Zone Clustering LMP with Location information using an Improved Fuzzy C-Mean

Se-Hwan Jang, Jin-Ho Kim, *Member, IEEE*, Sang-Hyuk Lee and June-Ho Park, *Member, IEEE*

***Abstract*-- For the efficient zone clustering of Large-scale power system, this paper introduces an improved fuzzy C-means (FCM) approach. Due to physical characteristics of Power System, each node of the power system has their own loctional marginal price (LMP) indicating the network-related characteristics of the system. In electricity market, it has been on the rise across much of a deregulation market to price market clearing prices considering LMP. It is inefficient to price on each node individually. Hence classification for the whole system into distinct several subsystems based on a dissimilarity measure is typically needed for the efficient operation of the whole system. Moreover, the location information on the system is taken into account in order to properly address the geometric mis-clustering problem such as grouping geometrically distant nodes with dissimilarity measures into a common cluster. Therefore, this paper proposes the improved FCM approach for clustering LMP with location information. We have conducted the clustering in real Korea power system to verify the usefulness of the proposed improved FCM approach.**

Index Terms—**FCM(fuzzy c-mean), Euclidean distance, LMP(locational marginal price), zone clustering, dissimilarity measure.**

I. Introduction

In the regional management of interconnected network systems, the efficient and economical operation of the networked systems in terms of system coherency is essential. Hence the research of system coherency has been made by numerous researchers [1-4]. However, most of the studies are focused on the dynamic grouping. At this point, we need a novel approach to partition the total system into several regions considering location information, such as locational marginal price(LMP), loss, regional distances, and so on. In this paper, It has been proposed to cluster the nodes of a networked power system considering LMP and the regional coherency.

Locational Marginal prices (LMPs) are used in many electric Markets to settle energy transactions. Generally, When a single slack bus is used in the power-flow formulation for calculating the prices, LMP can be decomposed into three components: 1) the reference price at the single slack bus (which is also the angle reference bus, 2) the marginal price of the transmission losses, and 3) the marginal price of the network constraints that are enforced in the power-flow model [5]. Namely, LMP implies the price at which the electricity is consumed at each node. Due to the physical characteristics of the electricity transmission network, electricity is lost when it is transmitted from supplying nodes (*i.e.*, supplying buses) to consuming nodes (consuming buses), and additional generation must be supplied to provide energy in excess of that consumed by customers. Moreover, the capacity limitation of the transmission network of electricity systems prevents full uses of system wide cheap electricity. Therefore, a electricity price at each node, *i.e.*, the price at which the electricity is consumed at each node is differently decided depending on the network topology and energy configuration.

Similarity measure has been known as the complementary meaning of the distance measure [6-10]. Hence, we consider the partitioning measure not only similarity measure but also location information, that is, distance measure. In the previous literatures, we had constructed similarity measure through distance measure or fuzzy entropy function [11]. Well known-Hamming distance was used to construct fuzzy entropy, so we composed the fuzzy entropy function through Hamming distance measure. With only similarity measure, we can obtain unpractical results, which partition physically distant locations into the same group. Hence we add the location information to complete combined dissimilarity measure.

In the next section, the axiomatic definition of similarity measure is introduced. Also, Modified similarity measure is proposed through distance measure and proved. In Section 3, the combined dissimilarity measures with the regional

This work has been supported in part by EERSI(R-2005-7-064), which is found by MOCIE (Ministy of Commerce, Industry and Energy)

S. H. Jang is with the Electrical Engineering Department, Pusan National University, Busan 609-735 Korea (Corresponding author, phone:+82-11-9544-3095;e-mail: elinia@naver.com).

S. H. Lee is with the Mecatronics Engineering Department, Changwon National University, Changwon 641-773 Korea (e-mail: leehyuk@changwon.ac.kr).

J. H. Kim is with the Electrical Engineering Department, Kyungwon University, Sungnam 461-701 Korea (e-mail: kimjh@kyungwon.ac.kr).

