The quantity $P(x_i)$ is calculated from the pairwise similarity measure $\mu(x_i, x_j)$, whereas $T(x_i)$ is obtained by aggregating the corresponding similarity values over the conditional attributes. Therefore, they are denoted by different symbols to distinguish their respective computational procedures.

Remark 3.28. The threshold $\delta = 0.5$ is selected in the numerical example for illustrative purposes. Specifically, two objects are considered sufficiently similar to belong to the same neighborhood whenever their similarity degree is greater than or equal to 0.5. Thus, this threshold represents a moderate level of similarity, excluding object pairs with relatively low similarity while retaining pairs with at least 50% similarity. A higher threshold yields smaller and more refined neighborhoods,

whereas a lower threshold produces larger and more inclusive neighborhoods. Therefore, other threshold values may be selected depending on the application domain and the desired level of granularity.

The results of Example 3.24 can be collected in Table 5.

Table 5. Evaluation of the information system.

Decision attributes	Topological space		Pretopological space		Analytical	
	Normal	Rest	Normal	Rest	Normal	Rest
$\delta = 1$	0.60	0.60	0.60	0.60	0.50	0.66
$\delta > 0.50$	0	0.13	0	0.13		

4. Reduction of information system attributes

Some data in an information system may be unnecessary or redundant and can be removed without significantly affecting the information available. Attribute reduction can be viewed as the process of identifying a minimally descriptive subset of attributes while preserving the same, or an approximately equivalent, classification of the objects. The term data reduction refers to the elimination of superfluous attributes from an information system so that the remaining attributes can be used to preserve or define its classification. Consequently, all redundant attributes can be removed, leaving only the attributes that are essential for representing the relevant information. Attribute reduction aims to identify a minimal subset of conditional attributes that preserves a specified evaluation criterion. In this study, the effect of deleting each attribute is assessed using the similarity-based structures introduced in Section 4.

4.1. The similarity matrix of the information system excluding the attribute Trans. Line A

Example 4.1. (Continuing from Example 3.23) Let $\mathcal{I}_{\mathcal{ST}} = (U, C, D, g, v_i \in V)$ be an information system, where $U = \{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}$, $|U| = 8$, and $|C| = 3$ (Trans. Line 2 = $\mathcal{B}$, Self-Gem = C and Self-load = $\mathcal{L}$). Construct the similarity matrix as in Table 6.

Table 6. Similarity matrix values of the information system excluding Trans. Line A attribute.

$C = \{\mathcal{B}, C, \mathcal{L}\}$	$\mathcal{E}_1$	$\mathcal{E}_2$	$\mathcal{E}_3$	$\mathcal{E}_4$	$\mathcal{E}_5$	$\mathcal{E}_6$	$\mathcal{E}_7$	$\mathcal{E}_8$	$\mathcal{P}(\mathcal{E}_i)$
$\mathcal{E}_1$	1	0.33	0.33	0	1	0.33	0.33	0.33	0.46
$\mathcal{E}_2$	0.33	1	1	0.67	0.33	0.33	0.33	0.33	0.54
$\mathcal{E}_3$	0.33	1	1	0.67	0.33	0.33	0.33	0.33	0.54
$\mathcal{E}_4$	0	0.67	0.67	1	0	0.67	0.33	0.67	0.501
$\mathcal{E}_5$	1	0.33	0.33	0	1	0.33	0.33	0.33	0.46
$\mathcal{E}_6$	0.33	0.33	0.33	0.67	0.33	1	0.67	1	0.58
$\mathcal{E}_7$	0.33	0.33	0.33	0.33	0.33	0.67	1	0.67	0.498
$\mathcal{E}_8$	0.33	0.33	0.33	0.67	0.33	1	0.67	1	0.58
Mean similarity score									0.52

Case 1. Pretopological space of information system excluding Trans. Line attribute

- (1) The neighborhoods $\sigma(x_i)$ for all $x_i \in U$ of the information system by using the similarity matrix at $\delta = 1$ are given by

$$\begin{aligned}\sigma\{\mathcal{E}_1\} &= \{\mathcal{E}_1, \mathcal{E}_5\}, & \sigma\{\mathcal{E}_2\} &= \{\mathcal{E}_2, \mathcal{E}_3\}, & \sigma\{\mathcal{E}_3\} &= \{\mathcal{E}_2, \mathcal{E}_3\}, & \sigma\{\mathcal{E}_4\} &= \{\mathcal{E}_4\}, \\ \sigma\{\mathcal{E}_5\} &= \{\mathcal{E}_1, \mathcal{E}_5\}, & \sigma\{\mathcal{E}_6\} &= \{\mathcal{E}_6, \mathcal{E}_8\}, & \sigma\{\mathcal{E}_7\} &= \{\mathcal{E}_7\}, & \sigma\{\mathcal{E}_8\} &= \{\mathcal{E}_6, \mathcal{E}_8\}.\end{aligned}$$

Then, the interior and closure of $N = \{\mathcal{E}_1, \mathcal{E}_3, \mathcal{E}_6, \mathcal{E}_8\}$ are given by $cl_Q(N) = \{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_5, \mathcal{E}_6, \mathcal{E}_8\}$, $int_Q(N) = \{\mathcal{E}_6, \mathcal{E}_8\}$, and the degree of accuracy is $\mu(N) = \frac{||\mathcal{E}_6, \mathcal{E}_8||}{||\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_5, \mathcal{E}_6, \mathcal{E}_8||} = 0.33$. The interior and closure of $R = \{\mathcal{E}_2, \mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_7\}$ are given by $cl_Q(R) = \{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_7\}$, $int_Q(R) = \{\mathcal{E}_4, \mathcal{E}_7\}$, and the degree of accuracy is $\mu(R) = \frac{||\mathcal{E}_4, \mathcal{E}_7||}{||\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_7||} = 0.33$.

- (2) The neighborhoods $\xi(x_i)$ for all $x_i \in U$ of the information system by using the similarity matrix at $\delta > 0.50$ are given by

$$\begin{aligned}\xi\{\mathcal{E}_1\} &= \{\mathcal{E}_1, \mathcal{E}_5\}, & \xi\{\mathcal{E}_2\} &= \{\mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4\}, & \xi\{\mathcal{E}_3\} &= \{\mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4\}, & \xi\{\mathcal{E}_4\} &= \{\mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_6, \mathcal{E}_8\}, \\ \xi\{\mathcal{E}_5\} &= \{\mathcal{E}_1, \mathcal{E}_5\}, & \xi\{\mathcal{E}_6\} &= \{\mathcal{E}_4, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}, & \xi\{\mathcal{E}_7\} &= \{\mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}, & \xi\{\mathcal{E}_8\} &= \{\mathcal{E}_4, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}.\end{aligned}$$

Then, the interior and closure of $N = \{\mathcal{E}_1, \mathcal{E}_3, \mathcal{E}_6, \mathcal{E}_8\}$ are given by $cl_R(N) = \{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}$, $int_R(N) = \phi$, and the degree of accuracy is $\mu(N) = \frac{|\phi|}{||\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8||} = 0$.

The interior and closure of $R = \{\mathcal{E}_2, \mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_7\}$ is given by $cl_R(R) = \{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}$, $int_R(R) = \phi$, and the degree of accuracy is $\mu(R) = \frac{|\phi|}{||\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8||} = 0$.

