

## Classification of The NTEV Problems on the Commercial Building

Mohd Abdul Talib Mat Yusoh, Saidatul Habsah Asman, Zuhaila Mat Yasin, Ahmad Farid Abidin  
Universiti Teknologi MARA (UiTM), 40450 Shah Alam, Selangor, Malaysia

---

### Article Info

#### Article history:

Received Oct 9, 2017

Revised Dec 20, 2017

Accepted Jan 4, 2018

---

#### Keywords:

K-Nearest Neighbor (K-NN)  
classifier tools  
Neutral to Earth Voltage (PQ)  
Power Quality (PQ)  
S-Transform (ST)

---

### ABSTRACT

Neutral to Earth Voltage (NTEV) is one of power quality (PQ) problems in the commercial building that need to be resolved. The classification of the NTEV problems is a method to identify the source types of disturbance in alleviating the problems. This paper presents the classification of NTEV source in the commercial building which is known as the harmonic, loose termination, and lightning. The Euclidean, City block, and Chebyshev variables for K-Nearest Neighbor (K-NN) classifying are being utilized in order to identify the best performance for classifying the NTEV problems. Then, S-Transform (ST) is applied as a pre-processing signal to extract the desired features of NTEV problem for classifier input. Furthermore, the performance of K-NN variables is validated by using the confusion matrix and linear regression. The classification results show that all the K-NN variables capable to identify the NTEV problems. While the K-NN results show that the Euclidean and City block variables are well performed rather than the Chebyshev variable. However, the Chebyshev variable is still reliable as the confusion matrix shows minor misclassification. Then, the linear regression outperformed the percentage close to a perfect value which is hundred percent.

Copyright © 2018 Institute of Advanced Engineering and Science.  
All rights reserved.

---

### Corresponding Author:

Mohd Abdul Talib Mat Yusoh,  
Universiti Teknologi MARA (UiTM),  
40450 Shah Alam, Selangor,  
Malaysia  
Email: mohd.abdul.talib@gmail.com

---

## 1. INTRODUCTION

High Neutral to Earth Voltage (NTEV) in the commercial building is one of the concerns in Power Quality (PQ) disturbances due to hazardous to the humans, animals, electric and electronic appliances, and electrical networks system [1],[2]. Many researchers have discussed the high NTEV in the commercial building is due to improper wiring, poor grounding system, nonlinear load, and cable damaged [3–6]. However, the discussion is still in assumption to determine the source types of NTEV as in [7] and have not been proved by any artificial intelligent techniques.

Normally, the source types of PQ disturbance in the commercial building can be identified by using the classification technique [8]. Nowadays, a lot of classifier tools are combined with the signal processing technique to classify the PQ disturbances in the commercial building [9–11]. The signal processing is utilized to extract the desired information which is applied as an input for the classifier tools [8]. Then, some of the disturbances introduce other problems in signal processing due to limitation and cannot be analyzed to extract the desired information accurately [12],[13]. Hence, an appropriate variable in classifier tool and signal processing technique need to be utilized to identify the types of NTEV problem in the commercial building. The NTEV occurrence needs to be aware of due to hazardous to humans, animals, electrical appliances, and electrical network system [14].

The main contribution of this paper is to classify the source types of NTEV problems in the commercial building due to the triplen harmonic and transient. The transient can be divided into two categories; transient due to the loose termination and transient due to the lightning strike. Then, the K-Nearest Neighbor (K-NN) is selected as the classifier tool and has been tested with different variables, which are Euclidean, City block, and Chebyshev. The different variables of K-NN are utilized to identify the best performance in classifying the NTEV problem. Further, S-Transform (ST) is chosen to extract the features of NTEV problem according to the statistical analysis algorithms. ST technique is selected due to the capability in processing the signal without missing any information features, variables window scale, and outperform wavelet transform (WT) and short-time fourier transform (STFT). Finally, the performances of K-NN are observed based on the different variables by using confusion matrix and linear regression. According to the confusion matrix and linear regression results, the performance of K-NN classifier tools is analysed based on the percentages of accuracy, misclassification and the relationship between the target and output results.

