Cluster analysis based on fuzzy relations

Miin-Shen Yang*, Hsing-Mei Shih

Department of Mathematics, Chung-Yuan Christian University, Chung-Li, Taiwan 32023, Taiwan

Received 13 August 1997; received in revised form 19 April 1999; accepted 14 May 1999

Abstract

In this paper, cluster analysis based on fuzzy relations is investigated. Tamura's max-min n-step procedure is extended to all types of max-t compositions. A max-t similarity-relation matrix is obtained by beginning with a proximity-relation matrix on the proposed max-t n-step procedure. Then a clustering algorithm is created for the max-t similarity-relation matrix. Three critical max-t compositions of max-min, max-prod and max-t are compared. The max-t composition is recommended as the first choice among them. Several examples give more perspectives for different choices of max-t compositions. Finally, the topic of incomplete data via max-t compositions is discussed. Max-t compositions can be effectively used to treat the t-connected incomplete data. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Cluster analysis; Similarity relation; Proximity relation; Tolerance relation; max-t transitivity; max-t composition

1. Introduction

Cluster analysis is one of the major techniques in pattern recognition. Fuzzy sets give an idea of uncertainty of belongedness which is described by a membership function. It is natural to apply fuzzy set theory in cluster analysis. Bellman et al. [1] and Ruspini [5] first initiated research in fuzzy clustering. Now fuzzy clustering has been widely studied and applied in diverse areas. See, for example, Bezdek [2], Trauwaert et al. [7], Dave [4], Yang et al. [9–11]. These fuzzy clusterings can be roughly divided into two categories. One category, based on objective functions, is effective and studied popularly in the literature. Another category is based on a relation matrix such as correlation coefficient, equivalence relation, similarity relation and fuzzy relations, etc. Although the second type of clustering methods is eventually the novel method of agglomerative hierarchical clustering, they are simple and useful in application systems. The investigation reported in this paper focuses on this type of fuzzy clustering. Especially, a clustering algorithm based on fuzzy relations is constructed.

Zadeh [12] proposed a max-min transitivity and defined a max-min similarity relation using the idea of a fuzzy set. Since this max-min similarity relation is a resolution form of equivalence relations, one can
consequently obtain a corresponding multi-level hierarchical clustering. Tamura et al. [6] constructed an n-step procedure using max-min compositions of fuzzy relations by beginning with a subjective relation (with reflexivity and symmetry only). A max-min similarity relation is obtained after the n-step max-min compositions are performed. Finally, one can achieve a multi-level hierarchical clustering based on this max-min similarity relation.

This paper extends Tamura’s procedure to all types of max-t compositions. A max-t similarity relation is obtained after n-step max-t compositions are implemented. Then a clustering algorithm is proposed to obtain a clustering based on this max-t similarity relation. Because the max-min composition is composed of higher relation values, one finds that clustering results based on Tamura’s n-step procedure of max-min compositions do not explore the data well. If one uses max-Δ compositions, then better results can be obtained. This is because the max-Δ composition is softer than the max-min composition. After comparisons of t-norms are made, one finds that this Δ-norm seems to be the best choice of compositions in the n-step procedure for clustering. Some examples are presented in this paper.

In Section 2 fuzzy relation and its properties are first reviewed. Then an n-step procedure using max-min compositions of fuzzy relations by beginning with a subjective relation (with transitivity and symmetry only). A max-min similarity relation is obtained after the n-step max-min compositions are implemented. Finally, conclusions are stated in Section 6.

2. An n-step procedure

A crisp (binary) relation R between two sets, X and Y, is defined as a subset of X × Y. Denoted by R(X, Y), this relation is associated with an indicator function μR(x, y) which belongs to {0, 1} for all (x, y) in X × Y. That is, μR(x, y) = 1 if (x, y) ∈ R(X, Y), and μR(x, y) = 0 if (x, y) ∉ R(X, Y). Zadeh [12] defined a fuzzy relation R between X and Y as a fuzzy subset of X × Y by an extension to allow μR(x, y) to be membership functions assuming values in the interval [0, 1]. The value of μR(x, y) represents the strength of the relationship between x and y. In cluster analysis, one is only interested in relations in a single set X, i.e. R(X, X).

A similarity relation which is a fuzzy-type generalization of an equivalence relation was defined in [12]. A fuzzy relation R in X is called a similarity relation if it satisfies the following: (1) (reflexivity) μR(x, x) = 1 for all x in X; (2) (symmetry) μR(x, y) = μR(y, x) for all x, y in X; (3) (transitivity) μR(x, z) ≥ maxy∈X{min{μR(x, y), μR(y, z)}} for all x, z in X. It is observed that Zadeh defined a similarity relation by a release of transitivity to a max-min transitivity. Zadeh also provided a max-prod transitivity. Some years afterward, Bezdek and Harris [3] proposed a max-Δ transitivity. These are described as follows:

- max-min transitivity: μR(x, z) ≥ maxy∈X{min{μR(x, y), μR(y, z)}};
- max-prod transitivity: μR(x, z) ≥ maxy∈X{μR(x, y)μR(y, z)};
- max-Δ transitivity: μR(x, z) ≥ maxy∈X{max{0, μR(x, y) + μR(y, z) − 1}}.

The t-norm (see Zimmermann [13, p. 30]) was defined as a general form of the fuzzy intersection. Zimmermann [13] also listed some specified t-norms as follows:

(1) \( t_0(x, y) = \begin{cases} \min\{x, y\} & \text{if max}\{x, y\} = 1, \\ 0 & \text{otherwise} \end{cases} \) (drastic product);

(2) \( t_1(x, y) = \max\{0, x + y - 1\} \) (bounded difference);

(3) \( t_{1.5}(x, y) = \frac{xy}{2 - (x + y - xy)} \) (Einstein product);

(4) \( t_2(x, y) = xy \) (algebraic product).
Example 1. Transitivity seems to be more meaningful than max-min transitivity. The following example is given to explain.

The most defined distance functions are easy in satisfying the triangle inequality but difficult in satisfying the ultra-metric inequality. These physical and mathematical interpretations appeared in [3]. In this sense max-transitivity is more restrictive than that of max-min transitivity. Thus (max-t₁ transitivity) ⊃ (max-t₂ transitivity) ⊃ ··· ⊃ (max-tₙ transitivity). It is seen that the max-t₁ transitivity is a max-Δ transitivity; the max-t₂ transitivity is a max-prod transitivity; and the max-tₙ transitivity is a max-min transitivity. The condition of max-min transitivity is more restrictive than that of max-Δ transitivity. If one defines a distance \( d(x, y) = 1 - \mu_R(x, y) \) for all \( x, y \) in \( X \), then one has “max-min transitivity ⇔ ultra-metric inequality [12]” and “max-Δ transitivity ⇔ triangle inequality [3]”. The ultra-metric inequality implies the triangle inequality. The most defined distance functions are easy in satisfying the triangle inequality but difficult in satisfying the ultra-metric inequality. These physical and mathematical interpretations appeared in [3]. In this sense max-Δ transitivity seems to be more meaningful than max-min transitivity. The following example is given to explain.

**Definition 1** (max-t transitivity). For a fuzzy relation \( R \) in \( X \) and for a \( t \)-norm, the condition:

\[
\mu_R(x, z) \geq \max_{y \in X} \{ t(\mu_R(x, y), \mu_R(y, z)) \},
\]

for all \( x, z \) in \( X \)

is called a max-t transitivity.