Case 2. Topological space on the information system excluding Trans. Line A attribute

- (1) The base for a quasi-discrete topological space method of the information system is by using the similarity matrix at $\delta = 1$ given as $\sigma_{\mathcal{A}} = \beta_{\sigma_{\mathcal{A}}} = \{\{\mathcal{E}_1, \mathcal{E}_5\}, \{\mathcal{E}_2, \mathcal{E}_3\}, \{\mathcal{E}_4\}, \{\mathcal{E}_6, \mathcal{E}_8\}, \{\mathcal{E}_7\}\}$. Then the accuracy measure is $\mu(N) = \frac{||\mathcal{E}_6, \mathcal{E}_8||}{||\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_5, \mathcal{E}_6, \mathcal{E}_8||} = 0.33$ and $\mu(R) = \frac{||\mathcal{E}_4, \mathcal{E}_7||}{||\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_7||} = 0.33$.

- (2) The subbase for the topological space method of the information system is by using the similarity matrix at $\delta > 0.50$ given as $\xi_{\mathcal{A}} = \{\{\mathcal{E}_1, \mathcal{E}_5\}, \{\mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4\}, \{\mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_6, \mathcal{E}_8\}, \{\mathcal{E}_4, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}, \{\mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}\}$, then $\beta_{\xi_{\mathcal{A}}} = \{\{\mathcal{E}_1, \mathcal{E}_5\}, \{\mathcal{E}_4\}, \{\mathcal{E}_6, \mathcal{E}_8\}, \{\mathcal{E}_4, \mathcal{E}_6, \mathcal{E}_8\}, \{\mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4\}, \{\mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_6, \mathcal{E}_8\}, \{\mathcal{E}_4, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}, \{\mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}\}$. Then, accuracy measure is $\mu(N) = \frac{||\mathcal{E}_6, \mathcal{E}_8||}{||\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8||} = 0.25$ and $\mu(R) = \frac{||\mathcal{E}_4||}{||\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8||} = 0.13$.

Case 3. The similarity matrix lattice of the information system excluding Trans. Line A attribute (analytical method)

Table 7 can be used to have the lattice of similarity values excluding Trans. Line A attribute.

Table 7. Lattice of similarity values excluding Trans. Line A attribute.

	$\mathcal{E}_1$	$\mathcal{E}_2$	$\mathcal{E}_3$	$\mathcal{E}_4$	$\mathcal{E}_5$	$\mathcal{E}_6$	$\mathcal{E}_7$	$\mathcal{E}_8$	ϵ_P
$\mathcal{P}(\mathcal{E}_i)$	0.46	0.54	0.54	0.50	0.46	0.58	0.50	0.58	0.52

In Figure 2, we give the lattice of the information system without the attributes A.

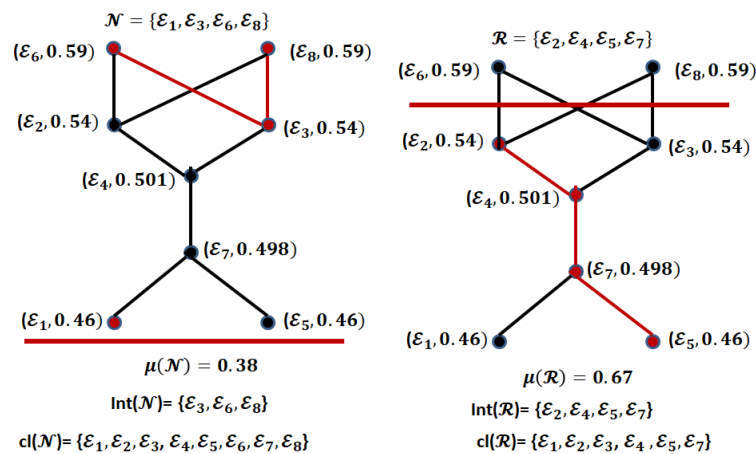


Figure 2. The lattice of the information system excluding Trans. Line A attribute.

Table 8 can be used to have the values of accuracy measure for information system excluding the attribute A.

Table 8. Evaluation of the information system excluding Trans. Line A attribute.

	Topological space		Pretopological space		Analytical	
Decision attributes	Normal	Rest	Normal	Rest	Normal	Rest
$\delta = 1$	0.33	0.33	0.33	0.33	0.38	0.67
$\delta > 0.50$	0.25	0.13	0	0.0		

4.2. The similarity matrix of the information system excluding the attribute Trans. Line B

Example 4.2. (Continuing from Example 3.23) Let $\mathcal{I}_{ST} = (U, C, D, g, v_i \in V)$ be an information system, where $U = \{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}$, $|U| = 8$, and $|C| = 3$ (Trans. Line 1 = $\mathcal{A}$, Self-Gem = C and Self-load = $\mathcal{L}$). Construct the similarity matrix as in Table 9.

Table 9. Similarity matrix values of the information system excluding Trans. Line B attribute.

$C = \{\mathcal{A}, C, \mathcal{L}\}$	$\mathcal{E}_1$	$\mathcal{E}_2$	$\mathcal{E}_3$	$\mathcal{E}_4$	$\mathcal{E}_5$	$\mathcal{E}_6$	$\mathcal{E}_7$	$\mathcal{E}_8$	$\mathcal{P}(\mathcal{E}_i)$
$\mathcal{E}_1$	1.00	0.67	0.67	0.00	0.67	0.33	0.33	0.33	0.5
$\mathcal{E}_2$	0.67	1.00	1.00	0.33	0.33	0.67	0.67	0.67	0.67
$\mathcal{E}_3$	0.67	1.00	1.00	0.33	0.33	0.67	0.67	0.67	0.67
$\mathcal{E}_4$	0.00	0.33	0.33	1.00	0.33	0.67	0.33	0.67	0.46
$\mathcal{E}_5$	0.67	0.33	0.33	0.33	1.00	0.00	0.00	0.00	0.33
$\mathcal{E}_6$	0.33	0.67	0.67	0.67	0.00	1.00	0.67	1.00	0.63
$\mathcal{E}_7$	0.33	0.67	0.67	0.33	0.00	0.67	1.00	0.67	0.54
$\mathcal{E}_8$	0.33	0.67	0.67	0.67	0.00	1.00	0.67	1.00	0.63
Mean similarity score									0.55

Case 1. Pretopological space of the information system excluding Trans. Line B attribute