**2. RESEARCH METHOD**

To classify the NTEV problems in the commercial building, several methods such as the ST operation, K-NN classifier tool should be followed

**2.1. The S-Transform (ST) Theory**

ST technique is a time-frequency pre-processing signal, which presents the result in complex number spectrum. Then, ST also produces the result in S-matrix which consist the multiple numbers of columns and rows. The ST of signal x(t) can be defined as follow [13]:

$$S(\tau, f) = \frac{|f|}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x(t) e^{\frac{-(t-\tau)^2 f^2}{2}} e^{-j2\pi ft} \tag{1}$$

According to (1), let  $\tau = kT$  and  $f = \frac{n}{NT}$ , the discrete ST is given by:

$$S\left[kT, \frac{n}{NT}\right] = \sum_{m=0}^{m=N-1} X\left[\frac{m+n}{NT}\right] e^{\frac{2n^2 m^2}{n^2}} e^{-2\pi mnk}, n \neq 0 \tag{2}$$

where,  
 k, m, n=0, 1, ..., N-1  
 T=sampling interval  
 N=total of sampling point

According to the ST result, the features of NTEV problem are extracted by using the statistical analysis technique, that include the standard deviation (3), mean (4), variance (5), skewness (6), kurtosis (7), and total harmonic distortion (THD) (8). Figure 1 shows the operation to classify the NTEV problem in the commercial building included ST, statistical analysis, and K-NN classifier tools. Then, the algorithms of statistical analysis are given as [15]:

$$\sigma = \max\left[\frac{1}{M-1} \sum_{i=1}^M (Y_{ij} - \bar{Y}_j)^2\right]^{1/2} \tag{3}$$

$$\bar{Y} = \mu = \max\left(\frac{1}{M} \sum_{j=1}^M Y_{ij}\right) \tag{4}$$

$$\sigma^2 = \max\left(\frac{1}{M-1} \sum_{i=1}^M (Y_{ij} - \bar{Y}_j)^2\right) \tag{5}$$

$$s = \max \left( \frac{\frac{1}{M} \sum_{j=1}^M (Y_{ij} - \bar{Y}_j)^3}{\left( \left( \frac{1}{M-1} \sum_{j=1}^M (Y_{ij} - \bar{Y}_j)^2 \right)^{1/2} \right)^3} \right) \tag{6}$$

$$k = \max \left( \frac{\frac{1}{M} \sum_{j=1}^M (Y_{ij} - \bar{Y}_j)^4}{\left( \frac{1}{M-1} \sum_{j=1}^M (Y_{ij} - \bar{Y}_j)^2 \right)^2} \right) \tag{7}$$

$$THD = \max \left( \frac{\left( \sum_{j=2}^M Y_{ij}^2 \right)^{1/2}}{Y_{ij=1}} \right) \tag{8}$$

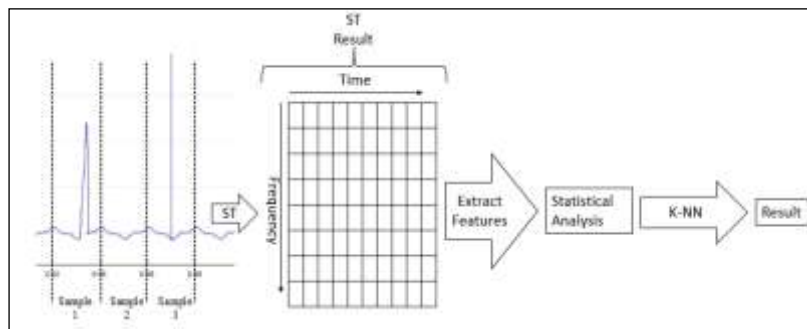


Figure 1. The Classification of NTEV Problem

Based on the equation in (3)-(8), the result of NTEV features can be obtained and utilized as an input for K-NN classifier.

**2.2. K-Nearest Neighbor (K-NN)**

K-NN can be categorized as the Nearest Neighbor (NN) family [16]. K-NN is reliable and easy to implement for classification by using non-parametric and lazy learning technique [17]. Basically, the K-NN classification operates based on the distance metric of input classifying and testing samples. Then, the performance of K-NN also depends on the selected number of K. Thus, the number of K is selected based on the best values before analyzing the performance of K-NN using the different variables. This paper shows the testing for different variables of K-NN in order to identify the best performance of K-NN in classifying the NTEV problem. The variables of K-NN such as the Euclidean (9), City block (10), Chebyshev (11) are used to analyze the NTEV problems in the commercial building. The mathematical algorithms are given below [18]:

$$\sqrt{\sum_{i=1}^k (x_i - y_i)^2} \tag{9}$$

$$\sum_{i=1}^k |x_i - y_i| \tag{10}$$

$$\max (|x_2 - x_1|, |y_2 - y_1|) \tag{11}$$

**3. RESULTS AND ANALYSIS**

Figure 2 shows the input of classification which utilized on different variables of K-NN classifier tools. The number of samples that have been used as input and testing data are 124 and 126, respectively. Based on the samples data, K-NN analysis is utilized to identify the best categories of K-NN variables in classifying the NTEV problems.

