Classification of Power Signals using ACO based K-Means Algorithm and Fuzzy C-Means Algorithm

Varsha Bal
M-Tech Scholar (EEE)
CUTM, Jatni
Bhubaneshwar, India

Satyasis Mishra
Professor (ECE)
CUTM, Jatni
Bhubaneshwar, India

ABSTRACT

This paper presents pattern classification of power signal disturbances using modified form of S-transform, which is obtained by taking the Inverse Fourier transform of S-Transform is known as time-time transform (TT-transform). The TT-Transform based used for power signals to extract features, visual localization, detection. TT-Transform has good ability in gathering frequency; it gathers the high frequency signals in diagonal position of the spectrum and suppressing the low frequency signals. Only the diagonal of TT-Transform has been used for signal characterization. The diagonal of TT-Transform represent a simple frequency filtered version of the original signal. The extracted features are fed as input to a fuzzy C-means clustering algorithm (FCA) to generate a decision tree. To improve the pattern classification of the fuzzy C-means decision tree, the cluster centers are updated using ant colony optimized technique (ACO). Further K-Means algorithm is used for updation of cluster centers using ant colony optimization technique (ACO) for classification accuracy and the results of both the algorithm are compared.

KEYWORDS
Nonstationary Power Signals, Fourier Transform, Short Time Fourier Transform (STFT),Wavelet Transform (WT),S-Transform (ST), TT-Transform,Ant Colony Optimization (ACO), K-means Algorithm (KMA), Fuzzy C-means Algorithm (FCM)

1. INTRODUTION
In electrical power systems, the power signals exhibit fluctuations in amplitude, phase, and frequency due to the operation of electronic devices. The sudden variation voltage signals such as spike and notch, transient are often seen in power system networks due to electronic capacitors switching operations across nodes in a power network. The frequency content of the power signal cannot be analyzed through conventional Fourier transform. Over a past decade several methods of advanced signal processing techniques like S-Transform [6], wavelet transform have been applied to detect, localize and classify the power disturbance signal patterns like voltage sag, voltage swell, transients, spikes, and voltage notches etc.
In the current research Wavelet transform (WT) is widely used for power signal disturbance. Even though WT has been shown good results in detection of power signal disturbances, but the effect of electrical noise is not filtered properly in some of the cases. The advantage of S-Transform is that it preserves the phase information of the signal, and also provides a variable resolution similar to wavelet transform. However S-Transform suffers from poor concentration of energy at higher frequency and hence poor frequency localization.TT-Transform, a new view of localizing the time features of a time series around a particular point on the time axis. The TT-Transform [1] is derived from the S-Transform that is inverse Fourier transform of S-Transform.

Features are extracted from the power signal disturbances using modified TT-Transform [1], a clustering analysis fuzzy C means algorithm is used to group the feature data into clusters and identifying the class of the signal. The Fuzzy C-means algorithm suffers to have good choice of the cluster centers when the noise present in the original power signal. So, there is possibility of local minima while clustering is done. In this paper Fuzzy C-means and ACO algorithm are taken together for clustering the data. ACO [3] is inspired by the foraging behavior of ants. Ants deposit pheromone along the traveled path which is used by other ants to follow the trail. ACO is a Heuristic method for solving a very general class of computational problems by combining user given heuristics in the hope of obtaining a more efficient procedure. Shortest path is discovered via pheromone trails because more number of ants travelled that path. More pheromone on path increases probability of that path being followed.
In this paper section II presents the modified S-Transform and TT-Transform section III presents the decision tree and ACO algorithm with fuzzy c means and k-means algorithm, section IV presents results followed by conclusion.

2. TIME FREQUENCY ANALYSIS OF POWER SIGNALS USING TT-TRANSFORM
The advantage of S-Transform [8] is that it preserves the phase information of the signal; however, S-Transform suffers from poor concentration of energy at higher frequency and hence poor frequency localization.
The standard S-Transform of a signal x (t) is given by convolution integral (Stockwell et al., 1996) as;

$$S(t) = \int_{-\infty}^{\infty} x(\tau)\frac{1}{\sigma\sqrt{2\pi}}.e^{\frac{-(t-\tau)^2}{2\sigma^2}}.e^{-2\pi f\tau}d\tau \qquad (1)$$

where *f* is frequency, *t* and *τ* are time variables.