- (1) The neighborhoods $\sigma(x_i)$ for all $x_i \in U$ of the information system by using the similarity matrix at $\delta = 1$ are given by
 $\sigma\{\mathcal{E}_1\} = \{\mathcal{E}_1\}$, $\sigma\{\mathcal{E}_2\} = \{\mathcal{E}_2, \mathcal{E}_3\}$, $\sigma\{\mathcal{E}_3\} = \{\mathcal{E}_2, \mathcal{E}_3\}$, $\sigma\{\mathcal{E}_4\} = \{\mathcal{E}_4\}$,
 $\sigma\{\mathcal{E}_5\} = \{\mathcal{E}_5\}$, $\sigma\{\mathcal{E}_6\} = \{\mathcal{E}_6, \mathcal{E}_8\}$, $\sigma\{\mathcal{E}_7\} = \{\mathcal{E}_7\}$, $\sigma\{\mathcal{E}_8\} = \{\mathcal{E}_6, \mathcal{E}_8\}$. Then, the accuracy measure is $\mu(N) = \frac{||\mathcal{E}_1, \mathcal{E}_6, \mathcal{E}_8||}{||\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_6, \mathcal{E}_8||} = 0.60$ and $\mu(R) = \frac{||\mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_7||}{||\mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_7||} = 0.60$.
- (2) The neighborhoods $\xi(x_i)$ for all $x_i \in U$ of the information system by using the similarity matrix at $\delta > 0.50$ are given by
 $\xi\{\mathcal{E}_1\} = \{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_5\}$, $\xi\{\mathcal{E}_2\} = \{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}$, $\xi\{\mathcal{E}_3\} = \{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}$,
 $\xi\{\mathcal{E}_4\} = \{\mathcal{E}_4, \mathcal{E}_6, \mathcal{E}_8\}$, $\xi\{\mathcal{E}_5\} = \{\mathcal{E}_1, \mathcal{E}_5\}$, $\xi\{\mathcal{E}_6\} = \{\mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}$,
 $\xi\{\mathcal{E}_7\} = \{\mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}$, $\xi\{\mathcal{E}_8\} = \{\mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}$. Then, the accuracy measure is
 $\mu(N) = \frac{|\phi|}{||\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8||} = 0$ and $\mu(R) = \frac{|\phi|}{||\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8||} = 0$.

Case 2. Topological space on the information system excluding Trans. Line B attribute

- (1) The base for the quasi-discrete topological space method of the information system is by using the similarity matrix at $\delta = 1$ given as $\sigma_{\mathcal{B}} = \beta_{\sigma_{\mathcal{B}}} = \{\{\mathcal{E}_1\}, \{\mathcal{E}_2, \mathcal{E}_3\}, \{\mathcal{E}_4\}, \{\mathcal{E}_5\}, \{\mathcal{E}_6, \mathcal{E}_8\}, \{\mathcal{E}_7\}\}$. Then, the measure of accuracy is $\mu(N) = \frac{||\mathcal{E}_1, \mathcal{E}_6, \mathcal{E}_8||}{||\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_6, \mathcal{E}_8||} = 0.60$ and $\mu(R) = \frac{||\mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_7||}{||\mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_7||} = 0.60$.
- (2) The subbase for the topological space method of the information system is by using the similarity matrix at $\delta > 0.50$ given as $\xi_{\mathcal{B}} = \{\{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_5\}, \{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}, \{\mathcal{E}_4, \mathcal{E}_6, \mathcal{E}_8\}, \{\mathcal{E}_1, \mathcal{E}_5\}, \{\mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}\}$, then $\beta_{\xi_{\mathcal{B}}} = \{\{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_5\}, \{\mathcal{E}_1\}, \{\mathcal{E}_2, \mathcal{E}_3\}, \{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3\}, \{\mathcal{E}_6, \mathcal{E}_8\}, \{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}, \{\mathcal{E}_4, \mathcal{E}_6, \mathcal{E}_8\}, \{\mathcal{E}_1, \mathcal{E}_5\}, \{\mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}\}$. Then, the accuracy measure is $\mu(N) = \frac{||\mathcal{E}_1, \mathcal{E}_6, \mathcal{E}_8||}{||\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8||} = 0.38$ and $\mu(R) = \frac{|\phi|}{||\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8||} = 0$.

Case 3. The similarity matrix lattice of the information system excluding Trans. Line B attribute (analytical method)

Table 10 can be used to have the lattice of similarity values excluding attribute B.

Table 10. Lattice of similarity values excluding Trans. Line B attribute.

	$\mathcal{E}_1$	$\mathcal{E}_2$	$\mathcal{E}_3$	$\mathcal{E}_4$	$\mathcal{E}_5$	$\mathcal{E}_6$	$\mathcal{E}_7$	$\mathcal{E}_8$	ϵ_P
$\mathcal{P}(\mathcal{E}_i)$	0.50	0.67	0.67	0.46	0.33	0.63	0.54	0.63	0.55

In Figure 3, we give the lattice of the information system without the attributes B.

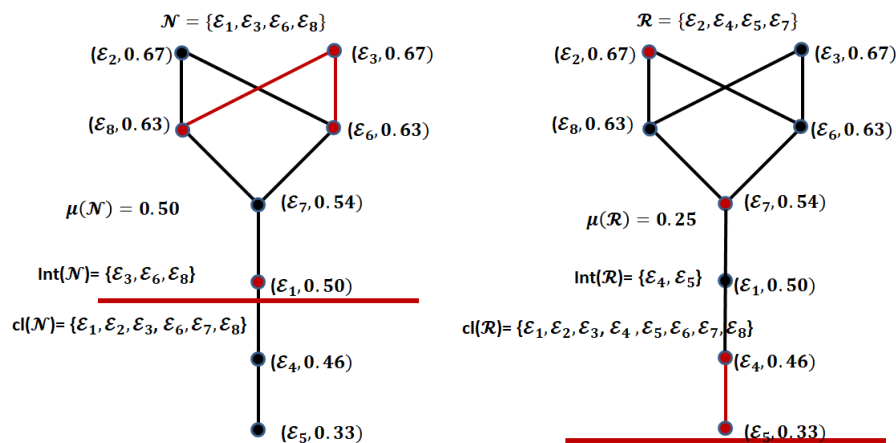


Figure 3. The lattice of the information system excluding Trans. Line B attribute.

Table 11 can be used to have the values of accuracy measure for information system excluding the attribute B.

Table 11. Evaluation of the information system excluding Trans. Line B attribute.

	Topological space		Pretopological space		Analytical	
	Normal	Rest	Normal	Rest	Normal	Rest
Decision attributes						
$\delta = 1$	0.60	0.60	0.60	0.60	0.50	0.25
$\delta > 0.50$	0.38	0	0	0		

4.3. The similarity matrix of the information system excluding the attribute Self-Gem C

Example 4.3. (Continuing from Example 3.23) Let $I_{ST} = (U, C, D, g, v_i \in V)$ be an information system, where $U = \{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}$, $|U| = 8$, and $|C| = 3$ (Trans. Line 1 = $\mathcal{A}$, Trans. Line 2 = $\mathcal{B}$ and Self-load = $\mathcal{L}$). Construct the similarity matrix as in Table 12.

Table 12. Similarity matrix values of the information system excluding Self-Gem C attribute.

