Article

Single valued Neutrosophic clustering algorithm Based on Tsallis Entropy Maximization

Qiaoyan Li [1], Yingcang Ma[1]* and Shuangwu Zhu[2]

[1] School of Science, Xi'an Polytechnic University, Xi'an 710048, P. R. China; liqiaoyan@xpu.edu.cn
[2] School of Textile and Materials, Xi'an Polytechnic University, Xi'an 710048, P. R. China; zhushuangwu@263.net
* Correspondence: mayingcang@126.com

Academic Editor: name
Version July 10, 2018 submitted to Journal Not Specified

Abstract: Data clustering is an important field in pattern recognition and machine learning. Fuzzy c-means is considered as a useful tool in data clustering. Neutrosophic set, which is extension of fuzzy set, has received extensive attention in solving many real life problems of uncertainty, inaccuracy, incompleteness, inconsistency and uncertainty. In this paper, we propose a new clustering algorithm, single valued neutrosophic clustering algorithm, which is inspired from fuzzy c-means, picture fuzzy clustering and the single valued neutrosophic set. A novel suitable objective function, which is depicted as a constrained minimization problem based on single valued neutrosophic set, is built and the Lagrange multiplier method is used to solve the objective function. We do several experiments with some benchmark data sets, and we also apply the method to image segmentation by Lena image. The experimental results show that the given algorithm can be considered as a promising tool for data clustering and image processing.

Keywords: single valued neutrosophic set; fuzzy c-means; picture fuzzy clustering; Tsallis entropy

1. Introduction

Data clustering is one of the most important topics in pattern recognition, machine learning and data mining. Generally, data clustering is the task of grouping a set of objects in such a way that objects in the same group (cluster) are more similar to each other than to those in other groups (clusters). In the past decades, lots of clustering algorithms have been proposed, such as k-means clustering[1], hierarchical clustering[2], spectral clustering[3], etc. The clustering technique has been used in many fields, including image analysis, bioinformatics, data compression, computer graphics, and so on[4–6].

The k-means algorithm is one of the typical hard clustering algorithms that widely used in real applications due to its simplicity and efficiency. Unlike the hard clustering, the fuzzy c-means (FCM) algorithm[7] is one of the most popular soft clustering algorithms, that is each data point belongs to a cluster to some degree that is specified by a membership degrees in $[0, 1]$, and the sum of over the clusters for each data be equal to 1. In recent years, many improved algorithms for FCM are proposed. There are three main ways to build the clustering algorithm. First, extensions of the traditional fuzzy sets. In this way, numerous fuzzy clustering algorithms based on the extension fuzzy sets, such as intuitionistic fuzzy set, type-2 fuzzy set, etc., are built. By replacing traditional fuzzy sets to intuitionistic fuzzy set, Chaira introduced the intuitionistic fuzzy c-means clustering method (IFCM) in [8], which integrated the intuitionistic fuzzy entropy with the objective function. Hwang and Rhee proposed Type-2 fuzzy sets (T2FS) in [9], which aim to design and manage uncertainty for fuzzifier m. Thong and Son proposed picture fuzzy clustering based on picture fuzzy set (PFS) in [10]. Second, Kernel-based method is applied to improve the fuzzy clustering quality. For example, Graves and Pedrycz present a kernel version of the FCM algorithm namely KFCM in[11]. Ramathilagam etl.

analysis the Lung Cancer database by incorporating hyper tangent kernel function[12]. Third, Adding regularization terms to the objective function is used to improve the clustering quality. For example, Yasuda proposed an approach to FCM based on entropy maximization in [13]. Of course, we can use them together to obtain more clustering quality.

Neutrosophic set is proposed by Smarandache [14] in order to deal with real-world problems. Now, neutrosophic set is gaining significant attention in solving many real life problems that involve uncertainty, impreciseness, incompleteness, inconsistent, and indeterminacy. A neutrosophic set has three membership functions and each membership degree is a real standard or non-standard subset of the nonstandard unit interval $]0^-, 1^+[$. Wang et al. [15] introduced single valued neutrosophic sets (SVNSs) which is a extension of intuitionistic fuzzy sets. Moreover, the three membership functions are independent and their values belong to the unit interval $[0, 1]$. In recent years, the studies of the SVNSs have been developed rapidly. Such as, Majumdar and Samanta [16] studied similarity and entropy of SVNSs. Ye [17]proposed correlation coefficients of SVNSs, and applied it to single valued neutrosophic decision-making problems, etc. Zhang etl. in [18] propose a new definition of inclusion relation of neutrosophic sets (call it type-3 inclusion relation), and a new method of ranking of neutrosophic sets is given. Zhang etl. in [19] study neutrosophic duplet sets, neutrosophic duplet semi-groups, and cancellable neutrosophic triplet groups.

The clustering methods by neutrosophic set have some studies. In paper [20], Ye propose a single-valued neutrosophic minimum spanning tree (SVNMST) clustering algorithm, and he also introduce single-valued neutrosophic clustering methods based on similarity measures between SVNSs [21]. Guo and Sengur give neutrosophic *c*-means clustering algorithm[22], which is inspired from FCM and the neutrosophic set framework. Thong and Son did significant work for the clustering based on PFS. In [10], a picture fuzzy clustering algorithm, called FC-PFS is proposed. In order to determine the number of clusters, they built an automatically determined the most suitable number of clusters based on particle swarm optimization and picture composite cardinality for a dataset[23]. They also extend the picture fuzzy clustering algorithm for complex data[24]. Unlike the method in[10], Son present a novel distributed picture fuzzy clustering method on picture fuzzy set [25]. We can note that the basic ideas of the fuzzy set, the intuitionistic fuzzy set and the SVNS are consistent in the data clustering, but there are differences in the representation of the objects, so that the clustering objective functions are different. Thus, the more adequate description can be better used for clustering. Inspired from FCM, FC-PFS, SVNS and maximization entropy method, we propose a new clustering algorithm, single valued neutrosophic clustering algorithm based on Tsallis entropy maximization(SVNCA-TEM) in this paper, and the experimental results show that the proposed algorithm can be considered as a promising tool for data clustering and image processing.

