Clustering of Personality using Indeterminacy Based Personality Test

Ilanthenral Kandasamy · Florentin Smarandache

Abstract Triple Refined Indeterminate Neutrosophic Set (TRINS) a case of the refined neutrosophic set was introduced in [8]. The uncertain and inconsistent information which are available in real world is represented with sensitivity and accuracy by TRINS. Better precision is provided in handling indeterminacy; by classifying indeterminacy (I) into three, based on membership; as indeterminacy leaning towards truth (I_T), indeterminacy (I) and indeterminacy leaning towards false (I_F). Based on the Open Extended Jung Type Scale (OEJTS) test and TRINS, an indeterminacy based personality test was introduced [8]. A significant role is played by clustering in several real world applications. Based on a generalized distance measure between TRINS and related distance matrix, a clustering algorithm is constructed. This article proposes a Triple Refined Indeterminate Neutrosophic Minimum Spanning Tree (TRIN-MST) clustering algorithm, to cluster the data represented by TRINS. Illustrative examples using the indeterminacy based personality test are given to exhibit the applications and effectiveness of the TRIN-MST clustering algorithm.

Keywords Personality Test · Neutrosophic Set · Indeterminacy · Triple Refined Indeterminate Neutrosophic Set (TRINS) · Minimum Spanning Tree (MST) · Clustering · TRIN-MST clustering algorithm

1 Introduction

Carl Jung had theorized the eight psychological types based on two main attitude types: extroversion and introversion, two observing functions: intuition

Ilanthenral Kandasamy
School of Computer Science and Engineering, VIT University, Vellore, India.
E-mail: ilanthenral.k@vit.ac.in

Florentin Smarandache
Department of Mathematics, University of New Mexico, USA.

and sensation and two judging functions: feeling and thinking in his collected work [5]. The Myers-Briggs Type Indicator (MBTI) [12], is based on the theory given by Carl Jung. The psychological variations are sorted into four contrary pairs, or "dichotomies", that provides 16 feasible psychological types. The MBTI is a reflective self-analytic questionnaire designed to find the psychological inclinations of people's world view and their decision making process. These personality tests are mostly objective in nature, where the test taker is forced to select a dominant choice. Quoting Carl Jung himself *"There is no such thing as a pure extrovert or a pure introvert. Such a man would be in the lunatic asylum."*, it is clear that there are degrees of variations, no person fits into a category 100%. Since it is not feasible for a person to put down his answer as single choice in reality, without ignoring the other degrees of variation; it necessitates a tool which can give more than one choice to represent their personality.

Clustering plays a vital role in several scientific and engineering fields in form of pattern recognition, data mining and machine learning. Clustering categorizes the items into groups (clusters), that is, it groups the similar ones into the same class, and groups the dissimilar ones into different classes. Clustering analysis has been a hard one traditionally, which allocates an item to a specific class. Since many items have no rigid restrictions, it cannot be directly ascertained which class they should belong to. Therefore, it is required to divide them softly.

Fuzzy set theory introduced by Zadeh [24] proposes a constructive analytic method for soft division of sets. Zadeh's fuzzy set theory [24] was extended to intuitionistic fuzzy set (A-IFS), in which every entity is assigned a non-membership degree and a membership degree by Atanassov [2]. A-IFS was further generalized into the concept of interval valued intuitionistic fuzzy set (IVIFS) by Atanassov and Gargov [1].

To characterize inconsistent, imprecise, uncertain, and incomplete information which are existing in real world, the notion of neutrosophic set from philosophical angle was given by Smarandache [15]. The neutrosophic set articulates independently truth, indeterminacy and falsity memberships. It's functions $T_A(x), I_A(x)$, and $F_A(x)$ are real standard or nonstandard subsets of $]^-0, 1^+[$, that is, $T_A(x) : X \rightarrow]^-0, 1^+[$, $I_A(x) : X \rightarrow]^-0, 1^+[$, and $F_A(x) : X \rightarrow]^-0, 1^+[$, respectively with the condition $^-0 \leq supT_A(x) + supI_A(x) + supF_A(x) \leq 3^+$.

To overcome the difficulty of adapting neutrosophic set in this form in engineering and scientific fields, Wang et al [19] introduced a Single Valued Neutrosophic Set (SVNS), which is another form of a neutrosophic set. Fuzzy sets and intuitionistic fuzzy sets cannot deal with inconsistent and indeterminate information, which SVNS is capable of. Owing to the fuzziness, uncertainty and indeterminate nature of many practical problems in the real world, neutrosophy has found application in various fields including Social Network Analysis (Salama et al [13]), Decision-making problems (Ye [20], [21], [23], [22]), Image Processing (Cheng and Guo[3], Sengur and Guo[14], Zhang et al [26]), Social problems (Vasantha and Smarandache [17], [18]) etc. Liu et al, have applied neutrosophy to group decision problems and multiple attribute decision making problems ([9], [?], [10], [11]) etc. To provide more accuracy and precision to indeterminacy, the value of indeterminacy present in the neutrosophic set has been classified into two; based on membership; as indeterminacy leaning towards truth and as indeterminacy leaning towards false. This modified refined neutrosophic set was defined as Double Refined Indeterminacy Neutrosophic Set (DRINS) alias Double Valued Neutrosophic Set (DVNS) by Kandasamy [6] and Kandasamy and Smarandache [7].

To increase the accuracy, precision and to fit in the Likert's scale which is usually used in personality test; here the indeterminacy concept is divided into three, as indeterminacy leaning towards truth, indeterminacy and indeterminacy leaning towards false. This refined neutrosophic set is known as the Triple Refined Indeterminate Neutrosophic Sets (TRINS).