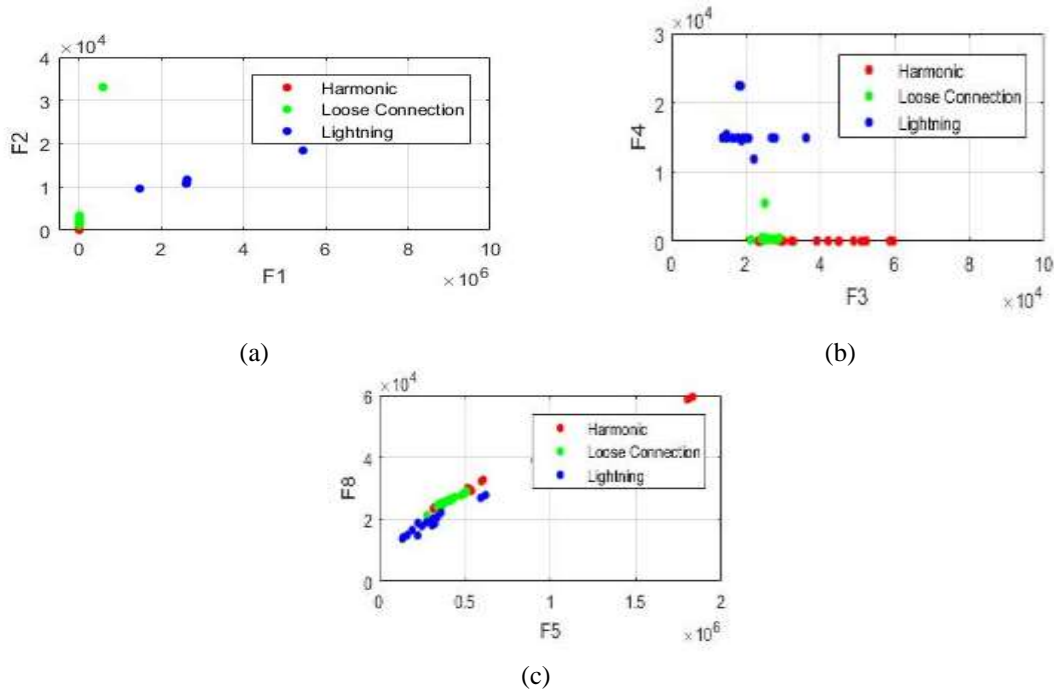
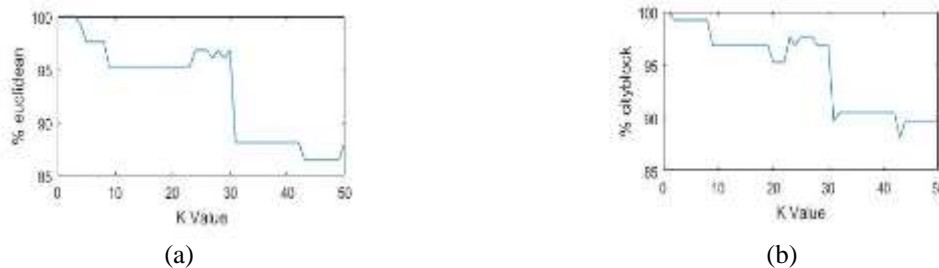


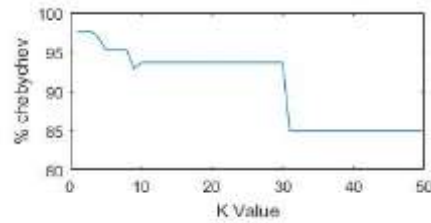
Figure 2: The plot classification input features: (a) F1 and F2, (b) F3 and F4, and (c) F5 and F8

**3.1. Performance of K-NN Variables Based on Different K Values**

The K-NN classification uses 50 iterations with a different number of K.

Figure 3 shows each different values of K produce the different results of K-NN variables. The best values of K for variables Euclidean, City block, and Chebyshev are between 1 until 4. The number of K between 1 until 4 shows the identical results with respect to their performance. Based on the best value of K, the K-NN is analyzed by using the confusion matrix technique and linear regression.