The rest of paper is organized as follows. Section 2 shows the related work on FCM, IFC and FC-PFS. Section 3 introduces the proposed method and using the Lagrange multiplier method to solve the objective function. The experiments on some benchmark UCI data set indicate that the proposed algorithm can be considered as a useful tool for data clustering and image processing in Section 4. The last section draws the conclusions .

2. Related works

In general, suppose data set $D = \{X_1, X_2, \cdots, X_n\}$ include n data points, each data $X_i = \{x_{i1}; x_{i2}; \cdots; x_{id}\} \in R^d$ is a d-dim feature vector. The aim of clustering is get k disjoint clusters $\{C_j|, j = 1, 2, \cdots, k\}$, and satisfies $C_{j'} \cap_{j' \neq j} C_j = \emptyset$ and $D = \cup_{j=1}^{k} C_j$. In the following, we will briefly introduce three fuzzy clustering methods, which are FCM, IFC and FC-PFS.

2.1. Fuzzy c-means

The FCM was proposed in 1984 [7] . FCM is a data clustering technique wherein each data point belongs to a cluster to some degree that is specified by a membership grade. A data point X_i to cluster C_j denoted by the term μ_{ij}, which shows the fuzzy membership degree of the i-th data point in the j-th cluster. We use $V = \{V_1, V_2, \cdots, V_k\}$ to describe the cluster centroids of the clusters and $V_j \in R^d$ is the cluster centroid of C_j. The FCM is based on minimization of the following objective function

$$J = \sum_{i=1}^{n} \sum_{j=1}^{k} u_{ij}^{m} \|x_i - V_j\|^2 \tag{1}$$

where m represents the fuzzy parameter and $m \geq 1$. The constraints for (1) are,

$$\sum_{l=1}^{k} \mu_{il} = 1, \ \mu_{ij} \in [0,1], \ \ i = 1, 2, \cdots, n, \ \ j = 1, 2, \cdots, k. \tag{2}$$

Using the Lagrangian method, the iteration scheme to calculate cluster centroids V_j and the fuzzy membership degrees μ_{ij} of the objective function (1) as follows.

$$V_j = \frac{\sum_{i=1}^{n} \mu_{ij}^{m} X_i}{\sum_{i=1}^{n} \mu_{ij}^{m}}, \ j = 1, 2, \cdots, k. \tag{3}$$

$$\mu_{ij} = (\sum_{l=1}^{k} (\frac{\|X_i - V_j\|}{\|X_i - V_l\|})^{\frac{2}{m-1}})^{-1}. \ i = 1, 2, \cdots, n. \ j = 1, 2, \cdots, k. \tag{4}$$

The iteration will not stop until reach the maximum iterations or $|J^{(t)} - J^{(t-1)}| < \epsilon$, where $J^{(t)}$ and $J^{(t-1)}$ are the objection function value at (t)-th and $(t-1)$-th iterations, and ϵ is a termination criterion between 0 and 0.1. This procedure converges to a local minimum or a saddle point of J. Finally, each data point is assigned into different cluster according to the fuzzy membership value, that is X_i belongs to C_l if $\mu_{il} = \max(\mu_{i1}, \mu_{i2}, \cdots, \mu_{ik})$.

2.2. Intuitionistic fuzzy clustering

The intuitionistic fuzzy set is an extension of fuzzy sets. Chaira proposed intuitionistic fuzzy clustering (IFC)[8], which is integrated the intuitionistic fuzzy entropy with the objective function of FCM. The objective function of IFS is:

$$J = \sum_{i=1}^{n} \sum_{j=1}^{k} \mu_{ij}^{m} \|X_i - V_j\|^2 + \sum_{j=1}^{k} \pi_j^{*} e^{1-\pi_j^{*}}. \tag{5}$$

where $\pi_j^{*} = \frac{1}{n} \sum_{i=1}^{n} \pi_{ij}$, and π_{ij} is hesitation degree of X_i for C_j. The constraints of IFC are similar to (2). Hesitation degree π_{ik} is initially calculated using the following form:

$$\pi_{ij} = 1 - \mu_{ij} - (1 - u_{ij}^{\alpha})^{1/\alpha}, \ \ \text{where} \ \ \alpha \in [0,1], \tag{6}$$

and the intuitionistic fuzzy membership values are obtained as follows

$$\mu_{ij}^{*} = \mu_{ij} + \pi_{ij}, \tag{7}$$

where $\mu_{ij}^{*}(\mu_{ij})$ denotes the intuitionistic (conventional) fuzzy membership of the i-th data in j-th class. The modified cluster centroid is:

$$V_j = \frac{\sum_{i=1}^{n} \mu_{ij}^{*m} X_i}{\sum_{i=1}^{n} \mu_{ij}^{*m}}, \ \ j = 1, 2, \cdots, k. \tag{8}$$

The iteration will not stop until reach the maximum iterations or the difference between $\mu_{ij}^{*(t)}$ and $\mu_{ij}^{*(t-1)}$ is not larger than a pre-defined threshold ϵ, that is $\max_{i,j} |\mu_{ij}^{*(t)} - \mu_{ij}^{*(t-1)}| < \epsilon$.