It is easy to demonstrate that \( t_{a} \leq t_{1} \leq t_{1.5} \leq t_{2} \leq t_{2.5} \leq t_{1} \). Thus (max-t₀ transitivity) ⊃ (max-t₁ transitivity) ⊃ ··· ⊃ (max-tₙ transitivity). It is seen that the max-t₁ transitivity is a max-Δ transitivity; the max-t₂ transitivity is a max-prod transitivity; and the max-tₙ transitivity is a max-min transitivity. The condition of max-min transitivity is more restrictive than that of max-Δ transitivity. If one defines a distance \( d(x, y) = 1 - \mu_R(x, y) \) for all \( x, y \) in \( X \), then one has “max-min transitivity ⇔ ultra-metric inequality [12]” and “max-Δ transitivity ⇔ triangle inequality [3]”. The ultra-metric inequality implies the triangle inequality. The most defined distance functions are easy in satisfying the triangle inequality but difficult in satisfying the ultra-metric inequality. These physical and mathematical interpretations appeared in [3]. In this sense max-Δ transitivity seems to be more meaningful than max-min transitivity. The following example is given to explain.

**Example 1.** Let \( X \) be the set of \( x, y \) and \( z \), in which three patterns are represented, as follows:

\[
\begin{array}{c|c|c}
 x & y & z \\
\hline
 | & | & \hline
| & | & \\
\end{array}
\]

Suppose that \( \mu_R(x, y) = 0.7 \) and \( \mu_R(y, z) = 0.7 \). Then one has max-min transitivity: \( \mu_R(x, z) \geq \min\{\mu_R(x, y), \mu_R(y, z)\} = 0.7 \), and max-Δ transitivity: \( \mu_R(x, z) \geq \max\{0, \mu_R(x, y) + \mu_R(y, z) - 1\} = 0.4 \).

An analysis of the three patterns reveals that they have a common pattern, “□”, but with different crossing lines, “—” or “|”. Patterns \( x \) and \( y \) have one different crossing line “|”, (i.e. dissimilarity = 1 − 0.7 = 0.3). Patterns \( y \) and \( z \) have one different crossing line, “—”. But patterns \( x \) and \( z \) have two different crossing lines. Therefore, the restriction of \( \mu_R(x, z) \geq 0.7 \) is too high. The condition \( \mu_R(x, y) \geq 0.4 \) is more reasonable. In this case the max-Δ transitivity will be more meaningful than the max-min transitivity.

One may ask if the condition of max-Δ transitivity is still too restrictive and whether it will be released. It is seen that the triangle inequality seems to be a necessary condition for a distance function which is useful in practice. If one does cluster analysis based on these fuzzy relations, then the max-Δ transitivity seems to be a minimum restriction. For example, max-t₀ transitivity makes no sense in clustering.

Tamura et al. [6] effectively constructed an \( n \)-step procedure for clustering based on max-min compositions by beginning with a proximity relation. A transitive closure is finally obtained after \( n \) steps. This transitive closure is a max-min similarity relation. The resolution form of this similarity relation is a convex combination of equivalence relations. Then a partition tree is achieved.

**Definition 2** (Proximity relation). A fuzzy relation \( R \) in \( X \) is called a proximity relation if it satisfies

1. (reflexivity) \( \mu_R(x, x) = 1 \) \( \forall x \in X \), and
2. (symmetry) \( \mu_R(x, y) = \mu_R(y, x) \) \( \forall x, y \in X \).
In real pattern recognition such as visual images, smells, and pictures, etc., human subjectivity provides important information. This subjective information may be represented by a proximity relation. A measure of subjective similarity can be a fuzzy relation, which is necessary for reflexivity and symmetry. This is why Tamura et al. [6] constructed an $n$-step procedure of max-min compositions by beginning with a proximity relation. Now, this $n$-step procedure is extended to all max-$t$ compositions.

**Definition 3 (max-$t$ composition).** Given a $t$-norm and an initial fuzzy-relation matrix, $R^{(0)} = \{\gamma^{(0)}_{ij}\}$, then $R^{(n)} = \{\gamma^{(n)}_{ij}\}$ is called a max-$t$ composition. This is why important information. This subjective information may be represented by a proximity relation. A measure of subjective similarity can be a fuzzy relation, which is necessary for reflexivity and symmetry.

**Theorem 1 (An $n$-step procedure).** Suppose that $R^{(0)}$ is a proximity-relation matrix. Then, by max-$t$ compositions, one has

$$I < R^{(0)} < R^{(1)} < \cdots < R^{(n)} = R^{(n+1)} = \cdots,$$

where $R^{(n)}$ is a max-$t$ similarity relation. If $n$ is not finite, then $\lim_{n \to \infty} R^{(n)} = R^{(\infty)}$ with $R^{(\infty)}$ a max-$t$ similarity relation, i.e.

$$I < R^{(0)} < R^{(1)} < \cdots < R^{(n)} < R^{(n+1)} < \cdots < R^{(\infty)}.$$

**Proof.** Since $R^{(0)}$ is a proximity-relation matrix, $I < R^{(0)}$. Let $R^{(1)} = \{\gamma^{(1)}_{ij}\}$ with $\gamma^{(1)}_{ij} = \max_k t(\gamma^{(0)}_{ijk}, \gamma^{(0)}_{kij})$. Then, one has $\gamma^{(0)}_{ij} = t(\gamma^{(0)}_{ij}, 1) = t(\gamma^{(0)}_{ij}, \gamma^{(0)}_{ij}) \leq \max_k t(\gamma^{(0)}_{ijk}, \gamma^{(0)}_{kij}) = \gamma^{(1)}_{ij}$. That is, $R^{(0)} \leq R^{(1)}$. Suppose that $R^{(0)}$ does not have max-$t$ transitivity. Then, there exist $i$ and $j$ such that for some $l$ one has

$$\gamma^{(0)}_{ij} < t(\gamma^{(0)}_{il}, \gamma^{(0)}_{lj}) \Rightarrow \gamma^{(0)}_{ij} < \max_k t(\gamma^{(0)}_{ijk}, \gamma^{(0)}_{kij}) = \gamma^{(1)}_{ij} \Rightarrow R^{(0)} < R^{(1)}.$$

If $R^{(1)}$ does not have max-$t$ transitivity either, then similarly one has $R^{(1)} < R^{(2)}$. If the max-$t$ transitivity is not reached after $(n - 1)$ compositions, then

$$I < R^{(0)} < R^{(1)} < \cdots < R^{(n-1)} < R^{(n)}.$$

Suppose that $R^{(n)}$ has max-$t$ transitivity. Then for all $i, j$ one has $\gamma^{(n)}_{ij} \geq \max_k t(\gamma^{(n)}_{ijk}, \gamma^{(n)}_{kij}) = \gamma^{(n+1)}_{ij}$ and $t(\gamma^{(n)}_{ij}, \gamma^{(n)}_{ij}) = t(1, \gamma^{(n)}_{ij}) = \gamma^{(n)}_{ij}$. Then $\gamma^{(n+1)}_{ij} = \max_k t(\gamma^{(n)}_{ijk}, \gamma^{(n)}_{kij}) \geq t(\gamma^{(n)}_{ij}, \gamma^{(n)}_{ij}) = \gamma^{(n)}_{ij}$. One also has $\gamma^{(n+1)}_{ij} = \gamma^{(n)}_{ij}$, i.e. $R^{(n+1)} = R^{(n)}$. Similarly, $R^{(n+2)} = R^{(n+1)}$. That is,

$$I < R^{(0)} < R^{(1)} < \cdots < R^{(n-1)} < R^{(n)} = R^{(n+1)} = \cdots.$$

If there is not a finite $n$ such that $R^{(n)} = R^{(n+1)} = \cdots$, then

$$I < R^{(0)} < R^{(1)} < \cdots < R^{(n)} < R^{(n+1)} < \cdots < R^*,$$

where

$$R^* = \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{bmatrix}.$$
It is known that \( \{R^k \mid k = 0, 1, 2, \ldots \} \) is monotone and bounded. Then \( R^{(\infty)} = \lim_{n \to \infty} R^{(n)} \) exists. Next, it is claimed that \( R^{(\infty)} \) is a max-\( t \) similarity relation.