The standard deviation *σ* in Eqn. (1) is a function of frequency *f* and in normal S-transform is defined as;

$$\sigma = \frac{1}{|f|} \qquad (2)$$

Since the window depends on the frequency, it provides a good localization in the time domain for higher frequencies. Further the integral of S-transform over time is the Fourier transform.

$$X(f) = \int_{-\infty}^{\infty} S(t, f)dt \qquad (3)$$

A significant improvement of S-Transform can be realized by defining the standard deviation of the window as;

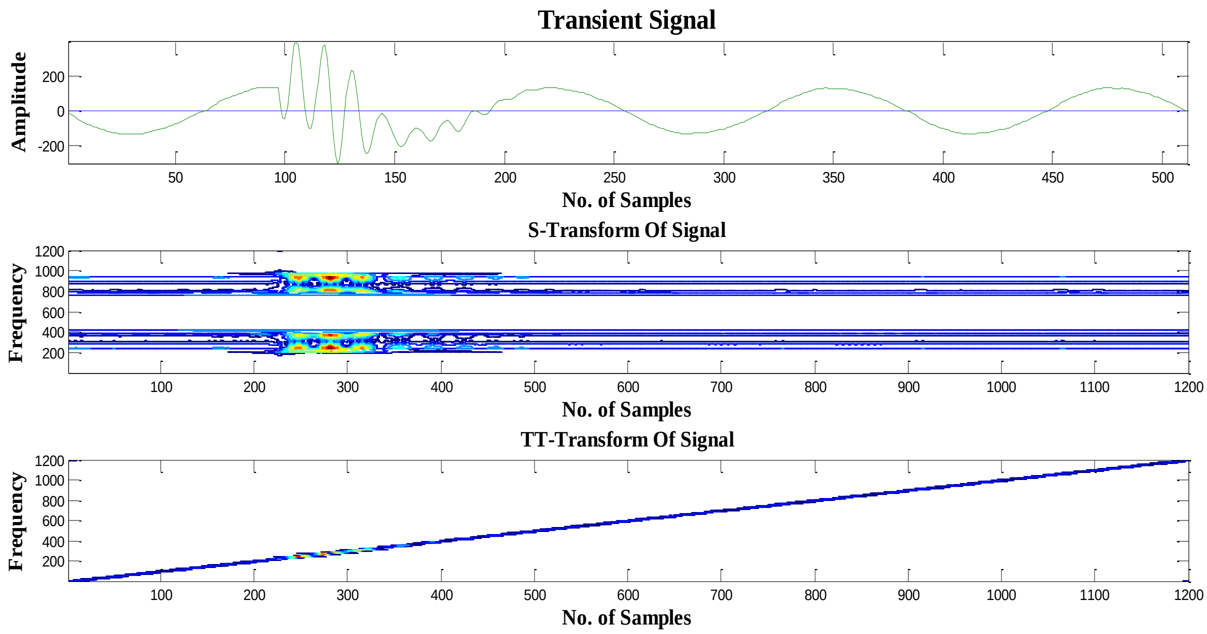

Fig .1 Localization of Transient signal using S-Transform and TT-Transform

$$\sigma = \frac{\hat{k}}{|f|^{\gamma}} \qquad (4)$$

resulting in a modified S-Transform as

$$S(t, f) = \frac{|f|^{\gamma}}{k\sqrt{2\pi}} \int_{-\infty}^{\infty} x(\tau) e^{\frac{-(t-\tau)^2 f^2}{2\sigma^2}} .e^{-2\pi f\tau} d\tau \qquad (5)$$

where, γ and α control the width of the window.