2.3. Picture fuzzy clustering

In [26] Cuong introduced the picture fuzzy set (is also called standard neutrosophic set [27]), which is defined on a non-empty set S, $\dot{A} = \{\langle x, \mu_{\dot{A}}(x), \eta_{\dot{A}}(x), \gamma_{\dot{A}}(x)\rangle | x \in S\}$, where $\mu_{\dot{A}}(x)$ is the positive degree of each element $x \in X$, $\eta_{\dot{A}}(x)$ is the neutral degree and $\gamma_{\dot{A}}(x)$ is the negative degree satisfying the constraints,

$$\begin{cases} \mu_{\dot{A}(x)}, \eta_{\dot{A}(x)}, \gamma_{\dot{A}(x)} \in [0,1], & \forall x \in S \\ \mu_{\dot{A}(x)} + \eta_{\dot{A}(x)} + \gamma_{\dot{A}(x)} \leq 1, & \forall x \in S \end{cases} \tag{9}$$

The refusal degree of an element is calculated as

$$\xi_{\dot{A}}(x) = 1 - (\mu_{\dot{A}}(x) + \eta_{\dot{A}}(x) + \gamma_{\dot{A}}(x)), \forall x \in S. \tag{10}$$

In paper [10] Thong and Son propose picture fuzzy clustering(FC-PFS), which is related to neutrosophic clustering. The objective function is:

$$J = \sum_{i=1}^{n}\sum_{j=1}^{k}(\mu_{ij}(2-\xi_{ij}))^m \|X_i - V_j\|^2 + \sum_{i=1}^{n}\sum_{j=1}^{k}\eta_{ij}(\log \eta_{ij} + \xi_{ij}). \tag{11}$$

where $i = 1, \cdots, n$, $j = 1, \cdots, k$. μ_{ij}, η_{ij} and ξ_{ij} are the positive, neutral and refusal degrees respectively that each data point X_i belongs to cluster C_j. Denote μ, η and ξ being the matrices whose elements are μ_{ij}, η_{ij} and ξ_{ij} respectively. The constraints for FC-PFS are defined as follows:

$$\begin{cases} u_{ij}, \eta_{ij}, \xi_{ij} \in [0,1], \\ u_{ij} + \eta_{ij} + \xi_{ij} \leq 1, \\ \sum_{l=1}^{k}(u_{il}(2-\xi_{il})) = 1, \\ \sum_{l=1}^{k}(\eta_{il} + \xi_{il}/k) = 1. \end{cases} \tag{12}$$

Using the Lagranian multiplier method, the iteration scheme to calculate $\mu_{ij}, \eta_{ij}, \xi_{ij}$ and V_j for the model (11,12) as the following equations:

$$\xi_{ij} = 1 - (\mu_{ij} + \eta_{ij}) - (1 - (\mu_{ij} + \eta_{ij})^{\alpha})^{1/\alpha}, \text{ where } \alpha \in [0,1], \ (i = 1, \cdots, n, j = 1, \cdots, k), \tag{13}$$

$$\mu_{ij} = \frac{1}{\sum_{l=1}^{k}(2-\xi_{ij})(\frac{\|X_i - V_j\|}{\|X_i - V_l\|})^{\frac{2}{m-1}}}, \ (i = 1, \cdots, n, j = 1, \cdots, k), \tag{14}$$

$$\eta_{ij} = \frac{e^{-\xi_{ij}}}{\sum_{l=1}^{k} e^{-\xi_{il}}}(1 - \frac{1}{k}\sum_{l=1}^{k}\xi_{il}), \ (i = 1, \cdots, n, j = 1, \cdots, k), \tag{15}$$

$$V_j = \frac{\sum_{i=1}^{n}(\mu_{ij}(2-\xi_{ij}))^m X_i}{\sum_{i=1}^{n}(\mu_{ij}(2-\xi_{ij}))^m}, \ (j = 1, \cdots, k). \tag{16}$$

The iteration will not stop until reach the maximum iterations or $\|\mu^{(t)} - \mu^{(t-1)}\| + \|\eta^{(t)} - \eta^{(t-1)}\| + \|\xi^{(t)} - \xi^{(t-1)}\| < \epsilon$.

3. The proposed model and solutions

Definition 1. *[15] Set U be a space of points (objects), with a generic element in U denoted by u. A SVNS A in U is characterized by three membership functions, a truth membership function T_A, an indeterminacy membership function I_A and a falsity-membership function F_A, where $\forall u \in U, T_A(u), I_A(u), F_A(u) \in [0,1]$. That is $T_A : U \to [0,1], I_A : U \to [0,1]$ and $F_A : U \to [0,1]$. There is no restriction on the sum of $T_A(u), I_A(u)$ and $F_A(u)$, thus $0 \leq T_A(u) + I_A(u) + F_A(u) \leq 3$.*

Moreover, the hesitate membership function is defined as $H_A : U \rightarrow [0,3]$, and $\forall u \in U, T_A(u) + I_A(u) + F_A(u) + H_A(u) = 3$.

Entropy is a key concept in the uncertainty field. It is a measure of the uncertainty of a system or a piece of information. It is an improvement of information entropy. The Tsallis entropy [28], which is a generalization of the standard Boltzmann-Gibbs entropy, is defined as follows.

Definition 2. *[28] Let $\mathcal{X}$ be a finite set, X be a a random variable taking values $x \in \mathcal{X}$, with distribution $p(x)$. The Tsallis entropy is defines as $S_m(X) = \frac{1}{m-1}(1 - \sum_{x\in\mathcal{X}} p(x)^m)$. where $m > 0$ and $m \neq 1$.*

For FCM, μ_{ij} denotes the fuzzy membership degree of X_i to C_j, and sanctifies $\sum_{j=1}^{k} \mu_{ij} = 1$. From Definition 2, the Tsallis entropy of μ can be described $S_m(\mu) = \sum_{i=1}^{n} \frac{1}{m-1}(1 - \sum_{j=1}^{k} \mu_{ij}^m)$. Being n is fixed number, Yasuda [13] use the following formulary to describe the the Tsallis entropy of μ:

$$S_m(\mu) = -\frac{1}{m-1}(\sum_{i=1}^{n}\sum_{j=1}^{k} \mu_{ij}^m - 1). \tag{17}$$

The maximum entropy principle has been widely applied in many fields, such as spectral estimation, image restoration, error handling of measurement theory, and so on. In the following, the maximum entropy principle is applied to the single valued neutrosophic set clustering. After the objection function of clustering is built, the maximum fuzzy entropy is used to regularized variables.