$C = \{\mathcal{A}, \mathcal{B}, \mathcal{L}\}$	$\mathcal{E}_1$	$\mathcal{E}_2$	$\mathcal{E}_3$	$\mathcal{E}_4$	$\mathcal{E}_5$	$\mathcal{E}_6$	$\mathcal{E}_7$	$\mathcal{E}_8$	$\mathcal{P}(\mathcal{E}_i)$
$\mathcal{E}_1$	1.00	0.67	0.67	0.00	0.67	0.67	0.67	0.67	0.63
$\mathcal{E}_2$	0.67	1.00	1.00	0.33	0.33	0.33	0.33	0.33	0.54
$\mathcal{E}_3$	0.67	1.00	1.00	0.33	0.33	0.33	0.33	0.33	0.54
$\mathcal{E}_4$	0.00	0.33	0.33	1.00	0.33	0.33	0.00	0.33	0.33
$\mathcal{E}_5$	0.67	0.33	0.33	0.33	1.00	0.33	0.33	0.33	0.46
$\mathcal{E}_6$	0.67	0.33	0.33	0.33	0.33	1.00	0.67	1.00	0.58
$\mathcal{E}_7$	0.67	0.33	0.33	0.00	0.33	0.67	1.00	0.67	0.5
$\mathcal{E}_8$	0.67	0.33	0.33	0.33	0.33	1.00	0.67	1.00	0.58
Mean similarity score									0.52

Case 1. Pretopological space of the information system excluding Self-Gem C attribute

- (1) The neighborhoods $\sigma(x_i)$ for all $x_i \in U$ of the information system by using the similarity matrix at $\delta = 1$ are given by
 $\sigma\{\mathcal{E}_1\} = \{\mathcal{E}_1\}$, $\sigma\{\mathcal{E}_2\} = \{\mathcal{E}_2, \mathcal{E}_3\}$, $\sigma\{\mathcal{E}_3\} = \{\mathcal{E}_2, \mathcal{E}_3\}$, $\sigma\{\mathcal{E}_4\} = \{\mathcal{E}_4\}$, $\sigma\{\mathcal{E}_5\} = \{\mathcal{E}_5\}$,
 $\sigma\{\mathcal{E}_6\} = \{\mathcal{E}_6, \mathcal{E}_8\}$, $\sigma\{\mathcal{E}_7\} = \{\mathcal{E}_7\}$, $\sigma\{\mathcal{E}_8\} = \{\mathcal{E}_6, \mathcal{E}_8\}$. Then the accuracy measure is $\mu(N) = \frac{||\{\mathcal{E}_1, \mathcal{E}_6, \mathcal{E}_8\}||}{||\{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_6, \mathcal{E}_8\}||} = 0.60$ and $\mu(R) = \frac{||\{\mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_7\}||}{||\{\mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_7\}||} = 0.60$.
- (2) The neighborhoods $\xi(x_i)$ for all $x_i \in U$ of the information system by using the similarity matrix at $\delta > 0.50$ are given by
 $\xi\{\mathcal{E}_1\} = \{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_5, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}$, $\xi\{\mathcal{E}_2\} = \{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3\}$, $\xi\{\mathcal{E}_3\} = \{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3\}$, $\xi\{\mathcal{E}_4\} = \{\mathcal{E}_4\}$,
 $\xi\{\mathcal{E}_5\} = \{\mathcal{E}_1, \mathcal{E}_5\}$, $\xi\{\mathcal{E}_6\} = \{\mathcal{E}_1, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}$, $\xi\{\mathcal{E}_7\} = \{\mathcal{E}_1, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}$, $\xi\{\mathcal{E}_8\} = \{\mathcal{E}_1, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}$.
Then, the accuracy measure is $\mu(N) = \frac{||\phi||}{||\{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_5, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}||} = 0$ and $\mu(R) = \frac{||\{\mathcal{E}_4\}||}{||\{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}||} = 0.13$.

Case 2. Topological space on the information system excluding Self-Gem C attribute

- (1) The base for the quasi-discrete topological space method of the information system is by using the similarity matrix at $\delta = 1$ given as $\sigma_C = \beta_{\sigma_C} = \{\{\mathcal{E}_1\}, \{\mathcal{E}_2, \mathcal{E}_3\}, \{\mathcal{E}_4\}, \{\mathcal{E}_5\}, \{\mathcal{E}_6, \mathcal{E}_8\}, \{\mathcal{E}_7\}\}$. Then, the accuracy measure is $\mu(N) = \frac{||\{\mathcal{E}_1, \mathcal{E}_6, \mathcal{E}_8\}||}{||\{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_6, \mathcal{E}_8\}||} = 0.60$ and $\mu(R) = \frac{||\{\mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_7\}||}{||\{\mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_7\}||} = 0.60$.
- (2) The subbase for the topological space method of the information system is by using the similarity matrix at $\delta > 0.50$ given as $\xi_C = \{\{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_5, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}, \{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3\}, \{\mathcal{E}_4\}, \{\mathcal{E}_1, \mathcal{E}_5\}, \{\mathcal{E}_1, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}\}$. Then, $\beta_{\xi_C} = \{\{\mathcal{E}_1\}, \{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_5, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}, \{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3\}, \{\mathcal{E}_4\}, \{\mathcal{E}_1, \mathcal{E}_5\}, \{\mathcal{E}_1, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}\}$. Then, the accuracy measure is $\mu(N) = \frac{||\{\mathcal{E}_1\}||}{||\{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_5, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}||} = 0.14$ and $\mu(R) = \frac{||\{\mathcal{E}_4\}||}{||\{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}||} = 0.13$.

Case 3. The similarity matrix lattice of the information system excluding Self-Gem C attribute (analytical method)

Table 13 is used to obtain the lattice of similarity values excluding attribute C.

Table 13. Lattice of similarity values excluding Self-Gem C attribute.

	$\mathcal{E}_1$	$\mathcal{E}_2$	$\mathcal{E}_3$	$\mathcal{E}_4$	$\mathcal{E}_5$	$\mathcal{E}_6$	$\mathcal{E}_7$	$\mathcal{E}_8$	ϵ_P
$\mathcal{P}(\mathcal{E}_i)$	0.63	0.54	0.54	0.33	0.46	0.58	0.50	0.58	0.52

In Figure 4, we give the lattice of the information system without the attributes C.

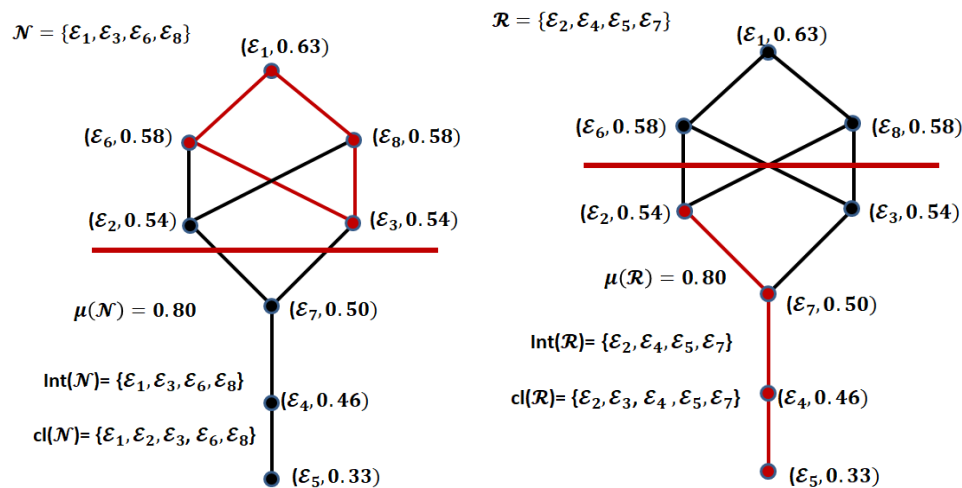


Figure 4. The lattice of the information system excluding Self-Gem C attribute.

We use Table 14 to evaluate the information system excluding the attribute C.

Table 14. Evaluation of the information system excluding Self-Gem C attribute.