Recall that \( i_{ij}^{(n)} = \max_k t(\gamma_{ik}^{(n-1)}, \gamma_{kj}^{(n-1)}) \). Now, a new term is defined as \( i_{ij}^{(n)} = \max_k t(\gamma_{ik}^{(n-1)}, \gamma_{kj}^{(n-1)}) \). Although \( R^{(n)} \) and \( R^{(n)} \) are different, \( \lim_{n \to \infty} R^{(n)} = \lim_{n \to \infty} R^{(n)} = R^{(\infty)} \).

\[
\begin{align*}
\gamma_{ij}^{(1)} & = \max_{k_1} t(\gamma_{ik_1}^{(1)}, \gamma_{kj_1}^{(1)}), \\
\gamma_{ij}^{(2)} & = \max_{k_2} t(\gamma_{ik_2}^{(2)}, \gamma_{kj_2}^{(2)}) = \max_{k_1, k_2} t(\gamma_{ik_1}^{(1)}, \gamma_{kj_1}^{(1)}, \gamma_{ik_2}^{(1)}, \gamma_{kj_2}^{(1)}), \\
& \vdots \\
\gamma_{ij}^{(n)} & = \max_{k_1, \ldots, k_{n-1}} t(\gamma_{ik_1}^{(1)}, \gamma_{kj_1}^{(1)}, \gamma_{ik_2}^{(1)}, \gamma_{kj_2}^{(1)}, \ldots, \gamma_{ik_{n-1}}^{(1)}, \gamma_{kj_{n-1}}^{(1)}), \\
\gamma_{ij}^{(m+n)} & = \max_{k_1, \ldots, k_{n-1}} t(\gamma_{ik_1}^{(1)}, \gamma_{kj_1}^{(1)}, \gamma_{ik_2}^{(1)}, \gamma_{kj_2}^{(1)}, \ldots, \gamma_{ik_{n-1}}^{(1)}, \gamma_{kj_{n-1}}^{(1)}).
\end{align*}
\]

Then \( \gamma_{ij}^{(m+n)} \geq t(\gamma_{il}^{(m)}, \gamma_{lj}^{(n)}) \) for all \( l \). As \( m \to \infty \) and \( n \to \infty \), one has

\[
\gamma_{ij}^{(\infty)} \geq t(\gamma_{il}^{(\infty)}, \gamma_{lj}^{(\infty)}) \quad \text{for all } l.
\]

One also has \( \gamma_{il}^{(\infty)} = 1 \) and \( \gamma_{lj}^{(\infty)} = \gamma_{lj}^{(\infty)} \). That is, \( R^{(\infty)} \) is a max-\( t \) similarity relation. \( \Box \)

Note that if a max-min composition is chosen in Theorem 1, then one obtains the same result as from Tamura et al. [6]. Theorem 1 extends Tamura’s \( n \)-step procedure to all types of max-\( t \) compositions. That is, beginning with a proximity-relation matrix \( R^{(0)} \), one can arrive at a max-\( t \) similarity relation matrix \( R^{(n)} \) based on max-\( t \) compositions.

**Example 2.** Given a proximity-relation matrix

\[
R^{(0)} = \begin{bmatrix}
1 & 0.8 & 1 \\
0.7 & 0.2 & 1
\end{bmatrix}.
\]

(a) By max-min composition, one has

\[
R^{(1)} = \begin{bmatrix}
1 & 0.8 & 1 \\
0.7 & 0.7 & 1
\end{bmatrix} = R^{(2)}.
\]

(b) By max-prod composition, one has

\[
R^{(1)} = \begin{bmatrix}
1 & 0.8 & 1 \\
0.7 & 0.56 & 1
\end{bmatrix} = R^{(2)}.
\]
By max-Δ composition, one has

\[ R^{(1)} = \begin{bmatrix} 1 & 0.8 & 1 \\ 0.7 & 0.5 & 1 \end{bmatrix}. \]

A proximity relation merely represents a subjective similarity relation. It is an initial relation which does not have transitivity. By max-\( t \) compositions, one can obtain a max-\( t \) transitivity. In Example 2, \( \gamma_{23}^{(0)} = 0.2 \).

By max-min compositions, the value obtained, \( \gamma_{23}^{(1)} = 0.7 \), is too high; but by max-Δ compositions, one has \( \gamma_{23}^{(1)} = 0.5 \), which is relatively close to \( \gamma_{23}^{(0)} = 0.2 \). In some sense, max-Δ compositions can attain transitivity and also remain nearest to the original subjective similarity. That is, the max-Δ composition seems to be more meaningful than the max-min composition.

**Theorem 2.** Let \( t_1 \) and \( t_2 \) be any two \( t \)-norns. Let \( R^{(0)} \) be a proximity-relation matrix. Suppose that \( I \preceq R^{(0)} \preceq \cdots \preceq R^{(n_1)} = R^{(n_1+1)} \) by the \( n \)-step procedure based on max-\( t_1 \) compositions and \( I \preceq R^{(0)} \preceq \cdots \preceq R^{(n_2)} = R^{(n_2+1)} \) by the \( n \)-step procedure based on max-\( t_2 \) compositions. If \( t_1 \)-norm \( \preceq t_2 \)-norm then \( n_1 \preceq n_2 \).

**Proof.** Since \( R^{(n_2)} \) is a max-\( t_2 \) similarity relation, \( \gamma_{ij}^{(n_2)} \geq \max_k t_2(\gamma_{ik}^{(n_2)}, \gamma_{kj}^{(n_2)}) \). But \( t_2(\gamma_{ik}^{(n_1)}, \gamma_{kj}^{(n_1)}) \geq t_1(\gamma_{ik}^{(n_1)}, \gamma_{kj}^{(n_1)}) \) for all \( k \). Then \( \gamma_{ij}^{(n_2)} \geq \max_k t_1(\gamma_{ik}^{(n_1)}, \gamma_{kj}^{(n_1)}) \). That is, \( R^{(n_2)} \) is also a max-\( t_1 \) similarity relation. Therefore, \( n_2 \geq n_1 \). \( \square \)

Theorem 2 states that the number of max-Δ compositions to reach its transitive closure \( R^{(n_1)} \) is less than that of the max-min compositions to reach its transitive closure \( R^{(n_2)} \), i.e. \( n_1 \leq n_2 \). Now, the resolution form of a max-\( t \) similarity relation is discussed.

**Definition 4.** For \( 0 \leq \alpha \leq 1 \), \( R_{\alpha} = \{(x, y) \mid \mu_R(x, y) \geq \alpha \} \) is called a \( \alpha \)-cut of the fuzzy relation \( R \). Thus, one has the proposition that if \( \alpha_1 \geq \alpha_2 \), then \( R_{\alpha_1} \subset R_{\alpha_2} \).