Supposing that there is a data set D consisting of n data points in d dimensions. Let $\mu_{ij}, \gamma_{ij}, \eta_{ij}$ and ξ_{ij} are the truth membership degree, falsity-membership degree, indeterminacy membership degree and hesitate membership degree respectively that each data point X_i belongs to cluster C_j. Denote μ, γ, η and ξ being the matrices whose elements are $\mu_{ij}, \gamma_{ij}, \eta_{ij}$ and ξ_{ij} respectively, where $\xi_{ij} = 3 - \mu_{ij} - \gamma_{ij} - \eta_{ij}$. The single valued neutrosophic clustering based on Tsallis entropy Maximization (SVNC-TEM) is minimization of the following objective function:

$$\begin{aligned} J = \; & \textstyle\sum_{i=1}^{n}\sum_{j=1}^{k}(\mu_{ij}(4 - \xi_{ij} - \gamma_{ij}))^m \|X_i - V_j\|^2 + \frac{\rho}{m-1}(\sum_{i=1}^{n}\sum_{j=1}^{k}(u_{ij}(4 - \gamma_{ij} - \xi_{ij}))^m - 1) \\ & + \textstyle\sum_{i=1}^{n}\sum_{j=1}^{k}\eta_{ij}(\log \eta_{ij} + \xi_{ij}/3), \end{aligned} \tag{18}$$

The constraints are given as follows:

$$\mu_{ij}, \gamma_{ij}, \eta_{ij} \in [0,1], \xi_{ij} \in [0,3], (i = 1,2,\cdots,n, j = 1,2,\cdots,k) \tag{19}$$

$$\sum_{l=1}^{k}(u_{il}(4 - \gamma_{il} - \xi_{il})) = 1, (i = 1,2,\cdots,n), \tag{20}$$

$$\sum_{l=1}^{k}(\eta_{il} + \xi_{il}/(3*k)) = 1, (i = 1,2,\cdots,n) \tag{21}$$

The proposed model in Formulary (18-21) is applied the maximum entropy principle on the SVNS. Now, let us summarize the major points of this model as follows.

- The first term of objection function (18) describes the weighted distance sum of each data point X_i to the cluster center V_j. Being μ_{ij} from the positive aspect and $(4 - \xi_{ij} - \gamma_{ij})$ (The 4 is selected in order to guarantee $\mu_{ij} \in [0,1]$ in the iterative calculation) from the negative aspect denote the membership degree for X_i to V_j, we use $\mu_{ij}(4 - \xi_{ij} - \gamma_{ij})$ represents the "integrated true" membership of the i-th data point in the j-th cluster. From the maximum entropy principle, the best represents the current state of knowledge is the one with largest entropy, so the second term of objection function (18) describes the negative Tsallis entropy of $\mu(4 - \gamma - \xi)$, which means that minimization of (18) is maximum Tsallis entropy. ρ is regularization parameter. If $\gamma = \eta = \xi = 0$, the proposed model returns to the FCM model.

- Formulary (19) guarantees the definition of the SVNS (Definition 1).
- Formulary (20) implies that the "integrated true" membership of a data point X_i to the cluster center V_j satisfies the sum-row constraint of memberships. For convenience, we set $T_{ij} = \mu_{ij}(4 - \xi_{ij} - \gamma_{ij})$ and X_i belongs to class C_l if $T_{il} = \max(T_{i1}, T_{i2}, \cdots, T_{ik})$.
- Equation (21) guarantees the working on the SVNS since at least one of two uncertain factors, namely indeterminacy membership degree and hesitate membership degree, always exist in the model.

Theorem 1. *The optimal solutions of the systems (18-21) are:*

$$V_j = \frac{\sum_{i=1}^{n}(\mu_{ij}(4-\gamma_{ij}-\xi_{ij}))^m X_i}{\sum_{i=1}^{n}(\mu_{ij}(4-\gamma_{ij}-\xi_{ij}))^m}, \tag{22}$$

$$\mu_{ij} = \frac{1}{\sum_{l=1}^{k}(4-\gamma_{ij}-\xi_{ij})\left(\frac{\|X_i-V_j\|^2+\frac{\rho}{m-1}}{\|X_i-V_l\|^2+\frac{\rho}{m-1}}\right)^{\frac{1}{m-1}}}, \tag{23}$$

$$\gamma_{ij} = 4-\xi_{ij}-\frac{1}{u_{ij}\sum_{l=1}^{k}\left(\frac{\|X_i-V_j\|^2+\frac{\rho}{m-1}}{\|X_i-V_l\|^2+\frac{\rho}{m-1}}\right)^{\frac{1}{m-1}}}, \tag{24}$$

$$\eta_{ij} = (1-\frac{1}{3k}\sum_{l=1}^{k}\xi_{il})\frac{e^{-\xi_{ij}}}{\sum_{l=1}^{k}e^{-\xi_{il}}}, \tag{25}$$

$$\xi_{ij} = 3-\mu_{ij}-\gamma_{ij}-\eta_{ij}. \tag{26}$$

Proof. The Lagrangain multiplier of optimization model (18-21) is:

$$\begin{aligned} J = \; & \sum_{i=1}^{n}\sum_{j=1}^{k}(u_{ij}(4-\gamma_{ij}-\xi_{ij}))^m\|X_i-V_j\|^2+\frac{\rho}{m-1}(\sum_{i=1}^{n}\sum_{j=1}^{k}(u_{ij}(4-\gamma_{ij}-\xi_{ij}))^m-1) \\ & +\sum_{i=1}^{n}\sum_{j=1}^{k}\eta_{ij}(\log\eta_{ij}+\xi_{ij}/3)+\sum_{i=1}^{n}\lambda_i(\sum_{j=1}^{C}\mu_{ij}(4-\gamma_{ij}-\xi_{ij})^m)-1) \\ & +\sum_{i=1}^{n}\chi_i(\sum_{j=1}^{k}(\eta_{ij}+\xi_{ij}/(3k))-1). \end{aligned} \tag{27}$$

Where λ_i and χ_i are Lagrangain multipliers.

