FORECASTING VIETNAM'S ELECTRIC LOAD PROFILE TO 2030

DỰ BÁO BIỂU ĐỒ PHỤ TẢI ĐIỆN VIỆT NAM TỚI NĂM 2030

Nguyen Hoang Minh Vu[1] Nguyen Ngoc Au[1], Vo Viet Cuong[1], Phan Thi Thanh Binh[2]
[1]Ho Chi Minh City University of Technology and Education, Vietnam
[2]Ho Chi Minh City University of Technology, Vietnam

Received 31/7/2018, Peer reviewed 6/8/2018, Accepted for publication 22/8/2018.

ABSTRACT

Long term load profile forecasting is really difficult but very nessesary for dynamic programing in power system planning. The purpose of this paper is to forecaste load profile of electric power system of Vietnam to 2030. Kmax - Kmin algorithm combining with expert selection are applied to findout load patterns of power system in 2006, 2010, 2012, and 2014. Similarity of load curve shapes of the patterns are recognised, and be used for forecating.

The results show that there are 8 load paterns in the past from 2006. In which, load paterns in years of 2010, 2012, and 2014 have same shapes or "rules". They are used to forcaste load profile of those paterns to 2030. Total load demands (GWH) come from the forcasted load profile are less than 2% difference with corresponding given load demands from previous study. Theses results are quite worldwide acceptable for this kind of study.

Keywords: *Electric load profile; Clustering; Load pattern; Forecasting; Vietnam.*

TÓM TẮT

Dự báo dài hạn biểu đồ phụ tải là công việc rất khó khăn nhưng hết sức cần thiết cho lập trình động trong quy hoạch hệ thống điện. Mục tiêu của bài báo này là dự báo biểu đồ phụ tải hệ thống điện Việt Nam tới năm 2030. Giải thuật Kmax – Kmin kết hợp với sự lựa chọn của chuyên gia được sử dụng để tìm các mẫu phụ tải của hệ thống điện cho các năm 2006, 2010, 2012, và 2014. Những mẫu có hình dạng giống nhau được nhận diện, và được sử dụng để dự báo.

Kết quả chỉ ra có 8 mẫu phụ tải từ năm 2006. Trong đó, các mẫu phụ tải của các năm 2010, 2012, và 2014 có hình dạng hay "luật" giống nhau. Chúng được sử dụng để dự báo biểu đồ phụ tải của các mẫu phụ tải tới năm 2030. Nhu cầu điện (GWH) có được từ biểu đồ phụ tải dự báo có khác biệt nhỏ, dưới 2%, so với những giá trị tương ứng được cho trong nghiên cứu trước. Kết quả như vậy được chấp nhận rông rải cho dạng nghiên cứu này.

Từ khóa: *Biểu đồ phụ tải điện; Phân nhóm; Mẫu phụ tải; Dự báo; Việt Nam.*

1. INTRODUCTION

Load profile clustering task is used for many purpose such as operating, planning power system, and building solution for demand side management. In planning, a dynamic program is the most used which can give the most acurrate optimum results by running simulation with hourly load value. Input data for the progam is hourly forecasted load profile. However, they create a huge number of variable if running with 8,670 values of load demand for a year. In order to reduce the variable for the programming, load profile clustering to find load patterns is needed.

There are various techniques has been purposed to deal with issue of load profile clustering. In particular, it is possible to identify unsupervised learning base techniques, such as Kohonen's self oganising map (SOM) which present a bi-dimensional map as results, but does not deliver final cluster directly, it need a post-processing stage for getting final results [1, 2, 3, 4]. Supervised learning based techniques (ANN), vector quantization, fuzzy logic based techniques, statistical techniques such as K-means (KM), hybrid techniques such as probability neuron network (PNN) [5], and fuzzy K-means (FKM). Recently definded techniques following the concept of entropy which is come from information theory such as follow the leader (FDL) and support vector clusting (SVC). These technics are not required a fine number of clusters as input, but use an internal distance threshold among cluster centroids [6, 7, 8].

On the application side, the clustering techniques differ is mainly depended on their usage. For techniques such as *k-means* (KM) and *fuzzy k-means* (FKM) [9, 10], accepting the final number of cluster as input, in a few case resulted of empty clusters. With internal steps depending on random number extractions, the desired number of clusters can be archived by running the program again. Basic k-means and fuzzy concept have been present in [11, 12, 13, 14]. An unsupervised learning based techniques of Adaptive Vector Quantization (AVQ) method is ued in [12].