Improved time-frequency localization is possible by suitably choosing k and γ. Normally k is chosen between 0.2 and 1 and $0.1 \leq \gamma \geq 1$. Defining a function;

$$p(t, f) = x(t) e^{\frac{(t-\tau)^2 f^{2\gamma}}{2\alpha^2}} dt \qquad (6)$$

and further defining the Fourier Transform (σ to α) of S (t, f) as;

$$B(f, \alpha) = X(\alpha + f) e^{\frac{2\pi^2 \alpha^2 k^2}{2\alpha^2}} \qquad (7)$$

The generalized S-transform can now be written as;

$$S(\tau, f) = \int_{-\infty}^{\infty} X(\alpha + f).e^{-\frac{2\pi^2 \alpha^2 k^2}{|f|^{2\gamma}}} .d\alpha \qquad (8)$$

The TT- transform [1] is obtained from the inverse Fourier transform of the S-transform as

$$TT(t, \tau) = \int_{-\infty}^{\infty} S(t, f) e^{2\pi i f\tau} df \qquad (9)$$

Also inverting the TT-transform, the original signal x (τ) is obtained as;

$$x(\tau) = \int_{-\infty}^{\infty} TT(t, \tau) dt \qquad (10)$$

3. UPDATION OF CENTER OF FUZZY C-MEANS ALGORITHM AND K-MEANS ALGORITHM BY ACO ALGORITHM

In this paper different features such as energy, standard deviation, autocorrelation, mean, and variance have been extracted from the power signals.

ACO, proposed by Dorigo et al., is an optimization technique which helps in refining the data clusters by optimizing the Euclidean distance of each data point from the centers. In ACO [7], at each iteration the pheromone values are updated by all the "m" ants that have built a solution in the iteration itself. The pheromone τ_{ij}, associated with the path joining center i and data points j, is updated as follows:

The pheromone τ_{ij}, associated with the path joining center i and data points j, is updated as follows:

$$\tau_{ij}(new) = (1-\rho)\tau_{ij}(old) + \sum_{k-1}^{m} \Delta\tau_{ij}^{k} \tag{11}$$

Where ρ, is the evaporation rate (ensures that pheromone is not accumulated infinitely, and denotes the proportion of old pheromone) "m" is the number of data points. (ρ=0.6) and is the quantity of pheromone deposited on path (i, j) that is from i to j by ant "k" at the current iteration.

$$\Delta\tau_{ij}^{k} = \begin{cases} Q/L_k, & if\ ant\ k\ used\ path(i,j) \text{inits tour} \\ 0, & \text{otherwise} \end{cases}$$

Where, Q is a constant and L_k is the length of the path constructed by ant k. (Q=1).

The main characteristic of ACO [3] is that, at each iteration the pheromone values are updated by all the *m* ants that have built a solution in the iteration itself. In the construction of a solution, ant selects the following path to be visited through a stochastic mechanism. When ant 'k' is in data point 'j', the probability of moving to center 'i' at time't' is given by the formula;

$$P_{ij}^{k}(t) = \frac{(\tau_{ij}(t))^{\alpha}.(\eta_{ij})^{\beta}}{\sum_{ie}(\tau_{ij}(t))^{\alpha}.(\eta_{ij})^{\beta}} \tag{12}$$

Where

$\tau_{ij}(t)$=Amount of pheromone concentration on path (i, j).

η_{ij} = (eta matrix)Heuristic desirability of choosing center i when the ant is at data point j.

j_i^k = It is the set of data point's unvisited center.

α = is a parameter to control the influence of τ_{ij}

β = is a parameter to control the influence of η_{ij} and $\alpha = 0.5, \beta = 0.8$

The parameters α and β control the relative importance of the pheromone versus the heuristic information η_{ij},(eta matrix) which is given by;

$$\eta_{ij} = 1/d_{ij} \tag{13}$$

Where d_{ij} express weighted Euclidean distance between centers i and the data point j.

$$d_{ij} = \left\| x_{jm} - c_{in} \right\|^2 \tag{14}$$

Where n = No. of centers, m = No. of data points.

The feature data of power signal disturbances are considered as group of ants. The random center signifies the food. Ants move in search of food and they have the tendency to follow the shortest path to reach the food and deposit pheromone on their trail. The center corresponding to the shortest path is selected and updated which signifies movement of the center towards the data point. The above procedure was repeated for all data points and thereby finding the updated centers. This results in the formation of two clusters corresponding to the respective updated centers. This process is implemented in the following algorithm for power signal classification.