In order to get V_j, taking the derivative of objective function with respect to V_j, we have $\frac{\partial J}{\partial V_j} = \sum_{i=1}^{n}(\mu_{ij}(4-\gamma_{ij}-\xi_{ij}))^m(-2X_i+2V_j)$. Being $\frac{\partial J}{\partial V_j} = 0$, so $\sum_{i=1}^{n}(\mu_{ij}(4-\eta_{ij}-\xi_{ij}))^m(-2X_i+2V_j) = 0$ $\Leftrightarrow \sum_{i=1}^{n}(\mu_{ij}(4-\eta_{ij}-\xi_{ij}))^m X_i = \sum_{i=1}^{n}(\mu_{ij}(4-\eta_{ij}-\xi_{ij}))^m V_j \Leftrightarrow V_j = \frac{\sum_{i=1}^{n}(\mu_{ij}(4-\eta_{ij}-\xi_{ij}))^m X_i}{\sum_{i=1}^{N}(\mu_{ij}(4-\eta_{ij}-\xi_{ij}))^m}$

Similarly, $\frac{\partial J}{\partial \mu_{ij}} = m\mu_{ij}^{m-1}(4-\xi_{ij}-\eta_{ij})^m\|X_i-V_j\|^2+\frac{\rho m}{m-1}\mu_{ij}^{m-1}(4-\xi_{ij}-\eta_{ij})^m)+\lambda_i(4-\xi_{ij}-\eta_{ij}) = 0 \Leftrightarrow \mu_{ij}^{m-1}(4-\gamma_{ij}-\xi_{ij})^{m-1}(m\|X_i-V_j\|^2+\frac{\rho m}{m-1})+\lambda_i = 0 \Leftrightarrow \mu_{ij} = \frac{1}{4-\gamma_{ij}-\xi_{ij}}(\frac{\lambda_i}{m\|X_k-V_j\|^2+\frac{\rho m}{m-1}})^{\frac{1}{m-1}}$. From (20), we can get $\sum_{l=1}^{k}(\frac{\lambda_i}{m\|X_i-V_l\|^2+\frac{m\rho}{m-1}})^{\frac{1}{m-1}} = 1$, that is $\lambda_i = (\frac{1}{\sum_{l=1}^{k}\frac{1}{(m\|X_i-V_l\|^2+\frac{m\rho}{m-1})^{\frac{1}{m-1}}}})^{m-1}$, so $\mu_{ij} = \frac{1}{\sum_{l=1}^{k}(4-\xi_{ij}-\eta_{ij})(\frac{\|X_i-V_j\|^2+\frac{\rho}{m-1}}{\|X_i-V_l\|^2+\frac{\rho}{m-1}})^{\frac{1}{m-1}}}$, thus (23) holds.

From (23), we can also get $\mu_{ij}(4-\gamma_{ij}-\xi_{ij}) = \frac{1}{\sum_{l=1}^{k}(\frac{\|X_i-V_j\|^2+\frac{\rho}{m-1}}{\|X_i-V_l\|^2+\frac{\rho}{m-1}})^{\frac{1}{m-1}}}$. so $\gamma_{ij} = 4-\xi_{ij}-\frac{1}{u_{ij}\sum_{i=1}^{C}(\frac{\|X_i-V_j\|^2+\frac{\rho}{m-1}}{\|X_i-V_k\|^2+\frac{\rho}{m-1}})^{\frac{1}{m-1}}}$, thus (24) holds.

Similarly, $\frac{\partial L}{\partial \eta_{ij}} = \log \eta_{ij} + 1 - \chi_i + \xi_{ij} = 0 \Leftrightarrow \eta_{ij} = e^{(}\chi_i - 1 - \xi_{ij}),$ From (21), we have, $\sum_{l=1}^{k} e^{\chi_i - 1 - \xi_{il}} + \frac{1}{3k}\sum_{l=1}^{k} \xi_{il} = 1 \Leftrightarrow e^{\chi_i - 1}\sum_{l=1}^{k} e^{-\xi_{il}} = 1 - \frac{1}{3k}\sum_{l=1}^{k} \xi_{il} \Leftrightarrow e^{\chi_i - 1} = \frac{1 - \frac{1}{3k}\sum_{l=1}^{k} \xi_{il}}{\sum_{l=1}^{k} e^{-\xi_{il}}}$. So we have, $\eta_{ij} = (1 - \frac{1}{3k}\sum_{l=1}^{k} \xi_{il})\frac{e^{-\xi_{ij}}}{\sum_{l=1}^{k} e^{-\xi_{il}}}$.

Finally, from Definition 1, we can get $\xi_{ij} = 3 - \mu_{ij} - \gamma_{ij} - \eta_{ij}$. Thus (26) holds.

□

The Theorem 1 guarantee the convergence of the proposed method and the detail descriptions of SVNC-TEM algorithm is presented in the following:

Algorithm: SVNC-TEM
Input: Data set $D = \{X_1, X_2, \cdots, X_n\}$ (n elements, d dimensions), number of clusters k,
Maximal number of iteration Max-Iter, Parameters: m, ϵ, ρ
Output: Cluster result
$t = 0$;
Initialize μ, γ, ξ, satisfy constraints (19,20);
Repeat
$t = t + 1$;
Update $V_j^{(t)}, (j = 1, 2, \cdots, k)$ using Eq. (22);
Update $\mu_{ij}^{(t)}, (i = 1, 2, \cdots, n, j = 1, 2, \cdots, k)$ using Eq. (23);
Update $\gamma_{ij}^{(t)}, (i = 1, 2, \cdots, n, j = 1, 2, \cdots, k)$ using Eq. (24);
Update $\eta_{ij}^{(t)}, (i = 1, 2, \cdots, n, j = 1, 2, \cdots, k)$ using Eq. (25);
Update $\xi_{ij}^{(t)}, (i = 1, 2, \cdots, n, j = 1, 2, \cdots, k)$ using Eq. (26);
Update $T_{ij}^{(t)} = \mu_{ij}^{(t)}(4 - \gamma_{ij}^{(t)} - \xi_{ij}^{(t)}), (i = 1, 2, \cdots, n, j = 1, 2, \cdots, k)$;
Update $J^{(t)}$ using Eq. (18);
Until $\lvert J^{(t)} - J^{(t-1)} \rvert < \epsilon$ or Max-Iter has reached.
Assign $X_i (i = 1, 2, \cdots, n)$ into the l-th class if $T_{il} = \max(T_{i1}, T_{i2}, \cdots, T_{ik})$.