Regardless of the specific details of the clustering methods, clustering algorithm output is only need to alocate daily load profiles (DLPs) to the clusters, which two-dimension of number of the clusters and list of DLPs in each cluster. In order to evaluate effectiveness of the clustering methods, validity indicators have been defined. Most of these indicators are based on Euclidean distance metrics. Different types of distance are defined in [15, 16].

Regardless which method is used, resulted clusters creates many confusions by their unexplainable. That is why they need a re-arrangement by a "filter" called expert who has many experiencies in load profile. The expert will re-grouping the clusters based on factors of time of month/season in a year, and holiday/working day.

The purpose of this paper is to forecate load profile of load patterns of electric of power system for years of 2020, 2025, and 2030 in Vietnam. Kmax - Kmin algorithm combining with expert selection are applied to findout load patterns of power system in 2006, 2010, 2012, and 2014. Similarity of load cure shapes of the patterns are recognised and be used for forecating those load patterns for the future.

2. METHODOLOGY

Figure 1 shows proceduce for load pattern clustering and load profile forecasting [17].

2.1 Data gathering & processing phase

Hourly load profile data of the last teen years should be collected.

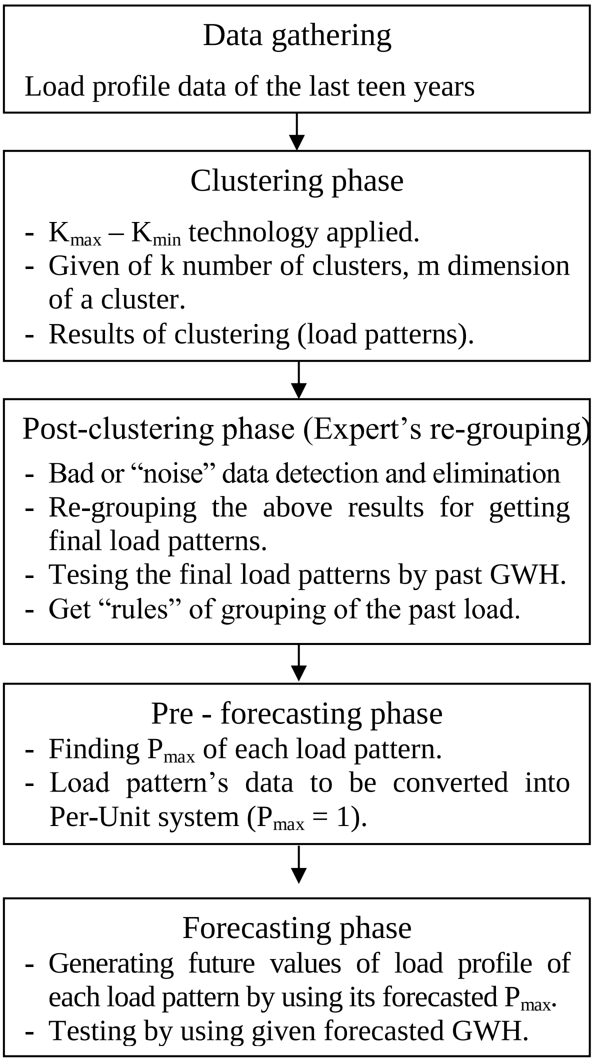

Fig. 1. *Load pattern clustering and load profile forecasting procedure*

Bad or "noise" data detection and elimination. The noise data is data of a day in which the power system was shutdowned.

2.2 Clustering phase

$K_{max} - K_{min}$ technology to be applied [18].

Input: $X = \{x_i\}$, $1 \le i \le m$ – the given samples in R^n. n – the problem's dimension. m – the number of samples.

Set k clusters.

Output: $\{y_j\}$, $1 \le j \le k$ – the cluster centers.

$\{m_j\}$, $1 \le j \le k$ – the cluster sizes.

$\{l_{ij}\}$, $1 \le i \le m_j$ – the indices of the original samples which belong to the j – th cluster, $1 \le j \le k$.

Step 1. Set $y_1 = x_1$, $y_2 = x_{j0}$, $l_{11} = 1$, $l_{12} = j_0$ where

$$M = \|x_{j_0} - y_1\| = \max_{2 \le i \le m} \|x_i - y_1\|$$

Set k =2, $i \neq j$ and $X' = X - \{y_1, y_2\}$, where $1 \le i, j \le k$, t= M/k

Step 2. Find j_0, $1 \le j_0 \le k$ and $x_{i_0} \in X'$ such that

$$d = \|x_{i_0} - y_{j_0}\| = \max_{x_i \in x'} \min_{1 \le j \le k} \|x_i - y_j\|$$

If d > t go to Step 4; otherwise go to Step 3.