**Proposition 1** (Zadeh [12, p. 181]). For any fuzzy relation \( R \) on \( X \times Y \), one has the resolution form

\[ R = \bigcup_{\alpha} \alpha R_{\alpha}, \quad 0 < \alpha \leq 1, \]

where \( \alpha R_{\alpha} \) is a fuzzy relation on \( X \times Y \) defined as

\[ \mu_{\alpha R_{\alpha}}(x, y) = \alpha \mu_R(x, y) = \begin{cases} \alpha & \text{if } (x, y) \in R_{\alpha} \\ 0 & \text{otherwise.} \end{cases} \]

**Proposition 2** (Zadeh [12, p. 186]). If \( R \) is a max-min similarity relation, then for any \( 0 < \alpha \leq 1 \), \( R_{\alpha} \) should be an equivalence relation.

**Definition 5** (Tolerance relation). A crisp relation \( R \) on \( X \) is called a tolerance relation if it satisfies the following conditions: for all \( x, y \) in \( X 

1. (Reflexivity): \( \mu_R(x, x) = 1 \).
2. (Symmetry): \( \mu_R(x, y) = 1 \) implies \( \mu_R(y, x) = 1 \).

**Proposition 3.** If \( R \) is a proximity relation, then for any \( 0 < \alpha \leq 1 \), \( R_{\alpha} \) is a tolerance relation.
Proof. For any $0 < \alpha \leq 1$,
(a) (Reflexivity): $\mu_R(x, x) = 1 \geq \alpha$ for all $x$ in $X$. Then $(x, x) \in R_\alpha$ for all $x$ in $X$, i.e. $\mu_{R_\alpha}(x, x) = 1$ for all $x$ in $X$.
(b) (Symmetry): $\mu_R(x, y) = \mu_R(y, x)$ for all $x, y$ in $X$. Let $(x, y) \in R_\alpha$, i.e., $\mu_R(x, y) \geq \alpha$. Then $\mu_R(y, x) \geq \alpha$; i.e., $(y, x) \in R_\alpha$.
Therefore, $R_\alpha$ is a tolerance relation. □

According to the propositions, any fuzzy relation always has a resolution form. A max-min similarity relation especially has a resolution form of equivalence relations. Therefore, it can yield a partition tree. This is why Tamura et al. [6] had constructed an $n$-step procedure for pattern classification based on max-min compositions. A max-prod or max-$\Delta$ similarity relation does not have a resolution form with equivalence relations but only with tolerance relations. This is because of:
(a) max-prod: $(x, y) \in R_\alpha$ and $(y, z) \in R_\alpha$ 
\[ \Rightarrow \mu_R(x, y) \geq \alpha \text{ and } \mu_R(y, z) \geq \alpha \] 
\[ \Rightarrow \mu_R(x, z) \geq \max_{y \in X} \{ \mu_R(x, y) \mu_R(y, z) \} \geq \alpha^2 \neq \alpha. \]
(b) max-$\Delta$: $(x, y) \in R_\alpha$, $(y, z) \in R_\alpha$ 
\[ \Rightarrow \mu_R(x, z) \geq \max_{y \in X} \{ \max \{ 0, \mu_R(x, y) + \mu_R(y, z) - 1 \} \} \] 
\[ \Rightarrow \max \{ 0, 2\alpha - 1 \} \neq \alpha. \]
Therefore, a max-prod or max-$\Delta$ similarity relation cannot yield a partition tree with equivalence classes by its resolution form. But a clustering algorithm can be constructed for these similarity relations and is presented in the next section.

Example 3. Given a proximity-relation matrix $R^{(0)}$ on $X = \{x_1, x_2, x_3\}$ with
\[
R^{(0)} = \begin{bmatrix}
1 & 0.8 & 1 \\
0.7 & 0.2 & 1
\end{bmatrix}
\]
as in Example 2. Then
(a)
\[
R^{(1)} = \begin{bmatrix}
1 & 0.8 & 1 \\
0.7 & 0.7 & 1
\end{bmatrix} = R^{(2)}
\]
is a max-min similarity relation based on max-min compositions. One has
\[
R^{(1)} = x_1 \left[ \begin{array}{ccc}
1 & 1 & 1 \\
1 & 1 & 1 \\
1 & 1 & 1
\end{array} \right] \cup x_2 \left[ \begin{array}{ccc}
1 & 1 & 0 \\
0 & 0 & 1
\end{array} \right] \cup x_3 \left[ \begin{array}{ccc}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{array} \right],
\]
where $0 < x_1 \leq 0.7$, $0.7 < x_2 \leq 0.8$ and $0.8 < x_3 \leq 1$.
Then $R^{(1)}$ yields partitions as follows:
\[ 0 < \alpha \leq 0.7 \Rightarrow \{x_1, x_2, x_3\}, \]
\[ 0.7 < \alpha \leq 0.8 \Rightarrow \{x_1, x_2\}, \{x_3\}, \]
\[ 0.8 < \alpha \leq 1 \Rightarrow \{x_1\}, \{x_2\}, \{x_3\}. \]
(b) 

\[ R^{(1)} = \begin{bmatrix} 1 & 0.8 & 1 \\ 0.7 & 0.5 & 1 \end{bmatrix} \]

is a max-\(A\) similarity relation based on max-\(A\) compositions. One has 

\[ R^{(1)} = \alpha_1 \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \cup \alpha_2 \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \cup \alpha_3 \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \cup \alpha_4 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} , \]

where \(0 < \alpha_1 \leq 0.5, 0.5 < \alpha_2 \leq 0.7, 0.7 < \alpha_3 \leq 0.8, 0.8 < \alpha_4 \leq 1\).

It is known that 

\[ \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \]

is a tolerance relation, not an equivalence relation. Although it is certain that \(x_2\) and \(x_3\) are not in the same class, it is not so clear that \(x_1\) and \(x_2\) are in a class or that \(x_1\) and \(x_3\) are in a class.

Taking \(\alpha = 0.5 < \alpha \leq 0.7\), then 

\[ R_\alpha = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \]

is only a tolerance-relation matrix. One cannot obtain a clustering result from \(R_\alpha\). According to the max-\(A\) similarity-relation matrix, one has 

\[ \gamma_{12}^{(1)} = 0.8 \geq \alpha, \quad \gamma_{13}^{(1)} = 0.7 \geq \alpha \quad \text{and} \quad \gamma_{23}^{(1)} = 0.5 < \alpha. \]

This means that \(x_1\) and \(x_2\) or \(x_1\) and \(x_3\) can be in a class, but \(x_2\) and \(x_3\) cannot be in a class. It is known that 0.8 is a maximum value of similarity. By the maximum similarity principle to have \(x_1\) and \(x_2\) in a class is preferred. Therefore, one has the following clustering results:

\[ 0 < \alpha \leq 0.5 \Rightarrow \{x_1, x_2, x_3\}, \]
\[ 0.5 < \alpha \leq 0.8 \Rightarrow \{x_1, x_2\}, \{x_3\}, \]
\[ 0.8 < \alpha \leq 1 \Rightarrow \{x_1\}, \{x_2\}, \{x_3\}. \]

In the next section this type of maximum similarity principle is modified to create a clustering algorithm for all types of fuzzy relations.

3. A clustering algorithm

An objective function-oriented approach in cluster analysis is usually constructed on the basis of a distance function. It is a most popular approach to clustering, such as \(k\)-means, fuzzy \(c\)-means, expectation and maximization (EM), etc. But there are many data presented in subjective relations which are not present in a distance function. In other words, one has only a proximity-relation matrix. Since proximity relations do not have transitivity, they cannot be used in clustering. A max-\(t\) similarity relation is obtained after an \(n\)-step
procedure is performed, based on the max-t compositions proposed in Section 2, by beginning with a given proximityrelation matrix. Then, this max-t similarity-relation matrix is used for clustering.