3.1 Procedure for Clustering

Let {X = (x_{j1}, x_{j2}, ….., x_{jm}), j = 1, 2, ….., N} is the data set that needs to be clustered. The centers 'C' have been randomly initialized as C = (C_{i1}, C_{i2} …C_{in}). Initially take the data points and the random centers as the input features. Create the cost matrix by taking the Euclidean distance of each data points corresponding to each center using

$$d_{ij} = \left\| x_{jm} - c_{in} \right\|^2 \tag{15}$$

Calculate the eta matrix which is the reciprocal of the cost matrix using

$$\eta_{ij} = 1/d_{ij} \tag{16}$$

Calculate the probability of each point with respect to each center by using the formula in,

$$P_{ij}^{k}(t) = \frac{(\tau_{ij}(t))^{\alpha}.(\eta_{ij})^{\beta}}{\sum_{leJ_i^k}(\tau_{il}(t))^{\alpha}.(\eta_{il})^{\beta}} \tag{17}$$

In Fuzzy C means clustering we determine the cluster center C_i and the membership matrix U and we thus determine distinct clusters. Fuzzy C Means method is based on minimization of the following objective function:

$$J_m = \sum_{i=1}^{N}\sum_{j=1}^{C} u_{ij}^{m} \left\| x_i - c_j \right\|^2 \tag{18}$$

Where m = 2, fuzziness coefficient, u_{ij} is the degree of membership of x_i in cluster j, is the i_{th} of n-dimensional measured data, c_j is the n-dimensional center of the cluster. For each data point the center having the maximum probability is chosen as max $P_{ij}^{k}(t)$ and the corresponding center was updated using the formula.

$$c_j = \frac{\sum_{i-1}^{N} u_{ij}^{m}.x_i}{\sum_{i-1}^{N} u_{ij}^{m}}, \quad u_{ij} = \sum_{k-1}^{c}\left[\frac{\left\|x_i\right\| - \left\|c_j\right\|}{\left\|x_i - c_k\right\|}\right]^{-2/m-1} \tag{19}$$

3.2 K-means Algorithm

K-means algorithm is one of the simplest unsupervised learning algorithms that solve the well known clustering problem. The procedure follows a simple and easy way to classify a given data set through a certain number of clusters. The main idea is to define k centroids, one for each cluster. These centroids should be placed as much as possible far away from each other. K-means is a clustering method that aims to find the positions μ_i, i=1,2...k of the clusters that minimize the *distance* from the data points to the cluster.

Initialize the center of the clusters μ_i, where i=1.2……k. The K-means clustering uses the square of the Euclidean distance $d_{ij}(\mathbf{x_i}, \mu_i)=\|\mathbf{x_i}-\mu_i\|^2$.Set the position of each cluster to the mean of all data points belonging to that cluster

$$\mu_i = \frac{1}{|c_1|} \sum_{J \in c_1} X_j \tag{20}$$

.where μ_i is the mean and C_I is the cluster center. The cluster center is updated by using the above formula and the maximum probability corresponding to that particular center is updated.

3.2.1 Decision tree:

The following disturbances have been considered for power signal clustering in the order as depicted in the tree shown in Fig.2.

A. Transient., **B**. Harmonic, **C.** Notch**, D**. Sag., **E**. Voltage Spike, **F**. Sag+ Harmonic **.G**. Voltage Swell, **H**. Flicker, **I.** Swell + Harmonic

In contrast to neural networks, decision trees represent rules. Rules can readily be expressed so that humans can understand them or even directly used in a database. For the classification of the power quality events a binary decision tree is considered as follows: For classification, the feature vectors are first divided into two classes as (B,C,D,E,F,G,H,I) and A. Class-1 as A refers to Transient and Class-2 as (B,C,D,E,F,G,H,I) refers to the remaining power signal disturbances. Further the Class-2 (B,C,D,E,F,G,H,I) is sub divided into another Class-X as B and Class-Y as (C,D,E,F,G,H,I). In class–X, B represents Harmonic and (C, D, E, F, G, H, I) represents the rest of the power signals. This process is continued for the rest of power signal disturbances according to decision tree which is shown in the Fig. 2.