Compared with FCM, the proposed algorithm needs additional time to calculate μ, γ, η and ξ in order to more precisely precise describe the object and get better performance. If the dimension of the given data set is d, the number of objects is n, the number of clusters is c and the number of iterations is t, then the computational complexity of the proposed algorithm is $O(dnct)$. We can see that the computational complexity is very high if d and n are large.

4. Experimental results

In the section, some experiments have intended to validate the effectiveness of proposed algorithm SVNC-TEM for data clustering. Firstly, we use an artificial data set to show SVNC-TEM can cluster well. Secondly, the proposed clustering method is used in image segmentation by an example. Lastly, we select five benchmark data sets and SVNC-TEM is compared with four state-of-the-art clustering algorithms, which are: k-means, FCM, IFC and FS-PFS.

In the experiments, the parameter m is selected as 2 and $\epsilon = 10^{-5}$. Maximum iterations Max-Iter= 100. The selected data sets have class labels, so the number of cluster k is known in advance. All the codes in the experiments are implemented in MATLAB R2015b.

4.1. An artificial data to cluster by SVNC-TEM algorithm

The activities of the SVNC-TEM algorithm will be illustrated to cluster on an artificial data, which is 2-dimensional data and has 100 data points, four classes. We use the example to show the clustering process of the proposed algorithm. The distribution of data points is illustrated in Figure1(a). Figures 1(b-e) show the clusters results when the number of iterations is t =1, 5, 10, 20 respectively. We can see that the clustering result is obtained when $t = 20$. Figure 1(f) show the final results of the clustering, the number of iterations is 32. We can see that the proposed algorithm gives right clustering results from Figure 1.

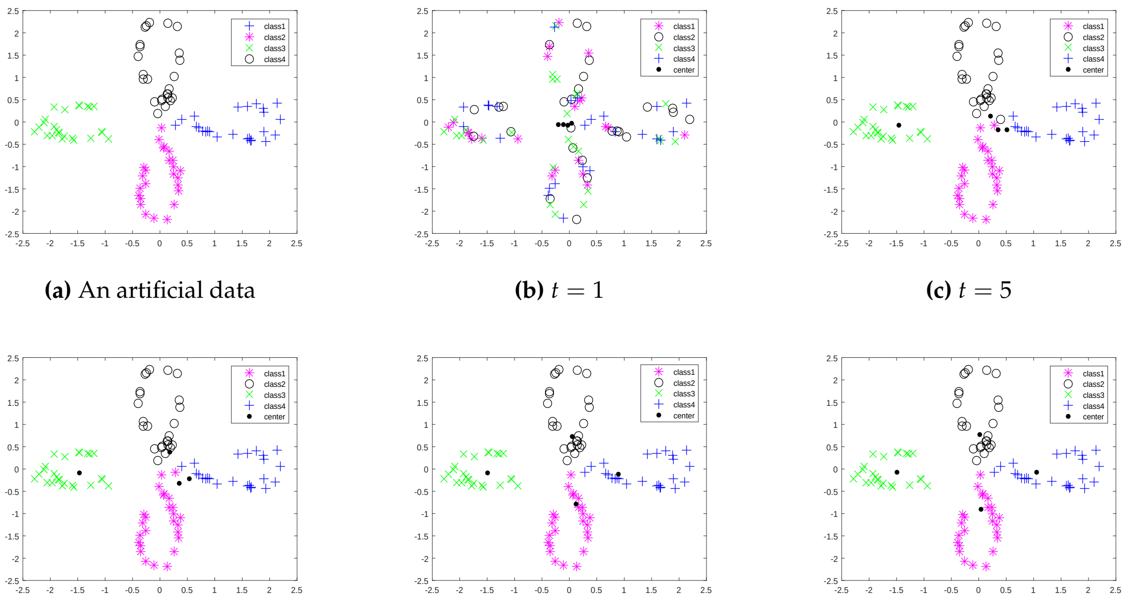
4.2. *Image segmentation by SVNC-TEM algorithm*

In this subsection, we use the proposed algorithm to image segmentation. As a simple example, the Lena image is used to test the proposed algorithm for image segmentation. Through this example, we wish to show that the proposed algorithm can be applied to image segmentation. Figure 2(a) is the original Lena image. Figure 2(b) shows the segmentation images when the number of clustering is $k = 2$, and we can see that the quality of the image has been greatly reduced. Figure 2(c-f) show the segmentation images when the number of clustering is $k = 5, 8$, 11 and 20 respectively. We can see that the quality of segmentation image has been improved very well with the increase of clustering number.

The above two examples demonstrate that the proposed algorithm can be effectively applied to the clustering and image processing. Next, we will further compare the given algorithm with other state-of-art clustering algorithms on benchmark data sets.

4.3. *Compare analysis experiments*

In order to verify the clustering performance, in the subsection, we experiment with five benchmark data sets of UCI Machine Learning Repository, which are IRIS, CMC, GLASS, BALANCE and BREAST. These data sets are used to test the performance of the clustering algorithm. Table 1 shows the details characteristic of the data sets.