Consider an example from a personality test "You tend to sympathize with others". The person need not be forced to opt for a single choice; cause it is natural that the behaviour is dependent on several external and internal factors, varying from the person's mood to surrounding. A person might not always react in a particular way, in a particular scenario. When a person takes a objective personality test, he opts with a dominant choice, whereas in reality there are degrees of variations possible. Consider, for example in a person taking an personality test, there is a degree to which the person will strongly agree to the statement (say 0.7), will just agree (0.1), neither agree or disagree (0.05), will disagree (0.1) and will strongly disagree (0.05). While taking an objective personality test he/she is forced to opt for a single choice, thereby the degrees of membership of others are completely lost. Whereas using TRINS this statement is represented as $\langle 07, 0.1, 0.05, 0.1, 0.05\rangle$, it can be evaluated more reasonably; thereby giving very useful necessary precision to the result. All the various choices are captured thereby avoiding the preferential choice that is executed in the classical method.

The rest of the paper is organized as follows: Section two recalls some basic concepts about neutrosophy, TRINS, and the indeterminacy based personality test. Section three introduces the clustering algorithm using the Minimum Spanning Tree known as the TRIN-MST clustering algorithm based distance measures. Illustrative examples of clustering of different people based on the answers of the personality test are provided in the section four to show the working and efficiency of the algorithm. The comparison is given in section five. The last section provides the conclusions and further work.

2 Preliminaries / Basic Concepts

2.1 Neutrosophy and Single Valued Neutrosophic Set (SVNS)

Neutrosophy is a section of philosophy, familiarized by Smarandache [15], that analyses the beginning, property, and scope of neutralities, as well as their connections with various concepts. It studies a concept, event, theory, proposition, or entity, "A" in relation to its contrary, "Anti-A" and that which is not A, "Non-A", and that which is neither "A" nor "Anti-A", denoted by "Neut-A".

Definition 1 [15] Let X be a space of points (objects), with a generic element in X denoted by x. A neutrosophic set A in X is described by a truth $T_A(x)$, an indeterminacy $I_A(x)$, and a falsity $F_A(x)$ membership functions. The functions $T_A(x)$, $I_A(x)$, and $F_A(x)$ are nonstandard or real standard subsets of $]^-0, 1^+[$, under the rule $^-0 \leq supT_A(x) + supI_A(x) + supF_A(x) \leq 3^+$.

Definition 2 [19] Let X be a space of points (objects) with generic elements in X denoted by x. An Single Valued Neutrosophic Set (SVNS) A in X is characterized by truth $T_A(x)$, indeterminacy $I_A(x)$, and falsity $F_A(x)$ membership functions . For each point x in X, there are $T_A(x), I_A(x), F_A(x) \in [0, 1]$, and $0 \leq T_A(x) + I_A(x) + F_A(x) \leq 3$. Therefore, an SVNS A can be represented by $A = \{\langle x, T_A(x), I_A(x), F_A(x)\rangle \mid x \in X\}$. The following expressions are defined in [19] for SVNSs A, B:

- $A \in B \iff T_A(x) \leq T_B(x), I_A(x) \geq I_B(x), F_A(x) \geq F_B(x) \forall\ x$ in X.
- $A = B \iff A \subseteq B$ and $B \subseteq A$.
- $A^c = \{\langle x, F_A(x), 1 - I_A(x), T_A(x)\rangle | x \in X\}$.

The refined neutrosophic logic defined by [16] is as follows:

Definition 3 T can be split into several types of truths: $T_1, T_2, \ldots, T_p$, and I into several types of indeterminacies: I_1, I_2, $\ldots, I_r$, and F into many several of falsities: $F_1, F_2, \ldots, F_s$, where all $p, r, s \geq 1$ are integers, and $p + r + s = n$. In the same way, but all subcomponents T_j, I_k, F_l are not symbols, but subsets of [0, 1], $\forall j \in \{1, 2, \ldots, p\}$, $k \in \{1, 2, \ldots, r\}$, $l \in \{1, 2, \ldots, s\}$.

2.2 Triple Refined Indeterminacy Neutrosophic Set (TRINS)

In TRINS, the indeterminacy concept is divided into three memberships, as indeterminacy leaning towards truth, indeterminacy and indeterminacy leaning towards false. This division aids in increasing the accuracy and precision of the indeterminacy and to fit in the Likert's scale which is usually used in personality test. This refined neturosophic set is defined as the Triple Refined Indeterminate Neutrosophic Sets (TRINS) [8].

Definition 4 Let X be a space of points (objects) with generic elements in X denoted by x. A Triple Refined Indeterminate Neutrosophic Set (TRINS) A in X is characterized by truth $T_A(x)$, indeterminacy leaning towards truth $I_{TA}(x)$, indeterminacy $I_A(x)$, indeterminacy leaning towards falsity $I_{FA}(x)$, and falsity $F_A(x)$ membership functions. Each membership function has a weight $w_m \in [0, 5]$ associated with it. For each $x \in X$, there are

$$T_A(x),\ I_{TA}(x),\ I_A(x),\ I_{FA}(x),\ F_A(x) \in [0, 1],$$
$$w_T^m(T_A(x)),\ w_{I_T}^m(I_{TA}(x)),\ w_I^m(I_A(x)),\ w_{I_F}^m(I_{FA}(x)),\ w_F^m(F_A(x)) \in [0, 5],$$

and $0 \leq T_A(x) + I_{TA}(x) + I_A(x) + I_{FA}(x) + F_A(x) \leq 5$. Therefore, a TRINS A can be represented by

$$A = \{\langle x, T_A(x), I_{TA}(x), I_A(x), I_{FA}(x), F_A(x)\rangle \mid x \in X\}.$$

The generalized Triple Refined Indeterminate Neutrosophic (TRIN) weight is defined as follows:

$$\begin{aligned} w^m(A) = \{\sum_{i=1}^{n} \{w_T^m(T_A(x_i)) + w_{I_T}^m(I_{TA}(x_i)) + \\ w_I^m(I_A(x_i)) + w_{I_F}^m(I_{FA}(x_i)) + w_F^m(F_A(x_i))\} \end{aligned} \tag{1}$$

Example 1 Let $X = [x_1, x_2]$ where x_1 is question 1 (makes a list / relies on memory) and x_2 is question 2 (Sceptical/wants to believe) from Table 1. The values of x_1 and x_2 are in $[0, 1]$ and when the weight of the membership is applied the values of $w_m(x_1)$ and $w_m(x_2)$ are in $[1, 5]$. This is obtained from the questionnaire of the user.