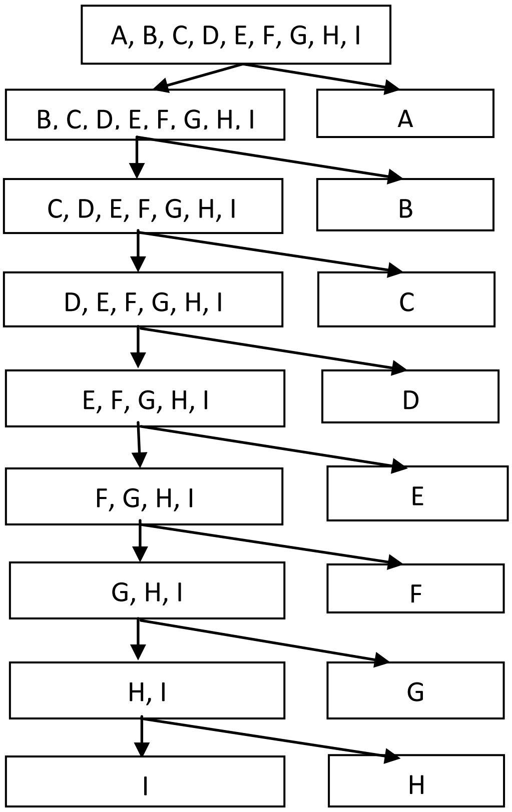

Fig2. Decision Tree for Classification

4. RESULTS AND DISCUSSION

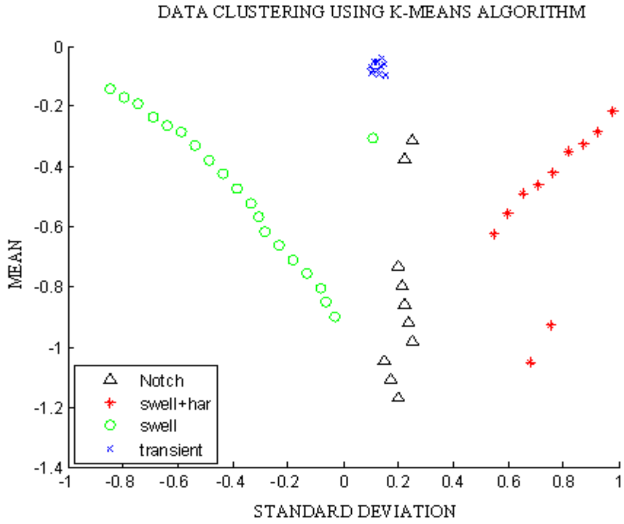

Fig.3 Classification of power signals using Fuzzy-C Means algorithm

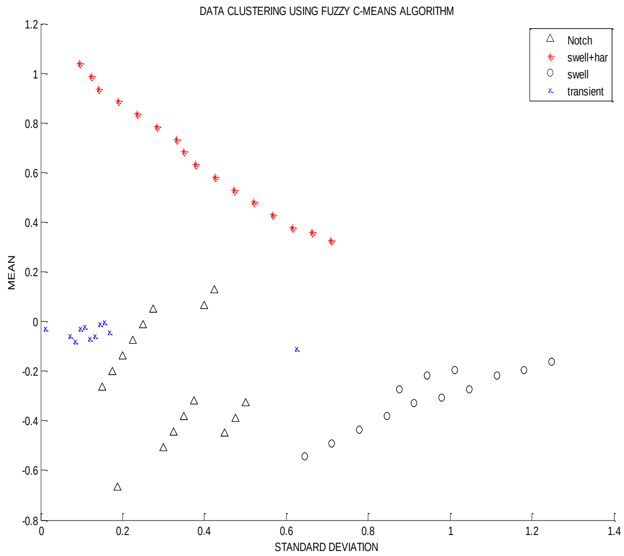

Fig4. Classification of Power Signals using K-means Algorithm

Table – 1

Sl. No.	Power Signal Disturbances	Accuracy in Percentage (%)	
		ACO-Fuzzy C-Means	ACO-K-Means
1	**Transient**	**94.66**	**100**
2	**Harmonic**	**92.58**	**98.35**
3	**Notch**	**92.57**	**97.42**
4	**Sag**	**97.48**	**98.37**
5	**Spike**	**94.87**	**97.47**
6	**Sag+Harmonic**	**93.25**	**99.44**
7	**Swell**	**95.98**	**98.56**
8	**Flicker, Sag+Harmonic**	**96.87**	**99.19**
%	**Accuracy**	**94.78**	**98.60**

From table-1, it is found that ACO k-means algorithm gives good percentage of accuracy than the ACO-Fuzzy C Means. In both the algorithm in the case of sag, it is found that the classification accuracy is nearly equal. Moreover the k-means algorithm is preferable for the classification of the power signals.

5. CONCLUSION

TT-Transform has potential to localize the power signal waveforms better than the S-Transform as TT-Transform localizes the spectrum diagonally. Ant Colony Optimization based k-means algorithm has achieved higher pattern recognition accuracy in classifying various power signal disturbances than the ACO-Fuzzy C Means classifier. From the simulation results, it is found that the proposed algorithm shows the better classification performance than the existing algorithms in power signal disturbance patterns classification.

6. REFERENCES

[1] C. R., Pinnegar, & L. Mansinha, (2203). A method of time-time analysis: The TT-transform. Elsevier Science on Digital Signal Processing, 13(4), 588–603

[2] P. K., Dash, Maya Nayak, , M.R Senapati,., & I. W. C. Lee, (2007). Mining for similarities in time series data using wavelet based feature vectors and neural networks. Engineering Applications of Artificial Intelligence, 20(2), 185–201.