Decision attributes	Topological space		Pretopological space		Analytical	
	Normal	Rest	Normal	Rest	Normal	Rest
$\delta = 1$	0.60	0.60	0.60	0.60	0.80	0.80
$\delta > 0.50$	0.14	0.13	0	0.13		

4.4. The similarity matrix of the information system excluding the attribute Self-Load L

Example 4.4. (Continuing from Example 3.23) Let $\mathcal{I}_{ST} = (U, C, D, g, v_i \in V)$ be an information system, where $U = \{\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4, \varepsilon_5, \varepsilon_6, \varepsilon_7, \varepsilon_8\}$, $|U| = 8$, and $|C| = 3$ (Trans. Line 1 = $\mathcal{A}$, Trans. Line 2 = $\mathcal{B}$ and Self-Gem = C). Construct the similarity matrix as in Table 15.

Table 15. Similarity matrix values of the information system excluding Self-Load L attribute.

$C = \{\mathcal{A}, \mathcal{B}, C\}$	ε_1	ε_2	ε_3	ε_4	ε_5	ε_6	ε_7	ε_8	$\mathcal{P}(\varepsilon_i)$
ε_1	1.00	0.33	0.33	0.00	0.67	0.67	0.67	0.67	0.54
ε_2	0.33	1.00	1.00	0.67	0.00	0.67	0.67	0.67	0.63
ε_3	0.33	1.00	1.00	0.67	0.00	0.67	0.67	0.67	0.63
ε_4	0.00	0.67	0.67	1.00	0.33	0.33	0.33	0.33	0.46
ε_5	0.67	0.00	0.00	0.33	1.00	0.33	0.33	0.33	0.38
ε_6	0.67	0.67	0.67	0.33	0.33	1.00	1.00	1.00	0.71
ε_7	0.67	0.67	0.67	0.33	0.33	1.00	1.00	1.00	0.71
ε_8	0.67	0.67	0.67	0.33	0.33	1.00	1.00	1.00	0.71
Mean similarity score									0.59

Case 1. Pretopological space of the information system excluding Self-Load L attribute

- (1) The neighborhoods $\sigma(x_i)$ for all $x_i \in U$ of the information system by using the similarity matrix at $\delta = 1$ are given by

$$\sigma\{\mathcal{E}_1\} = \{\mathcal{E}_1\}, \quad \sigma\{\mathcal{E}_2\} = \{\mathcal{E}_2, \mathcal{E}_3\}, \quad \sigma\{\mathcal{E}_3\} = \{\mathcal{E}_2, \mathcal{E}_3\}, \quad \sigma\{\mathcal{E}_4\} = \{\mathcal{E}_4\}, \quad \sigma\{\mathcal{E}_5\} = \{\mathcal{E}_5\}, \\ \sigma\{\mathcal{E}_6\} = \{\mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}, \quad \sigma\{\mathcal{E}_7\} = \{\mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}, \quad \sigma\{\mathcal{E}_8\} = \{\mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}. \text{ Then, the accuracy measure is } \mu(N) = \frac{||\mathcal{E}_1||}{||\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8||} = \frac{1}{6} = 0.17 \text{ and } \mu(R) = \frac{||\mathcal{E}_4, \mathcal{E}_5||}{||\mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8||} = 0.29.$$

- (2) The neighborhoods $\xi(x_i)$ for all $x_i \in U$ of the information system by using the similarity matrix at $\delta > 0.50$ are given by

$$\xi\{\mathcal{E}_1\} = \{\mathcal{E}_1, \mathcal{E}_5, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}, \quad \xi\{\mathcal{E}_2\} = \{\mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}, \quad \xi\{\mathcal{E}_3\} = \{\mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}, \\ \xi\{\mathcal{E}_4\} = \{\mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4\}, \quad \xi\{\mathcal{E}_5\} = \{\mathcal{E}_1, \mathcal{E}_5\}, \quad \xi\{\mathcal{E}_6\} = \{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}, \quad \xi\{\mathcal{E}_7\} = \{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}, \\ \xi\{\mathcal{E}_8\} = \{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}. \text{ Then, the accuracy measure is } \mu(N) = \frac{||\phi||}{||\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8||} = 0 \text{ and } \mu(R) = \frac{||\phi||}{||\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8||} = 0.$$

Case 2. Topological space on the information system excluding Self-Load L attribute

- (1) The base for the quasi-discrete topological space method of the information system by using the similarity matrix at $\delta = 1$ is given as $\sigma_{\mathcal{L}} = \beta_{\sigma_{\mathcal{L}}} = \{\{\mathcal{E}_1\}, \{\mathcal{E}_2, \mathcal{E}_3\}, \{\mathcal{E}_4\}, \{\mathcal{E}_5\}, \{\mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}\}$. Then, the accuracy measure is $\mu(N) = \frac{||\mathcal{E}_1||}{||\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8||} = 0.17$ and $\mu(R) = \frac{||\mathcal{E}_4, \mathcal{E}_5||}{||\mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8||} = 0.29$.

- (2) The subbase for the topological space method of the information system by using the similarity matrix at $\delta > 0.50$ is given as $\xi_{\mathcal{L}} = \{\{\mathcal{E}_1, \mathcal{E}_5, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}, \{\mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}, \{\mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4\}, \{\mathcal{E}_1, \mathcal{E}_5\}, \{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}\}$. Then, $\beta_{\xi_{\mathcal{L}}} = \{\{\mathcal{E}_1\}, \{\mathcal{E}_2, \mathcal{E}_3\}, \{\mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}, \{\mathcal{E}_1, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}, \{\mathcal{E}_1, \mathcal{E}_5, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}, \{\mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}, \{\mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4\}, \{\mathcal{E}_1, \mathcal{E}_5\}, \{\mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}, \{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8\}\}$. Then, the accuracy measure is $\mu(N) = \frac{||\mathcal{E}_4||}{||\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8||} = 0.13$ and $\mu(R) = \frac{||\phi||}{||\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_5, \mathcal{E}_6, \mathcal{E}_7, \mathcal{E}_8||} = 0$.

Case 3. The similarity matrix lattice of the information system excluding Self-Load L attribute (analytical method)

Table 16 is used to get the lattice of similarity values excluding attribute L.

Table 16. Lattice of similarity values excluding Self-Load L attribute.

	$\mathcal{E}_1$	$\mathcal{E}_2$	$\mathcal{E}_3$	$\mathcal{E}_4$	$\mathcal{E}_5$	$\mathcal{E}_6$	$\mathcal{E}_7$	$\mathcal{E}_8$	ϵ_P
$\mathcal{P}(\mathcal{E}_i)$	0.54	0.63	0.63	0.46	0.38	0.71	0.71	0.71	0.59

In Figure 5, we give the lattice of the information system without the attributes L.

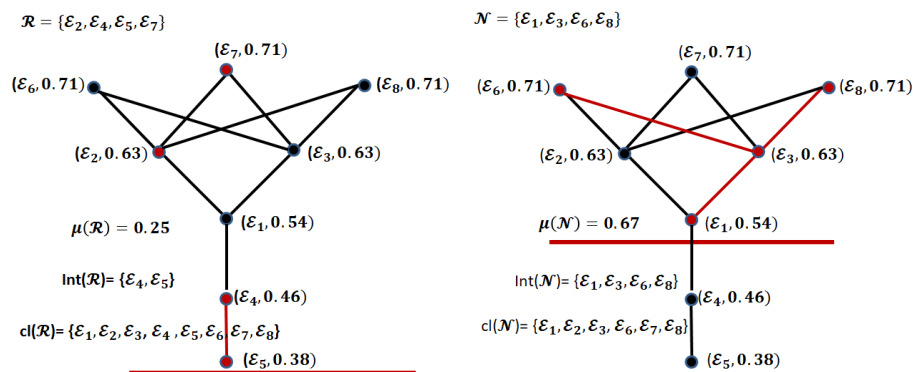


Figure 5. The lattice of the information system excluding Self-Load L attribute.