Consider question 1, instead of a forced single choice; their option for question 1 would be a degree of "make list", a degree of indeterminacy choice towards "make list" , a degree of uncertain and indeterminate combination of making list and depending on memory, an degree of indeterminate choice more of replying on memory, and a degree of "relying on memory".

A is a TRINS of X defined by

$$A = \langle 0.0, 0.4, 0.1, 0.0, 0.5\rangle/x_1 + \langle 0.5, 0.1, 0.1, 0.1, 0.2\rangle/x_2.$$

The associated membership weights are $w_T^m = 1$, $w_{I_T}^m = 2$, $w_I^m = 3$, $w_{I_F}^m = 4$, $w_F^m = 5$. Then the weighted TRINS $w_T^m(T_A(x))$, $w_{I_T}^m(I_{TA}(x))$, $w_I^m(I_A(x))$, $w_{I_F}^m(I_{FA}(x))$, $w_F^m(F_A(x)) \in [0, 5]$, will be

$$A = \langle 0.0, 0.8, 0.3, 0.0, 1.5\rangle/x_1 + \langle 0.5, 0.2, 0.3, 0.4, 1.0\rangle/x_2.$$

All set theoretic operators like commutativity, associativity, distributivity, idempotency, absorption and the DeMorgan's Laws were defined over TRINs [8].

2.3 Indeterminacy based Personality test

There are several types of personality tests, the most common type is the objective personality tests. The most famously used personality test is the Myers-Briggs Type Indicator test.The Open Extended Jungian Type Scales (OEJTS) test [4] is an open source alternative to the Myers Briggs type indicator test based on which the indeterminacy based personality test was constructed.

The dichotomies used in [4] are

1. Introversion (I) vs. Extroversion (E); sometimes is described as a persons orientation, they either orient within themselves or to the outside world. Other times the focus is put more openly on social communication and interactions, with some stating that social activities and interactions tires introverts whereas it increases the energy level of extroverts.

Table 1 Sample Questionnaire of the Indeterminacy based Personality Test

Q		Scale weight 1 2 3 4 5	
Q_1	makes lists	□ □ □ □ □	relies on memory
Q_2	sceptical	□ □ □ □ □	wants to believe
Q_3	bored by time alone	□ □ □ □ □	needs time alone
Q_7	energetic	□ □ □ □ □	mellow
Q_{11}	works best in groups	□ □ □ □ □	works best alone
Q_{15}	worn out by parties	□ □ □ □ □	gets fired up by parties
Q_{19}	talks more	□ □ □ □ □	listens more
Q_{23}	stays at home	□ □ □ □ □	goes out on the town
Q_{27}	finds it difficult to yell very loudly	□ □ □ □ □	yelling to others . . . comes naturally
Q_{31}	perform in public	□ □ □ □ □	avoids public speaking

2. Sensing (S) vs. Intuition (N); how a person takes in information. Sensors generally focus on the five senses while intuitives focus on possibilities.
3. Feeling (F) vs. Thinking (T); is based on what a person uses to take their decisions: whether it is interpersonal considerations or through dispassionate logic.
4. Judging (J) vs. Perceiving (P); was a dichotomy added by Myers and Briggs to choose between the 2^{nd} and 3^{rd} pair of functions. Individuals who desire a organized lifestyle are supposed to use their judging functions (thinking and feeling) while individuals who prefer a flexible lifestyle use their sensing functions (intuition and sensing).

The indeterminacy based personality test (based on OEJTS) evaluates four scales. A sample questionnaire of the indeterminacy based personality test using TRINS will be as given in Table 1.

The user is expected to fill the degree accordingly.

Example 2 Consider question 19, the different options would be

1. a degree of "talks more",
2. a degree of indeterminacy choice towards talking more,
3. a degree of uncertain and indeterminate combination of talking more and listening more,
4. a degree of indeterminate choice more of listening, and
5. a degree of "listens more".

Suppose the user thinks and marks a degree of "talks more" is 0.0, a degree of indeterminate choice towards "talks more" is 0.4 , a degree of uncertain and indeterminate combination of talking and listening is 0.1, an degree of indeterminate choice more of listening 0.3, and a degree of "listens more" is 0.2.

A is a TRINS of $Q = \{q_1\}$ defined by

$$A = \langle 0.0, 0.4, 0.1, 0.3, 0.2\rangle / q_1.$$

Table 2 Grouping of based on"dichotomies" of the Indeterminacy based personality test

Aspect	TRINS	Question Grouping
Extrovert	E	$Q_E = \{Q_{15}, Q_{23}, Q_{27}\}$
Introvert	I	$Q_I = \{Q_3, Q_7, Q_{11}, Q_{19}, Q_{31}\}$
Sensing	S	$Q_S = \{Q_{24}, Q_{28}\}$
Intuition	N	$Q_N = \{Q_4, Q_8, Q_{12}, Q_{16}, Q_{20}, Q_{32}\}$
Feeling	F	$Q_F = \{Q_2, Q_{14}, Q_{18}, Q_{26}, Q_{30}\}$
Thinking	T	$Q_T = \{Q_6, Q_{10}, Q_{22}\}$
Judging	J	$Q_J = \{Q_{17}, Q_{25}\}$
Perceiving	P	$Q_P = \{Q_1, Q_5, Q_{13}, Q_{21}, Q_{29}\}$

When the weight of each membership is applied, the TRINS A becomes

$$A = \langle 0.0, 0.8, 0.3, 1.2, 1.0\rangle / q_1 \ ; \ w(A) = 3.3.$$

In the general test, a whole number value from 1 to 5 will be obtained, whereas in the indeterminacy based personality test an accurate value is obtained. Thus the accuracy of the test is evident.