[3] P. K., Dash, B. K., Panigrahi, & G. Panda, (2003). Power quality analysis using S-transform. IEEE Transactions on power delivery, 18(2), 406–411.

[4] Dorigo, M., & Gambardella, L. M. (1997). Ant colony system: A cooperative learning approach to the travelling salesman problem. IEEE Transactions on Evolutionary Computation, 1(1), 53–66.

[5] Dorigo, M., & Gambardella, L. M. (1997). Ant colony system: A cooperative learning approach to the travelling salesman problem. IEEE Transactions on Evolutionary Computation, 1(1), 53–66.

[6] Lei Jiang, Wenhui Yang, " A Modified Fuzzy C-Means Algorithm for Segmentation of Magnetic Resonance Images", Proc. VIIth Digital Image Computing: Techniques and Applications, Sydney ,10-12 Dec. 2003.

[7] B. Biswal, P. K. Dash, S. Mishra, B. Biswal, P.K. Dash, S. Mishra, "A Hybrid Ant Colony Optimization Technique For Power Signal Pattern Classification", Elsevier Science, Expert Systems With Applications,Vol.38, No.5, pp.6368-6375, 2011.

[8] R. G., Stock well, L. Mansinha, & R. P Lowe,., (1996). Localization of the complex spectrum: The S-transform. IEEE Transactions on Signal Processing, 44(4), 998–1001.

[9] Mohamed N. Ahmed, Sameh M. Yamany et.al, "A Modified Fuzzy C-Means Algorithm For Bias Field Estimation and Segmentation of MRI Data", IEEE Trans. Med. Imag. Vol.21, No.3, pp.193–199, 2002.

[10] C. T. Su, C. F. Chang and J.P. Chiou, "Distribution Network Reconfiguration for loss reduction by Ant Colony search algorithm", Electric Power Systems Research, Vol.75, No. 2-3, pp. 190-199, Aug. 2005.

[11] O. J Oyelade, O. O Oladipupo, I. C Obagbuwa, "Application of k-Means Clustering algorithm for prediction of Students' Academic Performance"International Journal of Computer Science and Information Security,Vol. 7, No. 1, 2010.