Step 3. Set $k \leftarrow k + 1$, $y_{k+1} = x_{i_0}$, $l_{k_1} = i_0$, $X' \leftarrow X' - \{y_{k+1}\}$ and go to Step 2.

Step 4. Set $m_j = 1$, $1 \le j \le k$.

Step 5. For each $x_i \in X'$ find j: $1 \le j \le k$ for which

$$\|x_i - y_j\| = \min_{1 \le j \le k} \|x_i - y_j\|$$

and set $m_j \leftarrow m_j + 1$ and $l_{m_j j} = i$.

Step 6. For $1 \le j \le k$

Replace y_j by $(x_{l_1 j} + x_{l_2 j} + \ldots + x_{l_{m_j} j}) / m_j$

Step 7. For $1 \le j \le k$

Output y_j, m_j, $\{l_{ij}\}_{i=1}^{m_j}$ and stop.

Results of clustering (load patterns) appear.

2.3 Post-clustering phase (Expert's re-grouping)

The about results are mostly unexpainable by way of behaviour of loads. Because they were grouped by their internal distance thresholds among cluster centroids.

Re-grouping the above results for getting final clusters. Expert will re-grouping based on factors of time of month/season in a year, and holiday/working day. Then, each final load pattern is composed by averaging values of daily loads which are belong to the cluster.

Arcuracy testing for those above final load patterns by using load data of these patterns multiple with its number of days which belonging to the pattern. Sigma of those multiple is the GWH demand for a year.

After re-grouping by expert, "rules" of grouping of the past load will be deliver.

2.4 Pre - forecasting phase

Finding P_{max} of each of above load patterns among pattern's data. Then, data of load patterns to be converted into Per-Unit system with P_{max} of each pattern equal 1.

2.5 Forecasting phase

Building poportion of P_{max} of each pattern in P_{max} of a year. Finally, generating values of future load by using given forecasted P_{max}.

Testing results by using given forecasted GWH. Difference bettwen GWH composed from the forecasting load patterns and the given forecasted GWH is criteria.

3. IMPLEMENTATION - RESULTS

3.1 Results of data gathering

Gathering enough data for the research is quite a big challenge. Hourly load profile data of 2006, 2010, 2012, 2014 is collected. There are absent data of 2008 and 2016.

3.2 Results of clustering phase

After testing with some cases of K – number of clusters are 3, 5, 8, 9, and 12. The best one, which create reasonale of cluters, to be sellected are 9 for 2006 and 8 for 2010, 2012, 2014 as input. n-dimension of a cluster is 2 which are daily load profile, and days which belong to the cluster. Here after are some results of clustering phase:

Table 1. *Load patterns of clustering phase*

No.	Pattern	2006	2010	2012	2014
1	Tet hollidays	X	X	X	X
2	W: 1, 2		X	X	
3	W: 1, 2, 3	X			X
4	W: 3, 4		X		
5	W: 3, 4, 5			X	
6	W: 4, 5				X
7	W: 4, 5, 8	X			
8	W: 5, 6		X		
9	W: 6, 7	X			
10	W: 7, 8		X		
11	W: 6, 7, 8			X	X
12	W: 9, 10, 11, 12	X	X	X	X
13	S & N: 1, 2, 3	X			
14	S & N: 1, 2		X		X
15	S & N: 1, 2, 3, 4			X	
16	S & N: 3, 4, 5		X		X
17	S & N: 4, 5, 8	X			
18	S & N: 5, 6, 7, 8, 9		X	X	X
19	S: 6, 7	X			
20	S & N: 9, 10, 11, 12	X			
21	S & N: 10, 11, 12		X	X	X

Noted:

W: Working day; S: Sunday;

S & N: Sunday & National holiday

Figure 2 shows a typical load pattern of working days in Sep., Oct., Nov., Dec.. There is a noise (bad) data on Dec., 27, 2006 which may be caused by a shutdown of a part of the power system. Furthermore, even most of lines are belong to working days in Sep., Oct., Nov., Dec., but there are number of lines come from working day of Jan., Mar., Apr., and May. This issue is also happen to other clusters. It is just because the nature of number.

3.3 Results of Post-clustering phase (Expert's re-grouping)

Based on factors of time of month/season in a year, and holiday/working day, expert re-grouping the cluters in the previous phase. Final load patterns after re-grouping is presented in table 2 as below. Total reasonale pattern is 8.