Depending on the complete questionnaire with 32 questions the following grouping was done given in Table 2.

The weight of a TRINS E is given in Equations 1. The calculation for scoring is as follows:

$$IE = 30 - w(I) + w(E); SN = 12 - w(S) + w(N);$$

$$FT = 30 - w(F) + w(T); JP = 18 - w(J) + w(P).$$

The score results are based on the following rules:

1. If IE is greater than 24, extrovert (E), else introvert (I).
2. If SN is greater than 24, intuitive (N), else sensing (S).
3. If FT is greater than 24, thinking (T), else feeling (F).
4. If JP is greater than 24, perceiving (P), else judging (J).

3 TRIN-MST Clustering Algorithm using Distance Measures

3.1 Distance Measures of TRINS

The distance measures over TRINSs is defined in the following and the related algorithm for determining the distance is given. Consider two TRINSs A and B in a universe of discourse, $X = x_l, x_2, \ldots, x_n$, which are denoted by

$$A = \{\langle x_i, T_A(x_i), I_{TA}(x_i), I_A(x_i), I_{FA}(x_i), F_A(x_i)\rangle \mid x_i \in X\}, \text{ and}$$
$$B = \{\langle x_i, T_B(x_i), I_{TB}(x_i), I_B(x_i), I_{FB}(x_i), F_B(x_i)\rangle \mid x_i \in X\},$$

where $T_A(x_i), I_{TA}(x_i), I_A(x_i), I_{FA}(x_i), F_A(x_i), T_B(x_i), I_{TB}(x_i), I_B(x_i), I_{FB}(x_i), F_B(x_i) \in [0,5]$ for every $x_i \in X$. Let $w_i(i = 1,2,\ldots,n)$ be the weight of an element $x_i(i = 1,2,\ldots,n)$, with $w_i \geq 0(i = 1,2,\ldots,n)$ and $\sum_{i=1}^{n} w_i = 1$.

Then, the generalized TRIN weighted distance is defined as follows:

$$d_\lambda(A,B) = \{\frac{1}{5}\sum_{i=1}^{n} w_i[\mid T_A(x_i) - T_B(x_i) \mid^\lambda + \mid I_{TA}(x_i) - I_{TB}(x_i) \mid^\lambda + \mid I_A(x_i) - I_B(x_i) \mid^\lambda + \mid I_{FA}(x_i) - I_{FB}(x_i) \mid^\lambda + \mid F_A(x_i) - F_B(x_i) \mid^\lambda]\}^{1/\lambda} \quad (2)$$

where $\lambda > 0$.

Equation 2 reduces to the TRIN weighted Hamming distance and the TRIN weighted Euclidean distance, when $\lambda = 1, 2$, respectively. The TRIN weighted Hamming distance is given as

$$d_\lambda(A,B) = \frac{1}{5}\sum_{i=1}^{n} w_i[\mid T_A(x_i) - T_B(x_i) \mid + \mid I_{TA}(x_i) - I_{TB}(x_i) \mid + \mid I_A(x_i) - I_B(x_i) \mid + \mid I_{FA}(x_i) - I_{FB}(x_i) \mid + \mid F_A(x_i) - F_B(x_i) \mid] \quad (3)$$

where $\lambda = 1$ in Equation 2.

The TRIN weighted Euclidean distance is given as

$$d_\lambda(A,B) = \{\frac{1}{5}\sum_{i=1}^{n} w_i[\mid T_A(x_i) - T_B(x_i) \mid^2 + \mid I_{TA}(x_i) - I_{TB}(x_i) \mid^2 + \mid I_A(x_i) - I_B(x_i) \mid^2 + \mid I_{FA}(x_i) - I_{FB}(x_i) \mid^2 + \mid F_A(x_i) - F_B(x_i) \mid^2]\}^{1/2} \quad (4)$$

where $\lambda = 2$ in Equation 2.

The algorithm to obtain the generalized TRIN weighted distance $d_\lambda(A,B)$ is given in Algorithm 1.

The following proposition is given for the distance measure.

Proposition 1 *The generalized TRIN weighted distance $d_\lambda(A,B)$ for $\lambda > 0$ satisfies the following properties:*

1. *(Property 1)* $d_\lambda(A,B) \geq 0$;
2. *(Property 2)* $d_\lambda(A,B) = 0$ *if and only if* $A = B$;
3. *(Property 3)* $d_\lambda(A,B) = d_\lambda(B,A)$;
4. *(Property 4) If* $A \subseteq B \subseteq C, C$ *is an TRINS in* X, *then* $d_\lambda(A,C) \geq d_\lambda(A,B)$ *and* $d_\lambda(A,C) \geq d_\lambda(B,C)$.

It can be easily seen that $d_\lambda(A,B)$ satisfies the properties (Property 1) to (Property 4).

The TRIN distance matrix D is defined in the following.

Definition 5 Let $A_j(j = 1,2,\ldots,m)$ be a collection of m TRINs, then $D = (d_{ij})_{m\times m}$ is called a TRIN distance matrix, where $d_{ij} = d_\lambda(A_i, A_j)$ is the generalized TRIN distance between A_i and A_j, and its properties are as follows:

Algorithm 1: Generalized TRIN weighted distance $d_\lambda(A, B)$

Input: $X \leftarrow x_l, x_2, \ldots, x_n$,
$A \leftarrow \{\langle x_i, T_A(x_i), I_{TA}(x_i), I_A(x_i), I_{FA}(x_i), F_A(x_i)\rangle \mid x_i \in X\}$,
$B \leftarrow \{\langle x_i, T_B(x_i), I_{TB}(x_i), I_B(x_i), I_{FB}(x_i), F_A(x_i)\rangle \mid x_i \in X\}$,
$w_i(i = 1, 2, \ldots, n)$
Output: $d_\lambda(A, B)$
begin
  $d_\lambda A, B$
  $d_\lambda \leftarrow 0$
  for $i \leftarrow 1, n$ **do**
    $d_\lambda \leftarrow d_\lambda + w_i[\mid T_A(x_i) - T_B(x_i) \mid^\lambda + \mid I_{TA}(x_i) - I_{TB}(x_i) \mid^\lambda$
    $+ \mid I_A(x_i) - I_B(x_i) \mid^\lambda + \mid I_{FA}(x_i) - I_{FB}(x_i) \mid^\lambda + \mid F_A(x_i) - F_B(x_i) \mid^\lambda]$
  end
  $d_\lambda \leftarrow d_\lambda \; /5$
  $d_\lambda \leftarrow d_\lambda^{\{\frac{1}{\lambda}\}}$
end

Algorithm 2: TRIN weighted distance matrix D

Input: TVNS $A_1, \ldots \ldots, A_m$,
Output: Distance matrix D with elements d_{ij}
begin
  for $i \leftarrow 1, m$ **do**
    for $j \leftarrow 1, m$ **do**
      if $i = j$ **then**
        $d_{ij} \leftarrow 0$
      else
        $d_{ij} \leftarrow \{d_\lambda \; (A_i, A_j)\}$
    end
  end
end

1. $0 \leq d_{ij} \leq 5$ for all $i, j = 1, 2, \ldots, m$;
2. $d_{ij} = 0$ if and only if $A_i = A_j$;
3. $d_{ij} = d_{ji}$ for all $i, j = 1, 2, \ldots, m$.

The algorithm to calculate the TRIN weighted distance matrix D is given in Algorithm 2.

3.2 TRIN-MST Clustering Algorithm

Triple Refined Indeterminate Neutrosophic Minimum Spanning Tree (TRIN-MST) clustering algorithm is proposed as a generalization of the IFMST, SVN-MST and DVN-MST clustering algorithms here.

Consider $X = \{x_1, x_2, \ldots, x_n\}$ be an attribution space and the weight vector of an element $x_i(i = 1, 2, \ldots, n)$ be $w = \{w_1, w_2, \ldots, w_n\}$, where $w_i \geq 0(i = 1, 2, \ldots, n)$ and $\sum_{i=1}^{n} w_i = 1$. Let the m samples that need to be clustered be represented as $A_j(j = 1, 2, \ldots, m)$, as a collection of m TRINSs.

It is represented as: $A_j = \{\langle x_i, T_{A_j}(x_j), I_{TA_j}(x_j), I_{A_j}(x_j), I_{FA_j}(x_j), F_{A_j}(x_j)\rangle \mid x_j \in X\}$. The Triple Refined Indeterminate Neutrosophic Minimum Spanning Tree (TRIN-MST) clustering algorithm is provided in Algorithm 3, the detailed description is as follows:

Algorithm 3: Triple Refined Indeterminate Neutrosophic Minimum Spanning Tree (TRIN-MST) Clustering algorithm

```
Input: D = (d_ij)_{m×m}
Output: MST S and Clusters
begin
    Step 1: Calculation of distance matrix D(A_1, ..., A_m)
    // Distance matrix D is from Algo 2
    Step 2: Create TRIN graph G(V, E)
    for i ← 1, m do
        for j ← 1, m do
            if i != j then
                Draw the edge between A_i and A_j with d_ij
            end
        end
    end
    Step 3: Compute the MST of G(V, E): // Using Kruskal's algorithm
    Sort all the edges in increasing order of weight in E.
    while No. of edges in subgraph S of G < (V − 1) do
        Select the smallest edge (v_i, v_j).
        Delete (v_i, v_j) from E
        if (v_i, v_j) forms a cycle with spanning tree S then
            Discard the edge v_i, v_j
        else
            Include the edge v_i, v_j in S
        end
    end
    S is the MST of G(V, E).
    Step 4: Perform clustering of S
    for i ← 1, m do
        for j ← 1, m do
            if d_ij ≥ r then // r is the threshold
                Disconnect edge
            else
                Edge is not disconnected
            end
        end
    end
    Results in clusters automatically; it is tabulated
end
```

Step 1: Calculation of the distance matrix $D = d_{ij} = d_\lambda(A_i, A_j)$ is by using Algorithm 2 and taking $\lambda = 2$. The TRIN distance matrix $D = (d_{ij})_{m\times m}$ obtained is:

$$D = \begin{bmatrix} 0 & d_{12} & \dots & d_{1m} \\ \vdots & \vdots & & \vdots \\ d_{m1} & d_{m2} & \dots & 0 \end{bmatrix}.$$

Step 2: The TRIN graph $G(V,E)$ where every edge between A_i and $A_j(i,j = 1,2,\ldots,m)$ is assigned the TRIN weighted distance $d_{ij} \in D$, it represents the dissimilarity degree between the samples A_i and A_j.

Step 3: Construct the MST of $G(V,E)$.

1. In increasing order of weight of edges of $G(V,E)$ the sorting is done.
2. Take an empty subgraph S of $G(V,E)$ and select the smallest weighted edge e.
3. The smallest edge e is added to S and deleted from the sorted list.
4. The next smallest edge is selected and if no cycle is formed in S it is added to S and deleted from the list.
5. The process (iv) is repeated until the subgraph S has $(m-1)$ edges.

Thus, the MST of the TRIN graph $G(V,E)$ is obtained.

Step 4: Select a threshold r and disconnect all the edges of the MST with weights greater than r to obtain a certain number of clusters, list it as a table.

The clustering results induced by the subtrees do not depend on some particular MST [25], [27].

4 Illustrative Examples

To demonstrate the effectiveness of the proposed TRIN-MST clustering algorithm in the real world applications, a descriptive example is presented. The results of the indeterminacy based personality test conducted for eight different people which are represented by TRINs is clustered using the TRIN-MST clustering algorithm.