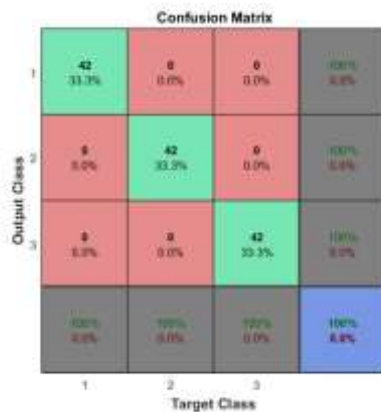
(c)

Figure 3 Performance of K value for K-NN types (a) Euclidean, (b) Cityblock, (c) Chebyshev

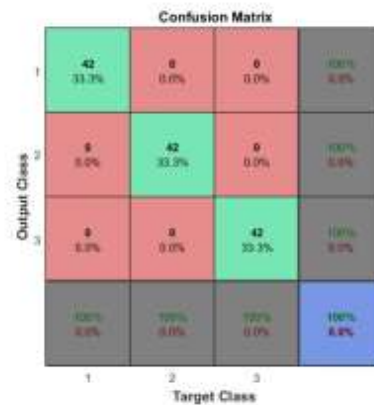
**3.2. Confusion Matrix Results**

Figure 4 shows the result of confusion matrix which has been tested by using the K-NN classifier of the different variables such as Euclidean, City block, and Chebyshev. The classes 1, 2, and 3 represents the NTEV problems due to the harmonic, loose connection, and lightning, respectively. Based on the figure, the numbers of samples for class 1, class 2, and class 3 are identical 42. The result of variable Euclidean shows that the class 1, class 2, and class 3 are 33.3%. The result of K-NN for variable City block seems similar with variable Euclidean. Then for the variable Chebyshev shows the results for class 1, class 2, and class 3 are 33.3%, 31%, and 33.3%, respectively. The correct number classifies for variable Chebyshev are 42, 39, and 42 for class 1, class 2, and class 3, respectively.

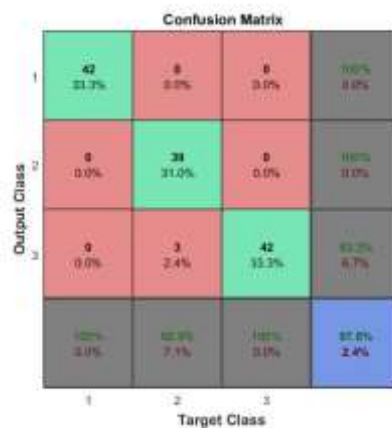
Overall results show that the K-NN classifier using the Euclidean and City block variables are the best results on classifying the NTEV problems with 100% accuracy.



(a)



(b)



(c)

Figure 4. The K-NN Confusion Matrix for variable (a) Euclidean (b) City block (c) Chebyshev

### 3.3. Regression Results

Table 1 elaborates the performance of K-NN based on the different variables by using linear regression. According to the table, the K-NN results for variables Euclidean and City block are 100% for all classes. That means the relationship between the target and the output are linear proportional. However, the variable Chebyshev shows the results for class 1 and class 3 are hundred percent linear. Then, the variable Chebyshev shows the class 2 result is 94.69% can be explained by the linear regression. The total results of K-NN are 100%, 100%, and 98.23% for variables Euclidean, City block, and Chebyshev, respectively. The variables Euclidean and City block show the results are similar to each other and represent as the best performance in analyzing the problems of NTEV due to the harmonic, loose termination, and lightning in the commercial building. However, the Chebyshev result is still dependable, since the relationship between the target and the output are close to the hundred percent.

Table 1. Performance of K-NN based on the different variables

Variables	Class 1 (%)	Class 2 (%)	Class 3 (%)	Total (%)
Euclidean	100	100	100	100
City block	100	100	100	100
Chebyshev	100	94.69	100	98.23

## 4. CONCLUSION

This paper is presented to classify the NTEV problem in the commercial building by using the K-NN classifier tools. The K-NN classifier tool uses the different continuous variables, which are Euclidean, City block, and Chebyshev. The ST is applied as the proposed technique to extract the features of NTEV problem, which is then utilized as classification input. The performance of K-NN classifier with different variables is observed based on the confusion matrix and linear regression results. The Euclidean and City block variables presented the best result to classify the NTEV problems in the commercial building compared to the Chebyshev variable. Even though the performance of the Chebyshev variable is lower compared to the other variables, it is still reliable as the confusion matrix shows minor misclassification which is 2.4%. Then, the linear regression shows that the percentage of Chebyshev result close to perfect value which is hundred percent.

## ACKNOWLEDGEMENTS

This work was supported in part by Universiti Teknologi Mara (UiTM) under the scholarship.