If max-min compositions are chosen in the n-step procedure, then one has a max-min similarity-relation matrix. Since a max-min similarity-relation matrix has a resolution form of equivalence relations, a multi-level partition tree is obtained. In Section 2, it is stated that the max-A composition is more meaningful and effective than the max-min composition as they are used in the n-step procedure. This is because a max-A similarity has transitivity and also keeps the closest values from the original proximity relation. But a max-A similarity does not have a resolution form of equivalence relations. Neither does it have a partition tree directly from its resolution form. In this section, a clustering algorithm for any max-t similarity relation based on a modified maximum similarity principle is presented. Beginning with a proximity-relation matrix, this clustering algorithm produces a level z partition through an n-step procedure based on max-t compositions.

The algorithm may be stated as follows:

**Clustering algorithm**

s0. Let R(0) = [γij]nxn be a given proximity-relation matrix and let I = {1, 2, …, n} be the index set. It is given that 0 < z ≤ 1.

s1. Obtain a max-t similarity-relation matrix R = [γij]nxn from the given proximity-relation matrix R(0) using the n-step procedure of max-t compositions proposed in Section 2.

s2. Set γij = 0 for all i = j and set γij = 0 for all γij < z.

s3. Choose s and t in I so that γst = max {γij | i ≠ j, i, j ∈ I}. Note that a tie is broken randomly.
   
   IF γst ≠ 0 THEN link s and t into the same cluster C = {s, t} and GOTO s4.
   ELSE PRINT all indices in I into separated clusters and STOP.

s4. Choose u in I \ C so that

\[ \sum_{i \in C} \gamma_{iu} = \max \left\{ \sum_{i \in C} \gamma_{ij} \mid j \in I \setminus C \text{ with } \gamma_{ij} \neq 0 \text{ for all } i \in C \right\} \]

A tie is broken randomly.
   
   IF there is such a u, THEN link u into C, i.e. C = {s, t, u}, and GOTO s4.
   ELSE PRINT the cluster C.

s5. Let I = I \ C and GOTO s3.

Now the algorithm is applied to the following examples.

**Example 4.** Given a 10 × 10 proximity-relation matrix, R(0), and given z = 0.55,

\[
R(0) = \begin{bmatrix}
1 & 0.2 & 0.5 & 0.8 & 0.6 & 0.6 & 0.9 & 0.3 & 0.2 & 0.3 \\
0.2 & 1 & 0.3 & 0.4 & 0.6 & 0.7 & 0.4 & 0.2 & 0.1 & 0.1 \\
0.5 & 0.3 & 1 & 0.2 & 0.9 & 0.3 & 0.6 & 0.3 & 0.2 & 0.3 \\
0.8 & 0.6 & 0.5 & 1 & 0.2 & 0.9 & 0.4 & 0.3 & 0.2 & 0.1 \\
0.6 & 0.7 & 0.3 & 0.7 & 1 & 0.3 & 0.2 & 0.1 & 0.5 & 0.4 \\
0.2 & 0.9 & 0.4 & 0.3 & 0.2 & 1 & 0.4 & 0.3 & 0.7 & 0.1 \\
0.3 & 0.2 & 0.1 & 0.5 & 0.4 & 0.1 & 1 & 0.8 & 0.1 & 0.9 \\
0.9 & 0.8 & 0.3 & 0.4 & 0.5 & 0.3 & 0.6 & 0.3 & 0.2 & 0.1 \\
0.4 & 0.3 & 0.7 & 0.1 & 0.8 & 0.7 & 0.1 & 0 & 1 & 0.3 \\
0.3 & 0.2 & 0.6 & 0.3 & 0.9 & 0.2 & 0.3 & 0.2 & 0.1 & 1 \\
\end{bmatrix}
\]
s0. Let \( I = \{1, 2, \ldots, 10\} \) and \( x = 0.55 \).

s1. Obtain a max-\( A \) similarity-relation matrix \( R = [\gamma_{ij}]_{10 \times 10} \) by the \( n \)-step procedure,

\[
R = \begin{bmatrix}
1 & 0.7 & 0.5 & 0.8 & 0.6 & 0.6 & 0.5 & 0.4 & 0.9 & 0.5 \\
0.7 & 1 & 1 & 0.7 & 0.5 & 0.5 & 0.4 & 0.6 & 0.6 & 0.5 \\
0.5 & 0.3 & 1 & 1 & 0.7 & 0.5 & 0.5 & 0.4 & 0.5 & 0.3 \\
0.8 & 0.6 & 0.5 & 1 & 1 & 0.5 & 0.5 & 0.4 & 0.4 & 0.3 \\
0.6 & 0.7 & 0.5 & 0.7 & 1 & 1 & 0.5 & 0.4 & 0.4 & 0.3 \\
0.6 & 0.9 & 0.4 & 0.6 & 0.5 & 1 & 0.7 & 0.4 & 0.3 & 0.2 \\
0.5 & 0.4 & 0.1 & 0.5 & 0.4 & 0.3 & 1 & 0.7 & 0.6 & 0.5 \\
0.9 & 0.8 & 0.4 & 0.7 & 0.5 & 0.7 & 0.6 & 1 & 0.6 & 0.5 \\
0.4 & 0.6 & 0.7 & 0.5 & 0.8 & 0.7 & 0.2 & 0.4 & 0.3 & 0.2 \\
0.5 & 0.6 & 0.6 & 0.9 & 0.5 & 0.3 & 0.4 & 0.7 & 1 & 0.5 \\
\end{bmatrix}
\]

s2. Set \( \gamma_{ij} = 0 \) for all \( i = j \), and set \( \gamma_{ij} = 0 \) for all \( \gamma_{ij} < x \). Then

\[
R_{0.55} = \begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.8 & 0.6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.6 & 0.7 & 0 & 0.7 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.6 & 0.9 & 0 & 0 & 0.6 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.9 & 0.8 & 0 & 0 & 0.7 & 0 & 0.7 & 0 & 0 & 0 \\
0 & 0.6 & 0.7 & 0 & 0.8 & 0.7 & 0 & 0 & 0 & 0 \\
0 & 0 & 0.6 & 0.6 & 0.9 & 0 & 0 & 0 & 0.7 & 0 \\
\end{bmatrix}
\]

s3. The values \( \gamma_{18} = \gamma_{26} = \gamma_{510} = 0.9 \) are maximum. If \( \gamma_{18} \) is randomly selected, then \( C = \{1, 8\} \).

s4. The values \( \gamma_{12} + \gamma_{82} = \gamma_{14} + \gamma_{84} = 1.5 \) are maximum. If \( j = 4 \) is randomly selected, then \( C = \{1, 8, 4\} \).

s5. Let \( I = I \setminus C = \{3, 5, 6, 7, 9, 10\} \).

s3. The value \( \gamma_{510} = 0.9 \) is maximum. Then \( C = \{5, 10\} \).

s4. The value \( \gamma_{59} + \gamma_{109} = 1.5 \) is maximum. Then \( C = \{5, 10, 9\} \).

s5. Let \( I = I \setminus C = \{3, 6, 7\} \).

s3. When \( \gamma_{36} = \gamma_{37} = \gamma_{67} = 0 \), then \( \{3\} \), \( \{6\} \), \( \{7\} \) are three separated clusters.