(a) An artificial data **(b)** $t = 1$ **(c)** $t = 5$

(d) $t = 10$ **(e)** $t = 20$ **(f)** $t = 32$

Figure 1. The demonstration figure of clustering process for an artificial data. (a) the original data (b-e) the clustering figures when the number of iterations t =1, 5, 10, 20 respectively. (f) The final clustering result.

(a) Lena image **(b)** $k = 2$ **(c)** $k = 5$

(d) $k = 8$ **(e)** $k = 11$ **(f)** $k = 20$

Figure 2. The image segmentation for Lena image. (a) the original Lena image (b-f) the clustering images when the number of clustering k =2, 5, 8, 11 and 20 respectively.

Table 1. description of experimental data sets

Dataset	No. of elements	No. of attributes	No. of classes	Elements in each classes
IRIS	150	4	3	[50, 50, 50]
CMC	1473	9	3	[629, 333, 511]
GLASS	214	9	6	[29, 76, 70, 17, 13, 9]
BALANCE	625	4	3	[49, 288, 288]
BREAST	277	9	2	[81, 196]

In order to compare the performance of the clustering algorithms, three evaluation criteria are introduced as following.

Given one data point X_i, denote p_i be the truth class and q_i be the predicted clustering class. The accuracy(ACC) measure is evaluated as follows:

$$ACC = \frac{\sum_{i=1}^{n} \delta(p_i, map(q_i))}{n}, \tag{28}$$

where n is the total number of data points, $\delta(x, y) = 1$ if $x = y$; otherwise $\delta(x, y) = 0$. $map(\bullet)$ is the best permutation mapping function that matches the obtained clustering label to the equivalent label of the data set. One of the best mapping functions is the Kuhn-Munkres algorithm [29]. The higher the ACC is, the better the clustering performance is.

Given two random variables X and Y, $MI(X;Y)$ is the mutual information of X and Y. $H(X)$ and $H(Y)$ are the entropies of P and Q, respectively. We use the normalized mutual information (NMI) as follows

$$NMI(X;Y) = \frac{MI(X;Y)}{\sqrt{H(X)H(Y)}}. \tag{29}$$

The clustering results $\hat{C} = \{\hat{C}_j\}_{j=1}^{\hat{k}}$ and the ground truth classes $C = \{C_j\}_{j=1}^{k}$ are regarded as two discrete random variables. So, NMI is specified as the following:

$$NMI(C;\hat{C}) = \frac{\sum_{i=1}^{\hat{k}}\sum_{j=1}^{k}|\hat{C}_i \cap C_j|\log\frac{n|\hat{C}_i \cap C_j|}{|\hat{C}_i||C_j|}}{\sqrt{(\sum_{i=1}^{\hat{k}}|\hat{C}_i\log\frac{|\hat{C}_i|}{n}|)(\sum_{j=1}^{k}|C_j|\log\frac{|C_j|}{n})}}. \quad (30)$$

The higher the NMI is, the better the clustering performance is.

The Rand index is defined as,

$$RI = \frac{2(a+d)}{n(n-1)}. \quad (31)$$

where a is the number of pairs of data points belonging to the same class in C and to the same cluster in $\hat{C}$. d is the number of pairs of data points belonging to the different class and to the different cluster. n is the number of data points. The larger the Rand index is, the better the clustering performance is.

We do a series of experiments to indicate the performance of the proposed method for data clustering. In the experiments, we set parameters of all approaches in same way to make the experimenters fair enough, that is, for parameter ρ, we set $\rho = \{0.01, 0.05, 0.07, 0.1, 0.15, 0.5, 1, 2, 5, 8, 9, 15, 20, 50\}$. For α, we set $\alpha = \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}$. For each parameter, we run the given method 50 times and select the best mean value to report. Tables 2-4 show the results with different evaluation measures respectively. In these tables, we use bold font to indicate the best performance.

We analyze the results from data set firstly. For IRIS data set, the proposed method gets the best performance for ACC, NMI and RI. For CMC data set, the proposed method has the best performance for ACC and RI. For GLASS and BREAST data sets, the proposed method gets the best performance for ACC and NMI. For BALANCE data set, the proposed method has the best performance for NMI and RI. On the other hand, from the three evaluation criteria, for ACC and NMI, the proposed method wins the other methods in four data sets. For RI, SVNC-TEM wins the other methods in three data sets. From the experimental results, we can see that the proposed method has better clustering performance than other algorithms.

Table 2. The ACC for different algorithms on different data sets

Data Set	k-means	FCM	IFC	FC-PFC	SVNC-TEM
IRIS	0.8803	0.8933	**0.9000**	0.8933	**0.9000**
CMC	0.3965	0.3917	0.3958	0.3917	**0.3985**
GLASS	0.3219	0.2570	0.3636	0.2935	**0.3681**
BALANCE	**0.5300**	0.5260	0.5413	0.5206	0.5149
BREAST	0.6676	0.5765	0.6595	0.6585	**0.6686**

Table 3. The NMI for different algorithms on different data sets

Data Set	k-means	FCM	IFC	FC-PFC	SVNC-TEM
IRIS	0.7514	0.7496	0.7102	0.7501	**0.7578**
CMC	0.0320	0.0330	0.0322	**0.0334**	0.0266
GLASS	0.0488	0.0387	0.0673	0.0419	**0.0682**
BALANCE	0.1356	0.1336	0.1232	0.1213	**0.1437**
BREAST	0.0623	0.0309	0.0285	0.0610	**0.0797**

Table 4. The RI for different algorithms on different data sets

Data Set	k-means	FCM	IFC	FC-PFC	SVNC-TEM
IRIS	0.8733	0.8797	0.8827	0.8797	**0.8859**
CMC	0.5576	0.5582	0.5589	0.5582	**0.5605**
GLASS	0.5373	**0.6294**	0.4617	0.5874	0.4590
BALANCE	0.5940	0.5928	0.5899	0.5904	**0.5999**
BREAST	0.5708	0.5159	**0.5732**	0.5656	0.5567

5. Conclusions

In the paper, we consider the truth membership degree, the falsity-membership degree, the indeterminacy membership degree and hesitate membership degree in a comprehensive way to data clustering by single valued neutrosophic set. We propose a novel data clustering algorithm SVNC-TEM and the experimental results show that the proposed algorithm can be considered as a promising tool for data clustering and image processing. The proposed algorithm has better clustering performance than other algorithms such as k-means, FCM, IFC and FC-PFS. Next, we will consider the proposed method to deal with outliers. Moreover, we will consider the clustering algorithm combines with spectral clustering and other clustering methods.