## REFERENCES

- [1] D. Dorr, "Determining voltage levels of concern for human and animal response to AC current," in *Power & Energy Society General Meeting, 2009. PES '09. IEEE*, 2009, pp. 1–6.
- [2] S. Ravlić and A. Marušić, "Simulation Models for Various Neutral Earthing Methods in Medium Voltage Systems," *Procedia Eng.*, vol. 100, pp. 1182–1191, 2015.
- [3] D. Salles *et al.*, "Assessing the Collective Harmonic Impact of Modern Residential Loads-2014;Part I: Methodology," *Power Deliv. IEEE Trans.*, vol. 27, no. 4, pp. 1937–1946, Oct. 2012.
- [4] A. M. Variz, J. L. . Pereira, S. Carneiro, and P. G. Barbosa, "Harmonic analysis of the power distribution Neutral-to-Earth Voltage (NEV) test case using four-wire three-phase harmonic current injection method," in *IEEE Power & Energy Society General Meeting*, 2009, pp. 1–7.
- [5] J. D. Bouford and D. J. Hurley, "A Recommended Standard for Voltages Appearing Across Publicly Accessible Surfaces," *IEEE Trans. Power Deliv.*, vol. 29, no. 1, pp. 48–51, Feb. 2014.
- [6] M. A. T. M. Yusoh, A. F. Abidin, Z. M. Yasin, M. F. Ghani, and U. Mat, "Modelling Neutral to Earth Voltage ( NTEV ) on the Commercial Building," *2016 IEEE Conf. Syst. Process Control*, pp. 91–96, 2016.
- [7] C. Jiang, Q. Shi, Y. Tian, X. Li, S. Lin, and Y. Liu, "Sensitivity study on the stray voltage of low voltage residential networks," *2016 IEEE 8th Int. Power Electron. Motion Control Conf. IPEMC-ECCE Asia 2016*, pp. 2571–2576, 2016.
- [8] K. Thirumala, A. C. Umarikar, and T. Jain, "Estimation of single-phase and three-phase power-quality indices using empirical wavelet transform," *IEEE Trans. Power Deliv.*, vol. 30, no. 1, pp. 445–454, 2015.
- [9] A. Asuhaimi Mohd Zin, M. Saini, M. W. Mustafa, A. R. Sultan, and Rahimuddin, "New algorithm for detection and fault classification on parallel transmission line using DWT and BPNN based on Clarke's transformation," *Neurocomputing*, vol. 168, pp. 983–993, 2015.
- [10] O. Ozgonenel, T. Yalcin, I. Guney, and U. Kurt, "A new classification for power quality events in distribution systems," *Electr. Power Syst. Res.*, vol. 95, pp. 192–199, 2013.
- [11] J. H. A. Morales, E. A. Orduña, and C. Rehtanz, "Classification of lightning stroke on transmission line using multi-resolution analysis and machine learning," *Electr. Power Energy Syst.*, vol. 58, pp. 19–31, 2014.

- 
- [12] A. Rodriguez, J. A. Aguado, F. Martin, J. J. Lopez, F. Munoz, and J. E. Ruiz, "Rule-based classification of power quality disturbances using S-transform," *Electr. Power Syst. Res.*, vol. 86, pp. 113–121, 2012.
- [13] A. F. Abidin, A. Mohamed, and H. Shareef, "Intelligent detection of unstable power swing for correct distance relay operation using S-transform and neural networks," *Expert Syst. Appl.*, vol. 38, no. 12, pp. 14969–14975, 2011.
- [14] C. DeNardo, "Stray and contact voltage an update on IEEE PES P1695 Working Group activities," in *Rural Electric Power Conference*, 2012, pp. A3-1-A3-7.
- [15] N. Huang, D. Xu, X. Liu, and L. Lin, "Power quality disturbances classification based on S-transform and probabilistic neural network," *Neurocomputing*, vol. 98, pp. 12–23, 2012.
- [16] M. A. Jabbar, B. L. Deekshatulu, and P. Chandra, "Classification of Heart Disease Using K- Nearest Neighbor and Genetic Algorithm," *Procedia Technol.*, vol. 10, pp. 85–94, 2013.
- [17] A. Kumar, R. Bhatnagar, and S. Srivastava, "ARSkNN-A k-NN classifier using mass based similarity measure," *Procedia Comput. Sci.*, vol. 46, no. Icict 2014, pp. 457–462, 2015.
- [18] R. A. Nugrahaeni and K. Mutijarsa, "Comparative analysis of machine learning KNN, SVM, and random forests algorithm for facial expression classification," in *2016 International Seminar on Application for Technology of Information and Communication (ISemantic)*, 2016, pp. 163–168.