A typical result of post - clustering phase for the pattern of working days in Sep., Oct., Nov., Dec. is shown in figure 3. This one to be gotten by eliminating the noise and the lines which are not belong to working days in Sep., Oct., Nov., Dec.. Testing results are shown in the table 3 in which the difference between GWH come from load patterns and real past GWH demands are less than 1% for year of 2010, 2012, and 2014. It is a quite very good acuracy of the clusting.

3.4 Pre - forecasting phase

Load pattern's data to be converted into Per-Unit system (P_{max} = 1) for pattern of working day in Sep., Oct., Nov., Dec. is shown in figure 4. A surpising result is that load patterns of 2010, 2012, and 2014 are nearly the same. It means that the "shape" of this pattern will be like this in the future. Load data of 2006 will be exclued from now on. The average line of these 3 lines will be the "rule" line for this pattern in the future.

Table 2. *Patterns of post - clustering phase*

No.	Pattern
1	Tet hollidays
2	W: 1, 2
3	W: 3, 4, 5
4	W: 6, 7, 8
5	W: 9, 10, 11, 12
6	S & N: 1, 2
7	S & N: 3, 4, 5, 6, 7, 8, 9
8	S & N: 10, 11, 12

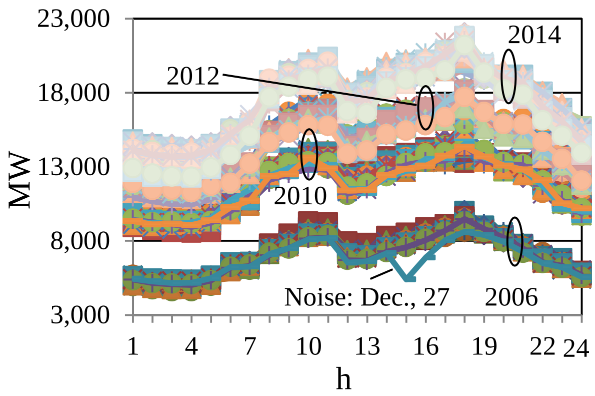

Fig. 2. *Load pattern of working day in Sep., Oct., Nov., Dec.*

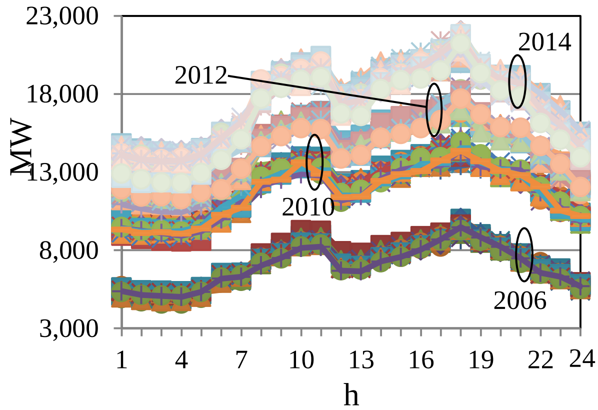

Fig. 3. *After re - grouping of the pattern of working day in Sep., Oct., Nov., Dec.*

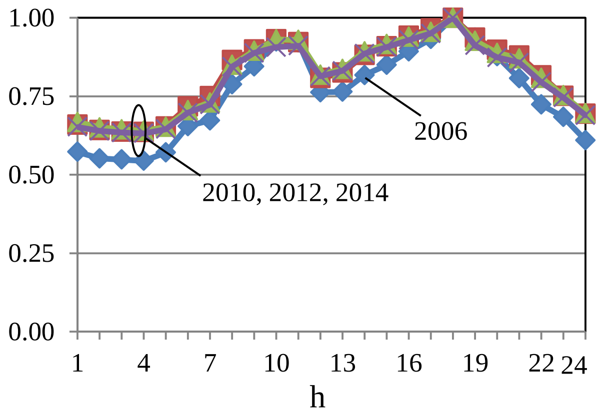

Fig. 4. *The pattern of working day in Sep., Oct., Nov., Dec. in the per unit system*

3.5 Results of forecasting phase

Average poportion of P_{max} of each pattern in P_{max} of a year in the three years of 2010, 2012, and 2014 is presented in figure 5. P. 5 is pattern No. 5 in table 2 (W: 9, 10, 11, 12). This line will be line for this patten in 2020, 2025, and 2030.