Thus, a level 0.55 partition \( \{1, 8, 4, 2\}, \{5, 10, 9\}, \{3\}, \{6\}, \{7\} \) is obtained.

It has been mentioned that the proposed clustering algorithm gives the same clustering results for a max-min similarity-relation matrix as that from resolution forms of equivalence relations. In fact, the algorithm can directly obtain clustering results for a given proximity-relation matrix by passing step s1 of the \( n \)-step procedure; but it makes no sense, because a proximity relation does not have transitivity. For example, if one directly clusters on the proximity-relation matrix \( R^{(0)} \) in Example 5, then one has a level 0.55 partition...
\{1,8\}, \{2,6\}, \{5,10\}, \{3,9\}, \{4\}, \{7\}. It is found that such a partition is assembled only with a maximum relation. Obviously, it is not a good result. Therefore, it is necessary to have an \(n\)-step procedure from step \(s_1\) in the algorithm to transfer a proximity-relation matrix to a max-similarity-relation matrix.

Next, an interesting example, calling attention to the characteristic beauty of Chinese characters, is presented.

**Example 5.** Seven different Chinese characters are selected, as follows:

\[
\begin{array}{ccccccc}
(1) & (2) & (3) & (4) & (5) & (6) & (7) \\
\begin{array}{c}
\Box \\
\Box \\
\Box \\
\Box \\
\Box \\
\Box \\
\Box \\
\end{array} & \begin{array}{c}
\Box \\
\Box \\
\Box \\
\Box \\
\Box \\
\Box \\
\Box \\
\end{array} & \begin{array}{c}
\Box \\
\Box \\
\Box \\
\Box \\
\Box \\
\Box \\
\Box \\
\end{array} & \begin{array}{c}
\Box \\
\Box \\
\Box \\
\Box \\
\Box \\
\Box \\
\Box \\
\end{array} & \begin{array}{c}
\Box \\
\Box \\
\Box \\
\Box \\
\Box \\
\Box \\
\Box \\
\end{array} & \begin{array}{c}
\Box \\
\Box \\
\Box \\
\Box \\
\Box \\
\Box \\
\Box \\
\end{array} & \begin{array}{c}
\Box \\
\Box \\
\Box \\
\Box \\
\Box \\
\Box \\
\Box \\
\end{array} \\
\end{array}
\]

These seven Chinese characters have all the basic pattern “□”, with different crossing lines added: “—”, “[”, “|”, “|”. Therefore, their subjective relations are stated as follows: (a) \(\gamma_{ij} = 0.75\) when patterns \(i\) and \(j\) have only one different crossing line; (b) \(\gamma_{ij} = 0.50\) when they have two different crossing lines; (c) \(\gamma_{ij} = 0.25\) when they have three different crossing lines; and (d) \(\gamma_{ij} = 0\) when they have four different crossing lines. Thus, the following proximity-relation matrix is obtained:

\[
R^{(0)} = \begin{bmatrix}
1 & 0.75 & 0.25 & 0.25 & 0.25 & 0.50 & 0.50 \\
0.75 & 1 & 0.50 & 0.50 & 0.50 & 0.75 & 0.75 \\
0.25 & 0.50 & 0.25 & 0.25 & 0.25 & 0.50 & 0.50 \\
0.25 & 0.50 & 0.50 & 0.25 & 0.25 & 0.50 & 0.50 \\
0.25 & 0.50 & 0.50 & 0.50 & 0.50 & 0.75 & 0.75 \\
0.50 & 0.75 & 0.75 & 0.75 & 0.75 & 0.75 & 1 \\
0.50 & 0.75 & 0.75 & 0.75 & 0.75 & 0.75 & 1 \\
0.50 & 0.75 & 0.75 & 0.75 & 0.75 & 0.75 & 1 \\
\end{bmatrix}
\]

By the \(n\)-step procedure, one has a max-min similarity-relation matrix \(R\) with

\[
R = \begin{bmatrix}
1 & 0.75 & 0.75 & 0.75 & 0.75 & 0.75 & 0.75 \\
0.75 & 1 & 0.75 & 0.75 & 0.75 & 0.75 & 0.75 \\
0.75 & 0.75 & 1 & 0.75 & 0.75 & 0.75 & 0.75 \\
0.75 & 0.75 & 0.75 & 1 & 0.75 & 0.75 & 0.75 \\
0.75 & 0.75 & 0.75 & 0.75 & 1 & 0.75 & 0.75 \\
0.75 & 0.75 & 0.75 & 0.75 & 0.75 & 1 & 0.75 \\
0.75 & 0.75 & 0.75 & 0.75 & 0.75 & 0.75 & 1 \\
\end{bmatrix}
\]

Then one obtains the following clustering results:

\[0 < \alpha \leq 0.75 \Rightarrow \{1,2,3,4,5,6,7\},\]
\[0.75 < \alpha \leq 1 \Rightarrow \{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}, \{7\}.\]

\(R^{(0)}\) is itself a max-A similarity relation. The following clustering results are obtained by the proposed clustering algorithm:

\[0 < \alpha \leq 0.25 \Rightarrow \{3,4,5,6,7\}, \{1,2\} \text{ or } \{3,7,5,4,2\}, \{1,6\}, \text{ etc.},\]
\[0.25 < \alpha \leq 0.50 \Rightarrow \{3,4,5,7\}, \{1,2\}, \{6\} \text{ or } \{4,5,6,3\}, \{2,7,1\}, \text{ etc.},\]
0.50 < x ≤ 0.75 ⇒ \{1, 2\}, \{3, 4\}, \{5, 7\}, \{6\} or \{1, 2\}, \{4, 5\}, \{3, 7\}, \{6\}, etc.,

0.75 < x ≤ 1 ⇒ \{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}, \{7\}.

By a comparison of the max-\( \Delta \) and the max-min similarity-relation matrix for clustering in this example, it is seen that the max-\( \Delta \) has results softer than the max-min. The max-\( \Delta \) also seems to have clustering results better than those of the max-min.

4. A practical example

In this section the proposed clustering algorithm is applied to a practical example. Portraits of 15 members from three families \( A \), \( B \) and \( C \) are collected. Numbers 1–15 are marked corresponding to the portraits, as follows:

<table>
<thead>
<tr>
<th></th>
<th>Dad</th>
<th>Mom</th>
<th>Children</th>
</tr>
</thead>
<tbody>
<tr>
<td>family ( A ):</td>
<td>4</td>
<td>12</td>
<td>2,7</td>
</tr>
<tr>
<td>family ( B ):</td>
<td>8</td>
<td>11</td>
<td>1,5,9,14</td>
</tr>
<tr>
<td>family ( C ):</td>
<td>3</td>
<td>15</td>
<td>6,10,13</td>
</tr>
</tbody>
</table>