Acknowledgments: This work is partially supported by National Natural Science Foundation of China (Grant No. 11501435), Instructional Science and Technology Plan Projects of China National Textile and Apparel Council (2016073).

Author Contributions: All authors have contributed equally to this paper.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. MacQueen J. B. Some methods for classification and analysis of multivariate observations. *In Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability*. University of California Press, Berkeley, 1967, 281-297.
2. Johnson S. C. Hierarchical clustering schemes. *Psychometrika*. **1967**, *2*, 241-254.
3. Ng A. Y.; Jordan M. I.; Weiss Y. On spectral clustering: Analysis and an algorithm. Advances in Neural *Information Processing Systems (NIPS)*. **2002**, *14*, 849-856.
4. Andenberg M.R.*Cluster Analysis for Applications* Academic Press: New York, 1973.
5. Ménard M.; Demko C.; Loonis P. The fuzzy $c+2$ means: solving the ambiguity rejection in clustering. **2000**, *33*, 1219-1237.
6. Aggarwal C. C.; Reddy C. K. *Data Clustering: Algorithms and Applications*. CRC Press, Chapman & Hall, 2013.
7. Bezdek J.C.; Ehrlich R.; Full W. FCM: the fuzzy c-means clustering algorithm. *Computers & Geosciences* . **1984**, *10*, 191-203.
8. Chaira T. A novel intuitionistic fuzzy C means clustering algorithm and its application to medical images. *Applied Soft Computing* . **2011** , *11* , 1711-1717.
9. Oliveira J. V. D.; Pedrycz W. *Advances in Fuzzy Clustering and Its Applications*. Wiley, Chichester, 2007.
10. Thong P. H.; Son L. H. Picture fuzzy clustering: a new computational intelligence method. *Soft computing* . **2016** , *20* , 3549-3562.
11. Graves D.; Pedrycz W. Kernel-based fuzzy clustering and fuzzy clustering: a comparative experimental study. *Fuzzy Sets and Systems*. **2010** , *161*, 522-543.
12. Ramathilagam S.; Devi R.; Kannan S.R. Extended fuzzy c-means: an analyzing data clustering problems. *Cluster Computing*. **2013**, *16* , 389-406.
13. Yasuda M. Deterministic Annealing Approach to Fuzzy C-Means Clustering Based on EntropyMaximization. *Advances in Fuzzy Systems*. **2011**, doi:10.1155/2011/960635
14. Smarandache F. *A Unifying Field in Logics. Neutrosophy: Neutrosophic probability, set and logic*. American Research Press, Rehoboth, 1998.

15. Wang H.; Smarandache F.; Zhang Y.Q. Sunderraman R. Single valued neutrosophic sets. *Multispaceand Multistruct*. **2010** , *4* , 410-413.
16. Majumdar P.; Samant S.K. On similarity and entropy of neutrosophic sets. *Journal of Intelligent & Fuzzy Systems* . **2014** , *26*, 1245-1252.
17. Ye J. Improved correlation coefficients of single valued neutrosophic sets and interval neutrosophic sets for multiple attribute decision making. *Journal of Intelligent & Fuzzy Systems*. **2014**, *27*, 2453-2462.
18. Zhang X.; Bo C.; Smarandache F.; Dai J. New Inclusion Relation of Neutrosophic Sets with Applications and Related Lattice Structure. *Journal of Machine Learning & Cybernetics* . **2018** .
19. Zhang, X.H.; Smarandache, F.; Liang, X.L. Neutrosophic duplet semi-group and cancellable neutrosophic triplet groups. *Symmetry*. **2017** , *9*, 275, doi:10.3390/sym9110275.
20. Ye j. Single-Valued Neutrosophic Minimum Spanning Tree and Its Clustering Method.*Journal of Intelligent Systems*. **2014**, *23*, 311-324.
21. Ye J. Clustering Methods Using Distance-Based Similarity Measures of Single-Valued Neutrosophic Sets.*Journal of Intelligent Systems*. **2014**, *23* , 379-389.
22. Guo Y.; Sengur A. NCM: Neutrosophic c-means clustering algorithm. *Pattern Recognition*. **2015** , *48*, 2710-2724.
23. Thong P. H.; Son L. H. A novel automatic picture fuzzy clustering method based on particle swarm optimization and picture composite cardinality. *Knowledge-Based Systems*. **2016** , *109*, 48-60.
24. Thong P. H.; Son L. H. Picturefuzzyclusteringforcomplexdata. *Engineering ApplicationsofArtificial Intelligence*. **2016** , *56*, 121-130.
25. Son L. H. DPFCM: A novel distributed picture fuzzy clustering method on picture fuzzy sets. *Expert Systems with Applications*. **2015** , *42*, 51-66.
26. B.C.Cuong, Picture Fuzzy Sets. *Journal of Computer Science and Cybernetics*. **2014** , *30*, 409-420.
27. Cuong B.C.; Phong P.H. Smarandache F. Standard Neutrosophic Soft Theory: Some First Results, *Neutrosophic Sets and Systems*. **2016** , *12*, 80-91.
28. Tsallis C. Possible generalization of Boltzmann-Gibbs statistics. *Journal of Statistical Physics*. **1988** , *52*, 479-487.
29. Lovasz L.; Plummer M.D. *Matching Theory*. American Mathematical Society, 2009.