We use Table 17 to evaluate the information system excluding the attribute L.

Table 17. Evaluation of the information system excluding Self Load L attribute.

Decision attributes	Topological space		Pretopological space		Analytical	
	Normal	Rest	Normal	Rest	Normal	Rest
$\delta = 1$	0.17	0.29	0.17	0.29	0.80	0.80
$\delta > 0.50$	0.13	0	0	0.0		

5. Results

- (1) For $\delta = 1$, the Pawlak approximation space, the induced pretopology, the corresponding topological space, and the space of lattice atoms are equivalent in the sense that they induce the same partition of the universe. Therefore, in the following table, we use the Pawlak approximation space to represent the case $\delta = 1$.
- (2) For $\delta > 0.50$, the induced pretopology and the corresponding topological space are distinct and generated by their respective subbases. Therefore, we present the two topologies in separate columns in the following table under $\delta > 0.50$.
- (3) An analytical representation is proposed by ordering the equivalence classes of the approximation space according to a score-based similarity measure to construct the lattice induced by the information system. This representation highlights the differences among objects sharing similar score values and facilitates the identification of the maximal and minimal lattice atoms. We present the analytical representation in a single column.
- (4) Table 18 presents the values of the evaluation indices obtained after successively removing each attribute from the original conditional attribute set $C = \{A, B, C, L\}$. Each row corresponds to the deletion of one conditional attribute, while each entry gives the resulting value of the corresponding evaluation index.

Table 18. Complete evaluation of the information system.

	$\delta = 1$		$\delta > 0.50$		$\delta > 0.50$		Analytical	
	Pawlak		Topology		Pretopology		Lattice	
Decision attributes	Normal	Rest	Normal	Rest	Normal	Rest	Normal	Rest
IS	0.60	0.60	0	0.13	0	0.13	0.50	0.67
$\mathcal{A}^-$	0.33	0.33	0.25	0.13	0	0	0.38	0.67
$\mathcal{B}^-$	0.60	0.60	0.38	0	0	0	0.50	0.25
$\mathcal{C}^-$	0.60	0.60	0.14	0.13	0	0.13	0.80	0.80
$\mathcal{L}^-$	0.17	0.29	0.13	0	0	0	0.80	0.80

An attribute is considered more important when its removal produces a larger variation in the evaluation index. In Algorithm 5.7, we give procedures to determine the core of the reduction of attribute value table. Let $\text{Acc}_Q(A)$, $\text{Acc}_R(A)$, and $\text{Acc}_L(A)$ denote the accuracy measures obtained at $\delta = 1$, $\delta > 0.50$ and by the analytical lattice approach, respectively, using the attribute subset A . For an attribute $a \in C$, we define its reduction impact with respect to each approach $M \in Q, R, L$ as

$$\Delta_M(a) = \frac{\text{Acc}_M(C)}{\text{Acc}_M(C \setminus \{a\})}.$$

The attribute a is considered dispensable if

$$\Delta_M(a) \leq 1 - \varepsilon, \quad \forall M \in \{Q, R, L\}, \quad \varepsilon \in (0, 1).$$

Otherwise, a is considered indispensable.

Remark 5.1. The tolerance ε controls the sensitivity of the attribute-reduction procedure. Under the proposed criterion, a larger value of ε imposes a stricter requirement for declaring an attribute dispensable, since a greater improvement in the accuracy measure is required after removing the attribute. Conversely, a smaller value of ε permits the removal of attributes that produce relatively smaller improvements in the evaluation measures. In practical applications, ε may be selected through validation on representative data or sensitivity analysis. For an illustrative study, a fixed small tolerance may be adopted, and its influence can be examined over a range of values.

Definition 5.2. Let C be the set of conditional attributes. A subset A of C is called a *reduct* if it preserves the evaluation capability of the full attribute set, namely, $\text{Eval}(A) = \text{Eval}(C)$, where $\text{Eval}(\cdot)$ denotes the adopted evaluation measure. A reduct A is *minimal* if no proper subset of A satisfies the above condition.

Definition 5.3. Let $A \subseteq C$. The subset A is called a *topological-lattice reduct* if

$$\text{Acc}_M(A) > \text{Acc}_M(C), \quad \forall M \in \{Q, R, L\},$$

and A is minimal with respect to this property; that is,

$$\forall a \in A, \quad \exists M \in \{Q, R, L\} \text{ such that } \text{Acc}_M(A \setminus \{a\}) \leq \text{Acc}_M(C).$$

Thus, A is a minimal subset whose removal of any attribute prevents the accuracy from being strictly greater than that obtained using the complete attribute set C in all three evaluation approaches.

Remark 5.4. *The reduct defined in this study should be distinguished from the classical Pawlak rough-set reduct. In Pawlak rough-set theory, a reduct is a minimal subset of conditional attributes that preserves the classification capability of the complete conditional attribute set with respect to the indiscernibility relation. In contrast, the reduct proposed here is determined with respect to the adopted evaluation measures derived from similarity-based topological, pretopological, and analytical lattice structures. When the similarity threshold is $\delta = 1$, the similarity relation coincides with equality of the corresponding attribute values. Hence, the induced equivalence classes coincide with the Pawlak indiscernibility classes. Consequently, under this condition, the proposed evaluation framework is directly connected to the classical Pawlak representation. For $\delta < 1$, however, similarity neighborhoods may contain objects that are not Pawlak-indiscernible. Therefore, the resulting topological-lattice reduct need not coincide with a classical Pawlak reduct.*

Remark 5.5. *For practical applications, the similarity threshold δ need not be fixed at 0.50. Instead, a data-driven threshold may be selected based on the observed distribution of the pairwise similarity values. For example, candidate thresholds can be evaluated over an appropriate range within $[0, 1]$, and the value yielding the highest decision accuracy, or providing the best trade-off between accuracy and neighborhood granularity, can be selected. The value $\delta = 0.50$ used in the numerical example is therefore an illustrative choice rather than a universal parameter.*

Proposition 5.6. *Let $n = |U|$ and $m = |C|$ denote the number of objects and conditional attributes, respectively. Algorithm 5.7 performs m single-attribute deletion tests. For each tested attribute, the similarity matrix is recomputed using the remaining $m - 1$ attributes, requiring $O(n^2(m - 1)) = O(n^2m)$ operations. Hence, the total cost of similarity-matrix construction over all deletion tests is $O(n^2m^2)$. For each deletion test, the computational cost also includes the construction of the corresponding neighborhood systems, topological and pretopological structures, interior and closure operators, and the analytical lattice. Let $T_{\text{top}}(n, m)$, $T_{\text{pre}}(n, m)$, and $T_{\text{lat}}(n, m)$ denote the computational costs of these structural and evaluation procedures for a single deletion test. Therefore, the overall computational complexity of Algorithm 5.7 can be expressed as $O\left(m\left[n^2m + T_{\text{top}}(n, m) + T_{\text{pre}}(n, m) + T_{\text{lat}}(n, m)\right]\right)$. In particular, the similarity-matrix construction contributes $O(n^2m^2)$ to the overall complexity.*