Subjective similarity is assigned as: very similar = 0.8, similar = 0.6, not so similar = 0.4, different = 0.2, quite different = 0. According to the above assignment of subjective similarities, one has the following proximity-relation matrix \( R^{(0)} \) of these 15 portraits:

\[
R^{(0)} = \begin{bmatrix}
1 & 0.2 & 1 \\
0 & 0.2 & 1 \\
0.4 & 0.6 & 0 & 1 \\
0.8 & 0.2 & 0 & 0 & 1 \\
0.4 & 0.2 & 0.6 & 0.4 & 0.2 & 1 \\
0.2 & 0.8 & 0.2 & 0.8 & 0.4 & 0.2 & 1 \\
0.8 & 0.8 & 0 & 0 & 0.6 & 0.4 & 0.4 & 1 \\
0.8 & 0.4 & 0.2 & 0 & 0.6 & 0.2 & 0.2 & 0.4 & 1 \\
0.2 & 0 & 0.6 & 0.2 & 0.2 & 0.8 & 0 & 0 & 0.4 & 1 \\
0.8 & 0 & 0 & 0 & 0.6 & 0.2 & 0.4 & 0 & 0.8 & 0.2 & 1 \\
0.2 & 0.8 & 0 & 0.2 & 0.2 & 0.2 & 0.8 & 0 & 0 & 0 & 1 \\
0.4 & 0.2 & 0.6 & 0 & 0.2 & 0.8 & 0 & 0.2 & 0.6 & 0 & 0 & 1 \\
0.8 & 0.2 & 0 & 0 & 0.8 & 0.4 & 0.2 & 0.6 & 0.8 & 0.4 & 0.8 & 0 & 0.4 & 1 \\
0 & 0 & 0 & 0 & 0.2 & 0.4 & 0.4 & 0 & 0.2 & 0.8 & 0 & 0 & 0.8 & 0.4 & 1 \\
\end{bmatrix}
\]

If the clustering algorithm is directly applied to this proximity relation matrix \( R^{(0)} \), it has ridiculous results, such as a couple or members of a family separated. This is because there is no transitivity in a proximity relation. Therefore, max-\( \tau \) compositions are used to compose \( R^{(0)} \) as a max-\( \tau \) similarity-relation matrix. If the
max-min $n$-step procedure is chosen, one has $R^{(0)} < R^{(1)} < R^{(2)} < R^{(3)} = R^{(4)}$ with

$$\begin{bmatrix}
1 \\
0.8 & 1 \\
0.4 & 0.4 & 1 \\
0.8 & 0.8 & 0.4 & 1 \\
0.8 & 0.8 & 0.4 & 0.8 & 1 \\
0.4 & 0.4 & 0.6 & 0.4 & 0.4 & 1 \\
0.8 & 0.8 & 0.4 & 0.8 & 0.4 & 1 \\
0.8 & 0.8 & 0.4 & 0.8 & 0.4 & 0.8 & 1 \\
0.8 & 0.8 & 0.4 & 0.8 & 0.4 & 0.8 & 0.8 & 1 \\
0.4 & 0.4 & 0.6 & 0.4 & 0.4 & 0.8 & 0.8 & 0.4 & 0.4 & 1 \\
0.8 & 0.8 & 0.4 & 0.8 & 0.8 & 0.8 & 0.8 & 0.4 & 0.4 & 1 \\
0.8 & 0.8 & 0.4 & 0.8 & 0.4 & 0.8 & 0.8 & 0.8 & 0.4 & 0.8 & 1 \\
0.4 & 0.4 & 0.6 & 0.4 & 0.4 & 0.8 & 0.8 & 0.4 & 0.4 & 0.8 & 0.8 & 0.8 & 0.4 & 0.8 & 0.4 & 1 \\
0.4 & 0.4 & 0.6 & 0.4 & 0.4 & 0.8 & 0.8 & 0.8 & 0.8 & 0.8 & 0.8 & 0.4 & 0.8 & 0.4 & 1 \\
0.8 & 0.8 & 0.4 & 0.8 & 0.8 & 0.8 & 0.8 & 0.8 & 0.8 & 0.8 & 0.8 & 0.8 & 0.8 & 0.8 & 0.8 & 0.8 & 0.4 & 0.4 & 1 \\
0.4 & 0.4 & 0.6 & 0.4 & 0.4 & 0.8 & 0.8 & 0.8 & 0.8 & 0.8 & 0.8 & 0.8 & 0.8 & 0.8 & 0.8 & 0.8 & 0.8 & 0.8 & 0.4 & 1 \\
0.4 & 0.4 & 0.6 & 0.4 & 0.4 & 0.8 & 0.8 & 0.8 & 0.8 & 0.8 & 0.8 & 0.8 & 0.8 & 0.8 & 0.8 & 0.8 & 0.8 & 0.8 & 0.8 & 1 \\
\end{bmatrix}$$

By the clustering algorithm, one obtains the clustering results as follows:

- $0 < x < 0.4 \Rightarrow \{1, 2, 3, \ldots, 15\}$,
- $0.4 < x < 0.6 \Rightarrow \{2, 14, 1, 12, 4, 9, 5, 7, 8, 11\}, \{6, 10, 13, 15, 3\}$,
- $0.6 < x < 0.8 \Rightarrow \{2, 14, 1, 12, 4, 9, 5, 7, 8, 11\}, \{6, 10, 13, 15, 3\}$,
- $0.8 < x \leq 1 \Rightarrow \{1\}, \{2\}, \{3\}, \ldots, \{15\}$.

If the max-$\Delta$ $n$-step procedure is chosen, one has $R^{(0)} < R^{(1)} < R^{(2)} = R^{(3)}$ with

$$\begin{bmatrix}
1 \\
0.6 & 1 \\
0 & 0.2 & 1 \\
0.4 & 0.6 & 0 & 1 \\
0.8 & 0.4 & 0 & 0.2 & 1 \\
0.4 & 0.2 & 0.6 & 0.4 & 0.2 & 1 \\
0.4 & 0.8 & 0.2 & 0.8 & 0.4 & 0.2 & 1 \\
0.8 & 0.8 & 0 & 0.4 & 0.6 & 0.4 & 0.6 & 1 \\
0.8 & 0.4 & 0.2 & 0.2 & 0.6 & 0.2 & 0.2 & 0.6 & 1 \\
0.2 & 0 & 0.6 & 0.2 & 0.2 & 0.8 & 0.2 & 0.2 & 0.4 & 1 \\
0.8 & 0.4 & 0 & 0.2 & 0.6 & 0.2 & 0.4 & 0.6 & 0.8 & 0.2 & 1 \\
0.4 & 0.8 & 0 & 0.6 & 0.2 & 0.2 & 0.8 & 0.6 & 0.2 & 0 & 0.2 & 1 \\
0.4 & 0.2 & 0.6 & 0.2 & 0.2 & 0.8 & 0.2 & 0.2 & 0.6 & 0.2 & 0 & 1 \\
0.8 & 0.4 & 0 & 0.2 & 0.8 & 0.2 & 0.2 & 0.6 & 0.8 & 0.4 & 0.8 & 0.2 & 0.4 & 1 \\
0.2 & 0.2 & 0.4 & 0.2 & 0.2 & 0.6 & 0.4 & 0 & 0.2 & 0.8 & 0.2 & 0.8 & 0.4 & 0.8 & 0.4 & 0.8 & 0.4 & 0.8 & 0.8 & 0.4 & 1 \\
\end{bmatrix}$$

By applying the clustering algorithm, one obtains the following results:

- $0 < x < 0.2 \Rightarrow \{13, 15, 10, 6, 3\}, \{1, 11, 14, 9, 5, 8, 2, 7, 12, 4\}$,
- $0.2 < x < 0.4 \Rightarrow \{13, 15, 10, 6, 3\}, \{1, 11, 14, 9, 5, 8, 2\}, \{7, 12, 4\}$,
- $0.4 < x < 0.6 \Rightarrow \{6, 10, 13, 15\}, \{1, 9, 11, 14, 5, 8\}, \{2, 7, 12, 4\}, \{3\}$,
- $0.6 < x < 0.8 \Rightarrow \{6, 10\}, \{1, 8\}, \{2, 12, 7\}, \{9, 11, 14\}, \{13, 15\}, \{3\}, \{4\}, \{5\}$,
- $0.8 < x \leq 1 \Rightarrow \{1\}, \{2\}, \{3\}, \ldots, \{15\}$. 
5. Incomplete data via max-$t$ compositions

An $n \times n$ proximity-relation matrix $R^{(0)}$ represents subjective relations of $n$ objects. For a given proximity-relation matrix, a suitable max-$t$ composition (e.g. max-$A$) is first chosen and a max-$t$ similarity-relation matrix is obtained based on the max-$t$ $n$-step procedure. Then the clustering algorithm produces clustering results from these $n$ objects. However, sometimes some data may have been missed with an incomplete proximity-relation matrix. In the case of the so-called $t$-connected incompleteness, max-$t$ compositions can be used to locate the missing information. In this section this type of incomplete data via max-$t$ compositions is discussed. It is desirable to use known information to estimate unknown missing information via max-$t$ compositions.