Example 3 The answers of the indeterminacy based personality test of eight different people $P_j(j = 1,2,\ldots,8)$ is taken for clustering. Six evaluation questions $(Q_7, Q_{11}, Q_{15}, Q_{19}, Q_{23}, Q_{27})$ were used for each person, given in Table 1. For the evaluation purpose, only questions related to the Extroversion (E) vs. Introversion (I) dichotomy has been considered. The questionnaire has been altered according so as to enable the using of distance measures.

The answers of the eight people $P_j(j = 1,2,\ldots,8)$ represented using TRIN is as follows:

$$
\begin{aligned}
P_1 = \{&\langle x_1, 0.9, .02, .05, 0.03, 0.0\rangle, \langle x_2, 0.8, 0.1, .05, .05, 0.0\rangle,\\
&\langle x_3, 0.7, 0.1, 0.1, 0.0, 0.1\rangle, \langle x_4, 0.8, 0.0, 0.0, 0.1, 0.1\rangle,\\
&\langle x_5, 0.75, 0.05, 0.1, 0.1, 0.0\rangle, \langle x_6, 0.8, 0.1, 0.05, 0.0, 0.05\rangle\},\\
P_2 = \{&\langle x_1, 0.8, 0.15, 0.0, 0.0, 0.05\rangle, \langle x_2, 0.7, 0.2, 0.1, 0.0, 0.0\rangle,\\
&\langle x_3, 0.75, 0.05, 0.1, 0.0, 0.1\rangle, \langle x_4, 0.7, 0.2, 0.0, 0.0, 0.1\rangle,\\
&\langle x_5, 0.8, 0.1, 0.1, 0.0, 0.0\rangle, \langle x_6, 0.6, 0.3, 0.1, 0.0, 0.0\rangle\},\\
P_3 = \{&\langle x_1, 0.5, 0.2, 0.1, 0.1, 0.1\rangle, \langle x_2, 0.6, 0.2, 0.1, 0.05, 0.05\rangle,\\
&\langle x_3, 0.5, 0.3, 0.05, 0.05, 0.1\rangle, \langle x_4, 0.75, 0.1, 0.05, 0.1, 0.0\rangle,\\
&\langle x_5, 0.5, 0.1, 0.3, 0.0, 0.1\rangle, \langle x_6, 0.6, 0.2, 0.0, 0.0, 0.2\rangle\},\\
P_4 = \{&\langle x_1, 0.0, 0.1, 0.0, 0.0, 0.9\rangle, \langle x_2, 0.0, 0.0, 0.0, 0.1, 0.9\rangle,\\
&\langle x_3, 0.1, 0.0, 0.0, 0.0, 0.9\rangle, \langle x_4, 0.0, 0.0, 0.0, 0.1, 0.9\rangle,\\
&\langle x_5, 0.0, 0.0, 0.1, 0.0, 0.9\rangle, \langle x_6, 0.0, 0.0, 0.0, 0.1, 0.9\rangle\},\\
P_5 = \{&\langle x_1, 0.7, 0.3, 0.0, 0.0, 0.0\rangle, \langle x_2, 0.6, 0.2, 0.1, 0.0, 0.1\rangle,\\
&\langle x_3, 0.5, 0.1, 0.2, 0.1, 0.1\rangle, \langle x_4, 0.7, 0.0, 0.1, 0.1, 0.1\rangle,\\
&\langle x_5, 0.5, 0.2, 0.1, 0.1, 0.1\rangle, \langle x_6, 0.3, 0.2, 0.2, 0.2, 0.1\rangle\},\\
P_6 = \{&\langle x_1, 0.5, 0.2, 0.1, 0.0, 0.1\rangle, \langle x_2, 0.0, 0.1, 0.1, 0.7, 0.1\rangle,\\
&\langle x_3, 0.0, 0.1, 0.1, 0.5, 0.3\rangle, \langle x_4, 0.1, 0.0, 0.1, 0.1, 0.7\rangle,\\
&\langle x_5, 0.0, 0.1, 0.2, 0.6, 0.1\rangle, \langle x_6, 0.0, 0.1, 0.1, 0.7, 0.1\rangle\},\\
P_7 = \{&\langle x_1, 0.4, 0.3, 0.2, 0.0, 0.1\rangle, \langle x_2, 0.5, 0.1, 0.2, 0.1, 0.1\rangle,\\
&\langle x_3, 0.6, 0.1, 0.1, 0.1, 0.1\rangle, \langle x_4, 0.5, 0.2, 0.1, 0.1, 0.1\rangle,\\
&\langle x_5, 0.5, 0.1, 0.1, 0.2, 0.1\rangle, \langle x_6, 0.7, 0.2, 0.0, 0.0, 0.1\rangle\},\\
P_8 = \{&\langle x_1, 0.1, 0.1, 0.1, 0.2, 0.5\rangle, \langle x_2, 0.0, 0.1, 0.1, 0.1, 0.7\rangle,\\
&\langle x_3, 0.0, 0.0, 0.0, 0.3, 0.7\rangle, \langle x_4, 0.1, 0.0, 0.2, 0.4, 0.3\rangle,\\
&\langle x_5, 0.1, 0.0, 0.0, 0.1, 0.8\rangle, \langle x_6, 0.2, 0.0, 0.0, 0.1, 0.7\rangle\}
\end{aligned}
$$

The weight vector is taken uniformly for the attribute $x_i (i = 1, 2, \ldots, 6)$ is given by $w = (0.167,\ 0.167,\ 0.167,\ 0.167,\ 0.167,\ 0.167)^T$. The TRIN-MST clustering algorithm provided in Algorithm 3 is used to group the eight people of $P_j (j = 1, 2, \ldots, 8)$ into clusters. In this sample P_1 and P_2 are extroverts and P_4 and P_8 are introverts according to the indeterminacy based personality test results.