Algorithm 5.7: Attribute Reduction for Computing Reducts and Core**Input:** Information system $\mathcal{I}_{ST} = (U, C, D, g, v_i \in V)$, where $C = \{a_1, \dots, a_m\}$ and $m = |C|$ **Output:** Topological spaces, reducts, and core attributes

-
- 1: Read the information system for each $u \in U$
 - 2: Construct the similarity matrix using the conditional attributes in C
 - 3: At $\delta = 1$, generate the Pawlak approximation space
 - 4: Compute the accuracy measure $Acc_Q(C)$
 - 5: At $\delta > 0.50$, construct the neighborhood system and generate the corresponding topology and pretopology
 - 6: Compute the accuracy measure $Acc_R(C)$
 - 7: Construct the analytical lattice using $P(u)$
 - 8: Compute the accuracy measure $Acc_L(C)$
 - 9: **for each** $a_i \in C$ **do**
 - 10: Set $C_i = C \setminus \{a_i\}$
 - 11: Recompute the similarity matrix using C_i
 - 12: Generate the corresponding Pawlak, topological, pretopological, and analytical lattice structures
 - 13: Compute $Acc_Q(C_i)$, $Acc_R(C_i)$, and $Acc_L(C_i)$
 - 14: Compute $\Delta_M(a_i) = \frac{Acc_M(C)}{Acc_M(C_i)}$ for all $M \in \{Q, R, L\}$
 - 15: **if** $\Delta_M(a_i) \leq 1 - \varepsilon$ for all $M \in \{Q, R, L\}$ **then**
 - 16: Declare a_i dispensable
 - 17: **else**
 - 18: Declare a_i indispensable
 - 19: **end if**
 - 20: **end for**
 - 21: Define the core as the set of indispensable attributes
 - 22: Construct the reducts from subsets of C satisfying the dispensability criterion
 - 23: Print $\text{Red}(\mathcal{I}_{ST})$ and the core attributes
-

In Figure 6, we give a floatchart for the algorithm.

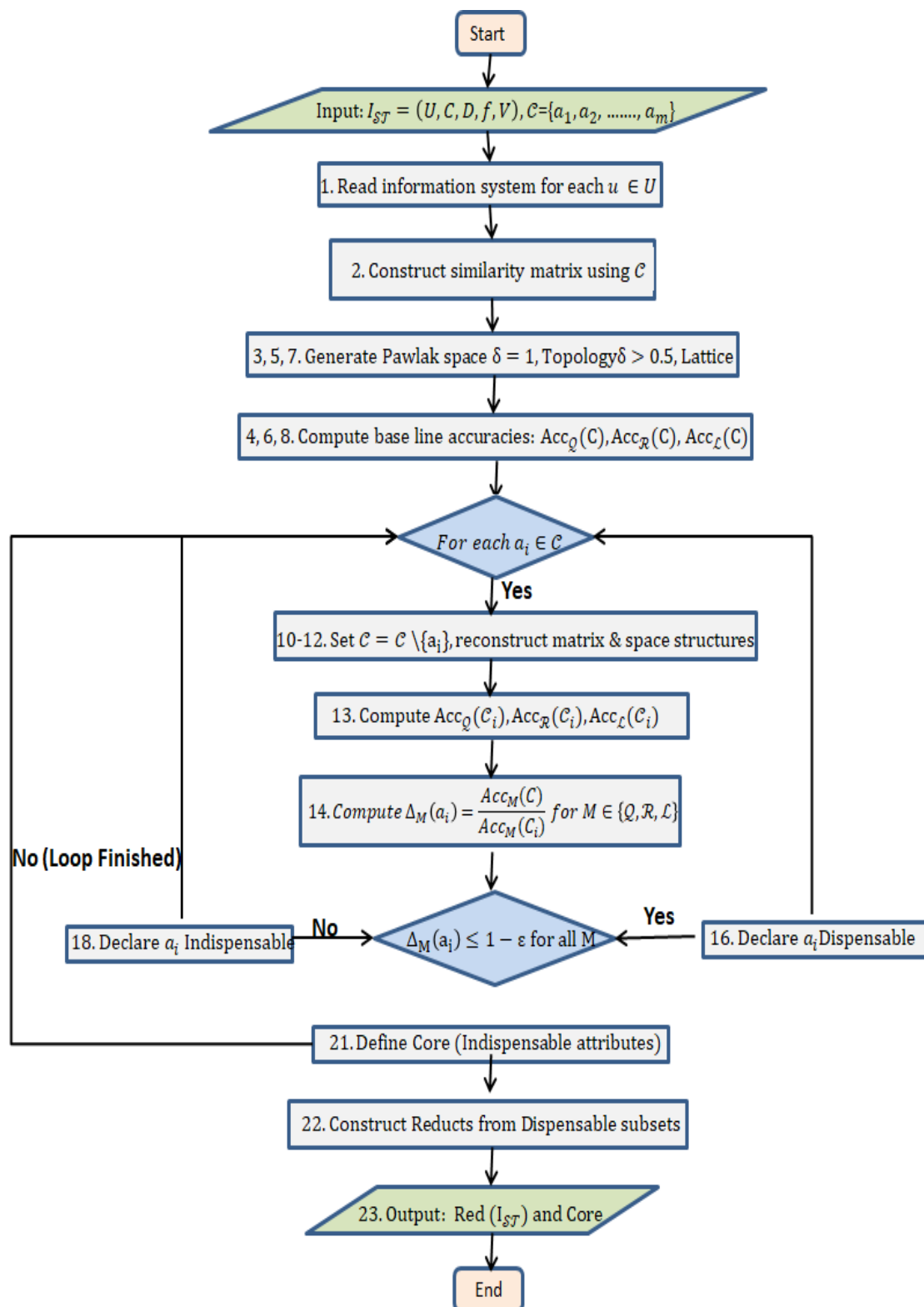


Figure 6. Flowchart of the topological spaces and core extraction algorithm.

Remark 5.7. *Algorithm 5.7 performs a single-attribute deletion analysis. Each conditional attribute is removed individually from the complete attribute set, and the resulting changes in the evaluation measures are recorded. The procedure is used to assess attribute significance and to identify indispensable attributes according to the adopted evaluation criterion. It does not perform recursive or iterative deletion of multiple attributes; therefore, each attribute is evaluated independently with respect to the complete attribute set.*

6. Conclusions and future work

This paper establishes a pretopological framework for analyzing information systems and evaluating the effectiveness of their attributes. Starting from an information tables, a similarity matrix based on the set of objects is constructed and used to generate different types of pretopological structures, including quasi-discrete, topological, and analytical structures, together with their corresponding interior and closure operators. In this context, the proposed approach determines precision measures for decision classes, thereby providing a mathematically sound measure of how effectively the local structure of the data supports classification. Based on these measures, an attribute reduction procedure is proposed to assess the effect of removing condition attributes by examining changes in the precision values. This procedure makes it possible to identify essential and redundant attributes while generally preserving the classification performance. The research further presents a lattice representation derived from the similarity matrix, providing an additional algebraic perspective on the relationships among objects and attributes. Overall, the research demonstrates that pretopology provides a flexible and rigorous framework for modeling information systems and guiding attribute selection, while also offering several avenues for extending the proposed approach to more complex and large-scale settings. Although the proposed framework has been illustrated using a representative information system, further validation using publicly available benchmark datasets is required to assess its scalability and robustness in high-dimensional, noisy, and incomplete data environments. Future work will include comprehensive comparisons with classical Pawlak attribute reduction, neighborhood rough set-based reduction, and other topology-based reduction methods using performance indicators such as the reduction rate, classification preservation accuracy, and computational complexity.