Forecasted P_{max} of 2020, 2025, and 2030 are shown in table 4. These values are collected from the previous paper of the authors. P_{max} values of each pattern in 2020, 2025, and 2025 are presented in table 5. These values are used to generate load profile of a pattern. Forecasted load profile of pattern of working day in Sep., Oct., Nov., Dec. in 2020, 2025, and 2025 are presented in figure 6. The rest 7 paterns are also to be builded follow this way. Testing results are shown in table 6. There are differences between GWH come from load profile of forecasted load and from given forecated of less than 2%.

Table 3. *Difference between GWH come from load patterns and real past GWH demands*

Year	Real past GWH	GWH from load patterns	Difference [%]
2010	99,199	99,896	0.72
2012	118,500	119,617	0.94
2014	144,316	143,283	- 0.76

Table 4. *Forecasted GWH & Pmax of Vietnam to 2030 [19, 20]*

Year	GWH	Pmax (MW)
2020	230,195	40,332
2025	349,949	60,835
2030	511,268	87,558

Table 5. *P_{max} values of each pattern in 2020, 2025, and 2025 [MW]*

Pattern	**2020**	**2025**	**2030**
Tet hollidays	24,603	37,109	53,410
W: 1, 2	35,492	53,535	77,051
W: 3, 4, 5	39,929	60,227	86,682
W: 6, 7, 8	39,929	60,227	86,682
W: 9, 10, 11, 12	40,332	60,835	87,558
S & N: 1, 2	31,056	46,843	67,420
S & N: 3, 4, 5, 6, 7, 8, 9	33,879	51,101	73,549
S & N: 10, 11, 12	36,299	54,752	78,802

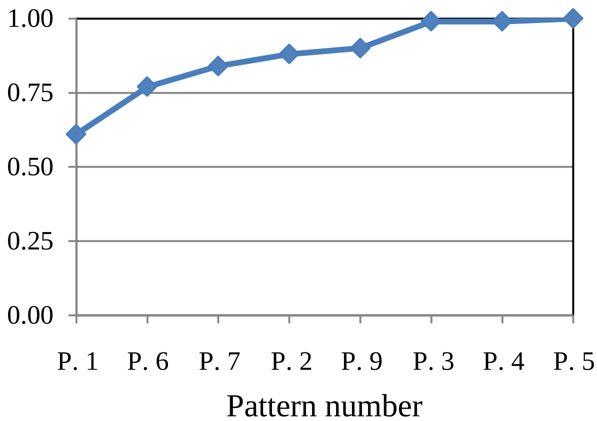

Fig. 5. *Poportion of P_{max} of each pattern in P_{max} of a year*

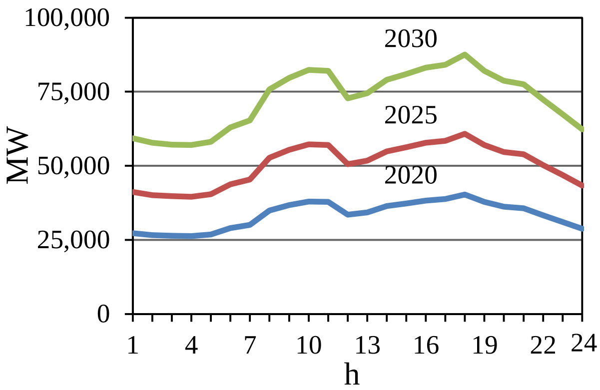

Fig. 6. *Forecasted load profile of pattern of working day in Sep., Oct., Nov., Dec. in 2020, 2025, and 2025*

Table 6. *Differences between GWH comes from forecasted load patterns and comes from given forecasted [19]*

Year	Given forecasted GWH	GWH from load patterns	Difference [%]
2020	230,195	229,177	- 0.93
2025	349,949	347,424	- 1.46
2030	511,268	506,462	- 1.92

4. CONCLUSION

Long term load profile forcasting is never be an easy task. Kmax - Kmin algorithm combining with expert selection are applied to findout load patterns of power system. Similarity of load curve shapes of the patterns are recognised, and be used for forecating those load profile of the load patterns to 2030.

Load profile data of 2006, 2010, 2012, and 2014 are used. The total of 8 past load paterns or “rules”, after exper’s selection, are found for load profile of 2010, 2012, and 2014. Testing results for these “rules” are less than 1% difference with the real past one.

The next step is using those 8 “rules” appling for getting load profile for years of 2020, 2025, and 2030. Testing results show the differences between GWH come from load profile of forecasted load patterns and from given forecated is less than 2%. These are very good results which make the results getting a quite good reliability.