The stepwise working of the TRIN-MST clustering algorithm is as follows:

Step 1: The distance matrix $D = d_{ij} = d_\lambda(P_i, P_j)$ is calculated by using Algorithm 2 (taking $\lambda = 2$). The TRIN distance matrix $D = (d_{ij})_{m\times m}$ is

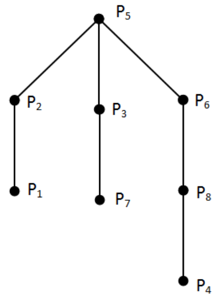

Fig. 1 TRIN-MST S of graph G

obtained as follows:

$$D = \begin{bmatrix} 0 & 0.0860 & 0.1384 & 0.5214 & 0.1512 & 0.4035 & 0.1588 & 0.4305 \\ 0.086 & 0 & 0.1284 & 0.507 & 0.1263 & 0.3938 & 0.139 & 0.4213 \\ 0.1384 & 0.1284 & 0 & 0.4534 & 0.1175 & 0.3539 & 0.1024 & 0.3536 \\ 0.5214 & 0.507 & 0.4534 & 0 & 0.453 & 0.3944 & 0.4394 & 0.186 \\ 0.1512 & 0.1263 & 0.1175 & 0.453 & 0 & 0.3155 & 0.134 & 0.3558 \\ 0.4035 & 0.3938 & 0.3539 & 0.3944 & 0.3155 & 0 & 0.3258 & 0.3238 \\ 0.1588 & 0.139 & 0.1024 & 0.4394 & 0.134 & 0.3258 & 0 & 0.3375 \\ 0.4305 & 0.4213 & 0.3536 & 0.186 & 0.3558 & 0.3238 & 0.3375 & 0 \end{bmatrix}.$$

Step 2: The TRIN graph $G(V, E)$ where every edge between P_i and $P_j (i, j = 1, 2, \ldots, 8)$ is assigned the TRIN weighted distance d_{ij}. An element d_{ij} of the TRIN distance matrix $D = (d_{ij})_{m \times m}$, represents the dissimilarity degree between the samples P_i and P_j.

Step 3: Construction of the MST of the TRIN graph $G(V, E)$, is done as follows:

1. The distances of edges of G sorted in increasing order by weights is: $d_{12} \leq d_{37} \leq d_{35} \leq d_{25} \leq d_{23} \leq d_{57} \leq d_{72} \leq d_{51} \leq d_{71} \leq d_{84} \leq d_{65} \leq d_{68} \leq d_{67} \leq d_{87} \leq d_{83} \leq d_{63} \leq d_{86} \leq d_{62} \leq d_{64}$.
2. An empty subgraph S of G is taken and the edge e with the smallest weight to S is added, if the end points of e are disconnected in S. Here the smallest edge is between P_1 and P_2; $d_{12} = 0.086$, it is added to S and deleted from the sorted list.
3. The next smallest edge is selected from G and if no cycle is formed in S it is deleted from the list and added to S.
4. Process (4) is repeated until the subgraph S has $(7 - 1)$ edges or spans eight nodes.

The MST of the TRIN graph $G(V, E)$ is obtained, it is illustrated in Figure 1.

Step 4: A threshold r is selected and all the edges of the MST with weights greater than r are disconnected to obtain a certain number of subtrees (clusters), as listed in Table 3.

Table 3 Clustering results of the eight different people personality using TRIN-MST clustering algorithm

Threshold r	Corresponding clustering result
$r = d_{68} = 0.3238$	$\{P_1, P_2, P_3, P_5, P_6, P_7\}, \{P_4, P_8\}$
$r = d_{56} = 0.3155$	$\{P_1, P_2, P_3, P_5, P_7\}, \{P_4, P_8\}, \{P_6\}$
$r = d_{84} = 0.186$	$\{P_1, P_2, P_3, P_5, P_7\}, \{P_4\}, \{P_8\}, \{P_6\}$
$r = d_{25} = 0.1263$	$\{P_1, P_2\}, \{P_3, P_5, P_7\}, \{P_4\}, \{P_6\}, \{P_8\}$
$r = d_{53} = 0.1175$	$\{P_1, P_2\}, \{P_3, P_7\}, \{P_4\}, \{P_5\}, \{P_6\}, \{P_8\}$
$r = d_{37} = 0.1024$	$\{P_1, P_2\}, \{P_3\}, \{P_4\}, \{P_5\}, \{P_6\}, \{P_7\}, \{P_8\}$
$r = d_{12} = 0.086$	$\{P_1\}, \{P_2\}, \{P_3\}, \{P_4\}, \{P_6, P_5\}, \{P_7\}, \{P_8\}$
$r = 0$	$\{P_1\}, \{P_2\}, \{P_3\}, \{P_4\}, \{P_5\}, \{P_6\}, \{P_7\}, \{P_8\}$

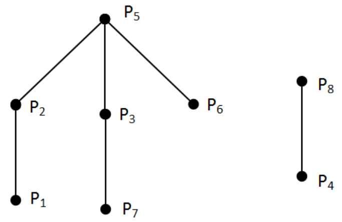

Fig. 2 Clusters of S for threshold $r = 0.3238$

Consider taking the threshold as $r = 0.3238$, then the clusters that are formed are $\{P_1, P_2, P_3, P_5, P_6, P_7\}$ and $\{P_4, P_8\}$. The resulting clusters are shown in Figure 2.

The results of the clustering algorithm clearly shows when the threshold r is 0.3238 the clusters are of Extroversion (E) vs. Introversion (I), it is seen that P_4 and P_8 are introverts and the rest are extroverts.