Author contributions

Mateb Aljabri: Conceptualization, methodology, formal analysis, investigation, visualization, writing- original draft; Abdelfattah El-Atik: Conceptualization, methodology, formal analysis, supervision, project administration, validation, writing- review and editing; Ahmed Zedan: Conceptualization, methodology, software, formal analysis, investigation, validation, writing- review and editing. All authors have read and approved the final version of the manuscript for publication.

Use of Generative-AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Conflict of interest

The authors declare no conflicts of interest.

References

1. W. Sierpinski, *General topology (Mathematical expositions No. 7)*, Toronto: University of Toronto Press, 1952. <https://doi.org/10.3138/9781487584894>
2. K. Kuratowski, *Topologie: Volume I*, Warszawa: Państwowe Wydawnictwo Naukowe, 1958.
3. W. Z. Wu, W. X. Zhang, Neighborhood operator systems and approximations, *Inf. Sci.*, **144** (2002), 201–217. [https://doi.org/10.1016/S0020-0255\(02\)00180-9](https://doi.org/10.1016/S0020-0255(02)00180-9)
4. Y. Y. Yao, Relational interpretations of neighborhood operators and rough set approximation operators, *Inf. Sci.*, **111** (1998), 239–259. [https://doi.org/10.1016/S0020-0255\(98\)10006-3](https://doi.org/10.1016/S0020-0255(98)10006-3)
5. Y. Y. Yao, Granular computing using neighborhood systems, In: *Advances in soft computing*, London: Springer, 1999, 539–553. https://doi.org/10.1007/978-1-4471-0819-1_40
6. Q. V. Bui, *Pretopology and topic modeling for complex systems analysis: Application on document classification and complex network analysis*, Ecole Pratique des Hautes Etudes, 2018. <https://doi.org/10.70675/092455f2zedd4z4a31zb90azae675c57487c>
7. A. A. Nasef, S. Jafari, M. Caldas, R. M. Latifd, A. A. Azzam, Preclosure operator and its applications in general topology, *J. Linear Topol. Algebra*, **7** (2018), 1–9.
8. Y. Y. Yao, A comparative study of fuzzy sets and rough sets, *Inf. Sci.*, **109** (1998), 227–242. [https://doi.org/10.1016/S0020-0255\(98\)10023-3](https://doi.org/10.1016/S0020-0255(98)10023-3)
9. N. Zhong, Rough sets in knowledge discovery and data mining, *J. Japan Soc. Fuzzy Theory Sys.*, **13** (2001), 581–591. <https://doi.org/10.3156/jfuzzy.13.6.581>
10. R. Intan, M. Mukaidono, Generalization of rough membership function based on α -coverings of the universe, In: *Advances in soft computing- AFSS 2002*, Berlin, Heidelberg: Springer, 2002, 129–136. https://doi.org/10.1007/3-540-45631-7_18
11. Y. Y. Yao, Generalized rough set models, In: *Rough sets in knowledge discovery*, Heidelberg: Physica-Verlag, 1998, 286–318.
12. A. Zedan, *A study of biapproximation spaces via topological concepts*, Ph.D. Thesis, South Valley University, 2010.
13. B. Ganter, R. Wille, *Formal concept analysis: Mathematical foundations*, 2 Eds., Cham: Springer, 2024. <https://doi.org/10.1007/978-3-031-63422-2>
14. A. El-Atik, A. Wahba, Topological approaches of graphs and their applications by neighborhood systems and rough sets, *J. Intell. Fuzzy Syst.*, **39** (2020), 6979–6992. <https://doi.org/10.3233/JIFS-200126>
15. T. M. Al-Shami, W. Q. Fu, E. A. Abo-Tabl, New rough approximations based on E-neighborhoods, *Complexity*, **2021** (2021), 6666853. <https://doi.org/10.1155/2021/6666853>

16. M. K. El-Bably, T. M. Al-Shami, Different kinds of generalized rough sets based on neighborhoods with a medical application, *Int. J. Biomath.*, **14** (2021), 2150086. <https://doi.org/10.1142/S1793524521500868>
17. T. Y. Lin, Neighborhood systems and approximation in relational databases and knowledge bases, In: *Proceedings of 4th international symposium on methodologies of intelligent systems*, 1989, 75–86.
18. T. M. Al-Shami, An improvement of rough sets accuracy measure using containment neighborhoods with a medical application, *Inf. Sci.*, **569** (2021), 110–124. <https://doi.org/10.1016/j.ins.2021.04.016>
19. T. M. Al-Shami, Improvement of the approximations and accuracy measure of a rough set using somewhere dense sets, *Soft Comput.*, **25** (2021), 14449–14460. <https://doi.org/10.1007/s00500-021-06358-0>
20. S. Liang, X. Yang, X. Chen, J. Li, Stable attribute reduction for neighborhood rough set, *Filomat*, **32** (2018), 1809–1815. <https://doi.org/10.2298/FIL1805809L>
21. T. Y. Lin, Granular computing on binary relations I: Data mining and neighborhood systems, II: rough set representations and belief functions, In: *Rough sets in knowledge discovery I*, Heidelberg: Physica-Verlag, 1998.
22. J. P. Auray, G. Duru, M. Mougeot, A pre-topological analysis of the input-output model, *Econom. Lett.*, **2** (1979), 343–347. [https://doi.org/10.1016/0165-1765\(79\)90048-X](https://doi.org/10.1016/0165-1765(79)90048-X)
23. J. Laborde, *Pretopology, a mathematical tool for structuring complex systems: Methods, algorithms and applications*, Université Paris sciences et lettres, 2019.
24. H. Saber Osman, S. A. El-Sheikh, A. E. Radwan, A. E. El-Atik, A model of π -pre-topological structures and related to human heart, *J. Intell. Fuzzy Syst.*, **44** (2023), 9431–9439. <https://doi.org/10.3233/JIFS-223891>
25. M. Dalud-Vincent, M. Brissaud, M. Lamure, Closed sets and closures in pretopology, *Int. J. Pure Appl. Math.*, **50** (2009), 391–402.
26. Z. Pawlak, Rough sets, *Int. J. Comput. Inf. Sci.*, **11** (1982), 341–356. <https://doi.org/10.1007/BF01001956>
27. Z. Pawlak, A. Skowron, Rough membership functions: A tool for reasoning with uncertainty, *Banach Cent. Publ.*, **28** (1993), 135–150. <https://doi.org/10.4064/-28-1-135-150>
28. A. El-Atik, A. Zedan, Some membership functions via neighborhood systems: Application to a rough set decision making, *Bol. Soc. Paran. Mat.*, **41** (2023), 1–11. <https://doi.org/10.5269/bspm.51936>
29. Z. Pawlak, *Rough sets: Theoretical aspects of reasoning about data*, Dordrecht: Springer, 1991. <https://doi.org/10.1007/978-94-011-3534-4>