REFERENCES

[1] Nazarko J, Styczynski ZA. Application of statistical and neural approaches to the daily load profiles modelling in power distribution systems. Proc. *IEEE Transm. and Distrib. Conference*, New Orleans, LA, 11-16 April 1999, 1:320-325.

[2] Kohonen T. Springer Series in Information Science. *Self-organizing maps*, 30. Berlin, Germany: Springer-Verlag; 1995.

[3] Chicco G, Napoli R, Piglione F, Scutariu M, Postolache P, Toader C. Load pattern-based classification of electricity customers. *IEEE Trans. Power Syst 2004*;19(2):1232-9.

[4] Valero S, Ortiz M, Senabre C, Alvarez C, Franco FJG, Gabaldon A. Methods for customer and demand response policies selection in new electricity markets. *IET Generation, Transm Distribution 2007*;1(1): 104-10.

[5] Gerbec D, Gasperic S, Smon I, Gubina F. Allocation of the load profiles to consumers using probabilistic neural networks. *IEEE Trans. Power Syst 2005*; 20(2):548-55.
[6] Chicco G, Napoli R, Postolache P, Scutariu M, Toader C. Customer characterisation options for improving the tariff offer. *IEEE Trans. Power Syst 2003*;18(1):381-7.
[7] Pao YH, Sobajic DJ. Combined use of unsupervised and supervised learning for dynamic security assessment. *IEEE Trans. Power Syst 1992*; 7:878-84.
[8] Batrinu F, Chicco G, Napoli R, Piglione F, Scutariu, M Postolache P, Toader C.Efficient iterative refinement clustering for electricity customer classification. *Proc. IEEE power Tech 2005*, St. Petersburg, Russia, 27-30 June 2005, paper no.139.
[9] Anderberg MR. Cluster analysis for applications. New York: *Academic Press*;1973.
[10] Bezdek JC, Harris JD. Fuzzy partitions and relations; an axiomatic basis for clustering. *Fuzzy Sets Syst 1978*; 1:111-27.
[11] Gerbec D, Gašperič S, Šmon I, Gubina F. Determining the load profiles of consumers based on fuzzy logic and probability neural networks. *IEE Proc Gener Transm Distrib 2004*;151(3):395-400.
[12] Tsekouras GJ, Hatziargyriou ND, Dialynas EN. Two-stage pattern recognition of load curves for classification of electricity customers. *IEEE Trans. Power Syst 2007*; 22(3):1120-8.
[13] Marques DZ, de Almeida KA, de Deus AM, da Silva Paulo ARG, da Silva Lima W. A comparative analysis of neural and fuzzy cluster techniques applied to the characterization of electric load in substations. *Proc. IEEE/PES Transmission and Distribution Conference and Exposition*: Latin America, 8-11 Nov. 2004, 908-913.
[14] Zadeh L. Similarity relations and fuzzy orderings. *Inf Sci 1971*; 3:177-200.
[15] Chicco G, Napoli R, Piglione F, Scutariu M, Postolache P, Toader C. Emergent electricity customer classification. *IEE Proc Gener Transm Distrib 2005*; 152(2):164-72.
[16] Chicco G, Napoli R, Piglione F, Scutariu M, Postolache P, Toader C. Application of clustering techniques to load pattern-based electricity customer classification. *Proc. 18th CIRED, Torino, Italy*, 6-9 June 2005, Session 5, paper No. 467.
[17] Gianfranco Chicco, Overview and performance assessment of the clustering methods for electrical load pattern grouping, *Energy* Vol. 42, 2012, 68 – 80.
[18] Menahem Friedman , Abraham Kandel, Introduction to Pattern Recognition: Statistical, Structural, Neural and Fuzzy Logic Approaches, *World Scientific*, ISBN-13: 978-9810233129, ISBN-10: 9810233124, 1999.
[19] N. H. M. Vu, N. T. P. Khanh, V. V.Cuong, P. T. T.Binh, Forecast on Vietnam Electricity Consumption to 2030. *Engineering, Technology & Applied Science Research*. Vol. 8, No. 3, 2018, 2869-2874.
[20] Nguyen Hoang Minh Vu, Vo Viet Cuong, Phan Thi Thanh Binh, Peak Load Forecasting for Vietnam National Power System to 2030. *Journal of Science and Technology*. No. 123, 2017.

Corresponding Author:
Nguyen Hoang Minh Vu
University of Technology and Education, HCMC
Email: vuminh.1974@gmail.com or vu.nguyenhoangminh@uah.edu.vn