5 Comparison

The existing classic personality test forces the test taker to select only one option and it is mostly what the user thinks he/she does often. The other options are lost to the test taker. It fails to capture the complete picture realistically. The dominant choice is selected, the selection might have very small margin. In such cases the accuracy of the test fails. Whereas when the indeterminacy based personality test is considered, it provides five different options to the test taker using TRINS for representing the choice. It is important to understand why TRINS is the best mechanism for this kind of personality test. TRINS is an instance of a neutrosophic set, which approaches the problem more logically with accuracy and precision to represent the existing uncertainty, imprecise, incomplete, and inconsistent information. It has the additional feature of being able to describe with more sensitivity the indeterminate and inconsistent information. TRINS alone can give scope for a person to express accurately the exact realistic choices instead of opting for a dominant choice. While, the SVNS, fuzzy set and intuitionistic fuzzy set fail to describe the existing indeterminacy this accurately. The clustering of the personality using the TRIN-MST clustering algorithm is capable of clustering people according to their personality with more accuracy and precision.

6 Conclusions

The user of any objective personality test (MBTI or OEJTS) is forced to select only one option which is generally the most dominant choice/option. The other options and their related influence are totally lost. Indeterminacy based personality test which gives the user the freedom to select and give weightage to every option was used to provide better accurate and realistic results. There is no forced option/selection done here. This test is better than the existing personality test since this test captures the indeterminate and inconsistent information from the user.

In this paper, the Triple Refined Indeterminate Neutrosophic Set Minimum Spanning Tree (TRIN-MST) clustering algorithm to cluster TRINS information was proposed. An illustrative example of clustering using the TRIN-MST clustering algorithm of the indeterminacy based personality test results of eight people was carried out. Similar personalities were clustered together, the clusters of extroverts and introverts was clearly seen in the illustration provided.

References

1. Atanassov, K., Gargov, G.: Interval valued intuitionistic fuzzy sets. Fuzzy sets and systems **31**(3), 343–349 (1989)
2. Atanassov, K.T.: Intuitionistic fuzzy sets. Fuzzy sets and Systems **20**(1), 87–96 (1986)
3. Cheng, H.D., Guo, Y.: A new neutrosophic approach to image thresholding. New Mathematics and Natural Computation **4**(03), 291–308 (2008)
4. Jorgenson, E.: Development of the open extended jungian type scales 1.2. Creative Commons Attribution NonCommercial-Share Alike 4.0 International License pp. 1–13 (2015)
5. Jung, C.G.: Psychological types. Routledge (2014)
6. Kandasamy, I.: Double-valued neutrosophic sets, their minimum spanning trees, and clustering algorithm (2016). Accepted for publication in Journal of Intelligent Systems
7. Kandasamy, I., Smarandache: Multicriteria decision making using double refined indeterminacy neutrosophic cross entropy and indeterminacy based cross entropy. Applied Mechanics and Materials **859**, 129–143 (2016)
8. Kandasamy, I., Smarandache: Triple refined indeterminate neutrosophic sets for personality classification (2016). Accepted for IEEE Symposium Series on Computational Intelligence 2016
9. Liu, P., Chu, Y., Li, Y., Chen, Y.: Some generalized neutrosophic number hamacher aggregation operators and their application to group decision making. International Journal of Fuzzy Systems **16**(2), 242–255 (2014)
10. Liu, P., Shi, L.: The generalized hybrid weighted average operator based on interval neutrosophic hesitant set and its application to multiple attribute decision making. Neural Computing and Applications **26**(2), 457–471 (2015)

11. Liu, P., Tang, G.: Some power generalized aggregation operators based on the interval neutrosophic sets and their application to decision making. Journal of Intelligent & Fuzzy Systems (Preprint), 1–12 (2016)
12. Myers, I.B.: The myers-briggs type indicator. Consulting Psychologists Press Palo Alto, CA (1962)
13. Salama, A., Haitham, A., Manie, A., Lotfy, M.: Utilizing neutrosophic set in social network analysis e-learning systems. International Journal of Information Science and Intelligent System **3**(2) (2014)
14. Sengur, A., Guo, Y.: Color texture image segmentation based on neutrosophic set and wavelet transformation. Computer Vision and Image Understanding **115**(8), 1134–1144 (2011)
15. Smarandache, F.: A Unifying Field in Logics: Neutrosophic Logic. Neutrosophy, Neutrosophic Set, Probability, and Statistics. American Research Press, Rehoboth (2000)
16. Smarandache, F.: n-valued refined neutrosophic logic and its applications in physics. Progress in Physics **4**, 143–146 (2013)
17. Vasantha, W., Smarandache, F.: Fuzzy cognitive maps and neutrosophic cognitive maps. Infinite Study (2003)
18. Vasantha, W., Smarandache, F.: Analysis of social aspects of migrant labourers living with hiv/aids using fuzzy theory and neutrosophic cognitive maps: With special reference to rural tamil nadu in india. arXiv preprint math/0406304 (2004)
19. Wang, H., Smarandache, F., Zhang, Y., Sunderraman, R.: Single valued neutrosophic sets. Review p. 10 (2010)
20. Ye, J.: Multicriteria decision-making method using the correlation coefficient under single-valued neutrosophic environment. International Journal of General Systems **42**(4), 386–394 (2013)
21. Ye, J.: A multicriteria decision-making method using aggregation operators for simplified neutrosophic sets. Journal of Intelligent & Fuzzy Systems **26**(5), 2459–2466 (2014)
22. Ye, J.: Similarity measures between interval neutrosophic sets and their applications in multicriteria decision-making. Journal of Intelligent & Fuzzy Systems **26**(1), 165–172 (2014)
23. Ye, J.: Single valued neutrosophic cross-entropy for multicriteria decision making problems. Applied Mathematical Modelling **38**(3), 1170–1175 (2014)
24. Zadeh, L.A.: Fuzzy sets. Information and control **8**(3), 338–353 (1965)
25. Zahn, C.T.: Graph-theoretical methods for detecting and describing gestalt clusters. Computers, IEEE Transactions on **100**(1), 68–86 (1971)
26. Zhang, M., Zhang, L., Cheng, H.: A neutrosophic approach to image segmentation based on watershed method. Signal Processing **90**(5), 1510–1517 (2010)
27. Zhang, X., Xu, Z.: An mst cluster analysis method under hesitant fuzzy environment. Control and Cybernetics **41**(3), 645–666 (2012)