**Definition 6** (Incomplete relation matrix). A relation matrix $R = [r_{ij}]_{n \times n}$ is called incomplete if some elements $r_{ij}$ of $R$ are unknown or missing.

**Definition 7** ($t$-connected). An incomplete proximity-relation matrix $R^{(0)} = [r_{ij}^{(0)}]_{n \times n}$ is called $t$-connected if there is an integer $n$ so that $R^{(n)}$ (based on max-$t$ compositions) becomes complete.

**Example 6.** For a $4 \times 4$ proximity-relation matrix $R^{(0)}$,

(a) \[
\begin{bmatrix}
1 & 0.2 & \times & 0.9 \\
0.2 & 1 & \times & 0.2 \\
\times & \times & 1 & \times \\
0.9 & 0.2 & \times & 1
\end{bmatrix}
\text{ and } \begin{bmatrix}
1 & \times & \times & 0.9 \\
\times & 1 & 0.7 & \times \\
\times & 0.7 & 1 & \times \\
0.9 & \times & \times & 1
\end{bmatrix}
\]

are not $t$-connected.

(b) \[
\begin{bmatrix}
1 & 0.2 & \times & 0.9 \\
0.2 & 1 & 0.7 & \times \\
\times & 0.7 & 1 & \times \\
0.9 & \times & \times & 1
\end{bmatrix} \text{, } \begin{bmatrix}
1 & 0.2 & 0.5 & \times \\
0.2 & 1 & \times & 0.2 \\
\times & 0.2 & 1 & 0.3 \\
\times & \times & \times & 1
\end{bmatrix}
\text{ and } \begin{bmatrix}
1 & 0.2 & \times & 0.9 \\
0.2 & 1 & \times & 0.2 \\
\times & 0.2 & 1 & 0.3 \\
0.9 & 0.2 & 0.3 & 1
\end{bmatrix}
\]

are $t$-connected.

For any $t$-connected incomplete proximity-relation matrix $R^{(0)} = [r_{ij}^{(0)}]_{n \times n}$, one can estimate all missing data $\hat{\gamma}_{ij}^{(0)}$ by max-$t$ compositions $\hat{\gamma}_{ij}^{(n)} = \max_k t(\hat{\gamma}_{ik}^{(n-1)}, \hat{\gamma}_{kj}^{(n-1)})$. If the estimate $\hat{\gamma}_{ij}^{(n)}$ is larger than the original missing data $\gamma_{ij}^{(0)}$, then the missing data $\gamma_{ij}^{(0)}$ can be located exactly by the estimate $\hat{\gamma}_{ij}^{(n)}$. Suppose that the missing data
\( \gamma_{ij}^{(0)} \) is larger than the estimate \( \gamma_{ij}^{(n)} \). That is, the original subjective relationship \( \gamma_{ij}^{(0)} \) is significant but it is missing. Then the estimate \( \gamma_{ij}^{(n)} \) can also give a conservative estimate of \( \gamma_{ij}^{(0)} \). Therefore, estimates based on max-\( t \) compositions suggest a reasonable estimate for missing relation data. Such an estimate is shown in the next example.

**Example 7.** Given a proximity-relation matrix \( R^{(0)} \) with

\[
R^{(0)} = \begin{bmatrix}
1 & 0.2 & 0.9 \\
0.2 & 1 & 0.7 \\
0.9 & 0.7 & 1
\end{bmatrix},
\]

by max-\( A \) compositions one has \( R^{(0)} < R^{(1)} < R^{(2)} = R^{(3)} \) with

\[
R^{(2)} = \begin{bmatrix}
1 & 0.7 & 0.9 \\
0.7 & 1 & 0.4 \\
0.9 & 0.4 & 1
\end{bmatrix}.
\]

(a) Suppose that \( \gamma_{21}^{(0)} = 0.2 \) is missing. Then the max-\( A \) estimate is

\[
\hat{\gamma}_{12}^{(1)} = \max \{ \max_k \{ 0, \gamma_{1k}^{(0)} + \gamma_{k2}^{(0)} - 1 \} \} = 0.7 > \gamma_{12}^{(0)}
\]

with

\[
\hat{R}^{(2)} = \begin{bmatrix}
1 & 0.7 & 0.9 \\
0.7 & 1 & 0.4 \\
0.9 & 0.4 & 1
\end{bmatrix} = R^{(2)}.
\]

Thus, one exactly locates the missing data based on max-\( A \) compositions.

(b) Suppose that \( \gamma_{42}^{(0)} = 0.8 \) is missing. Then the max-\( A \) estimate is

\[
\hat{\gamma}_{42}^{(1)} = 0.1 < \gamma_{42}^{(0)} \quad \text{and} \quad R^{(0)} < \hat{R}^{(1)} < \hat{R}^{(2)} = \hat{R}^{(3)}
\]

with

\[
\hat{R}^{(2)} = \begin{bmatrix}
1 & 0.7 & 0.9 \\
0.7 & 1 & 0.4 \\
0.9 & 0.6 & 1
\end{bmatrix} < R^{(2)}.
\]

Thus, one obtains the conservative estimates 0.6 of 0.8 and 0.3 of 0.5.

According to the analysis, if the missing data are not so significant for representing the information in relationships, then one can exactly locate all missing information based on max-\( t \) compositions. However, if the significant data in high relationships are missing, then one can locate only part of the information. This is quite reasonable. It demonstrates that max-\( t \) compositions can be effectively used to treat the missing data. Thus, the proposed clustering algorithm is also extended to treat all \( t \)-connected incomplete proximity-relation matrices.
6. Conclusions

Tamura et al. [6] proposed a max-min n-step procedure for clustering a proximity-relation matrix. In this paper, Tamura’s n-step procedure has been extended to all types of max-t compositions. A max-t similarity-relation matrix is obtained by beginning with a proximity-relation matrix based on the max-t n-step procedure. Since a max-t similarity-relation matrix may not have a resolution form of equivalence relations as a max-min similarity, a clustering algorithm has been provided for any max-t similarity-relation matrix. In comparisons of t-norms, it has been suggested that max-Α compositions may be the first choice among all max-t compositions. Examples have provided more perspective to the different choices of max-t compositions. For the t-connected incomplete data, max-t compositions can be used effectively to estimate the missing information. That incomplete data via max-t compositions extends the proposed clustering algorithm to treat the case of missing data. The approach proposed in this paper enriches the category of clustering based on fuzzy relations and was practically used in the multi-floor layout problem [8].

Acknowledgements

The authors are grateful to the referees for their suggestions and comments.

References