J. H. Park is with the Electrical Engineering Department, Pusan National University, Busan 609-735 Korea (e-mail:parkjh@pusan.ac.kr).

coherency are proposed. In Section 4, numerical examples are shown. In the example, we obtain a proper partitioning result, which consider both LMP and location information. Conclusions are followed in Section 5. Notations of Liu's are used in this paper [6].

II. Fuzzy C-means clustering and similarity measure

FCM algorithm was proposed by Bezdek in 1973 as an improvement over Hard C-means clustering (HCM) [11]. FCM play a roll of partitioning arbitrary vectors into fuzzy groups, also it finds a cluster center for each group such that a cost function of dissimilarity measure is minimized. Well known fact about FCM and HCM indicates that FCM employs fuzzy partitioning such that a data point can belong to several groups with the degree of membership grades between 0 and 1.

We will illustrate the FCM result briefly [12]. Membership U matrix is satisfied as follows:

$$\sum_{i=1}^{c} u_{ij} = 1,\ \forall j = 1,....,n \tag{1}$$

The cost function for FCM is defined below:

$$J(U, c_1, ..., c_0) = \sum_{i=1}^{c} J_i = \sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}^m d_{ij}^2 \tag{2}$$

where u_{ij} is between 0 and 1, C_i is the center of fuzzy group, d_{ij} = || c_i - x_j || is the Euclidean distance between i-th cluster center and the j-the data point, and m□[1,∞] is the weighting value. With Lagrange multiplier, the necessary conditions for (2) to reach a minimum are

$$\bar{J}(U, c_1, ..., c_c, \lambda_1, ..., \lambda_n) = J(U, c_1, ..., c_c) + \sum_{j=1}^{n} \lambda_j (\sum_{i=1}^{c} u_{ij} - 1)$$

$$= \sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}^m d_{ij}^2 + \sum_{j=1}^{n} \lambda_j (\sum_{i=1}^{c} u_{ij} - 1)$$

where λ_j , j=1 to n, are the Lagrange multipliers for the n constraints in (1). By differentiating respective input arguments, the necessary conditions for (2) to reach its minimum can be found as follow:

$$c_i = \frac{\sum_{j=1}^{n} u_{ij}^m x_j}{\sum_{j=1}^{n} u_{ij}^m} \tag{3}$$

And

$$u_{ij} = \frac{1}{\sum_{k=1}^{c} (\frac{d_{ij}}{d_{kj}})^{2/(m-1)}} \tag{4}$$

With these results, well known FCM algorithm is listed below:

Step 1: Initialize membership matrix U with random values satisfied by (1) between 0 and 1.

Step 2: Calculate c_i according to (3), i=1,…,c

Step 3: Compute (2). Stop if either it is below a certain tolerance.

Step 4: Compute a new U using (4). Go to step 2

Now for minimizing of (2), the less distance is d_{ij}, the smaller cost function become. Hence distance means the similarity between two data points. Finding similarity is determined from the types of data, time series signal, image, sound, *etc*.. Now we introduce a similarity measure for the fuzzy sets. And the proposed similarity measure can be applied to our problem.

We define modified similarity measure, which is different from that of Liu's.

Definition(Liu 1992). A real function : $P^2 \to R^+$ or $F^2 \to R^+$ is a modified similarity measure for regional point, if s has the following properties :

(S1) $s(A,B) = s(B,A),\ \forall A, B \in P(X)$ or $F(X)$

(S2) $s(A,A^C)$ satisfies minimum value, $\forall A \in P(X)$ or $F(X)$, where A^C is the farthest point from A.

(S3) $s(D,D)=max_{A,B \in P} s(A,B),\ \forall A, B \in P(X)$ or $F(X)$

(S4) $A, B, C \in P(X)$ or $F(X)$, or , if A□ B□ C, then $s(A,B) \ge s(A,C)$ and $s(B,C) \ge s(A,C)$.

With Definition, we propose the following theorem as the modified similarity measure.

Theorem 2.1 For any set $A, B \in P(X)$ or $F(X)$, if d satisfies Hamming distance measure, then

$$S(A,B) = 4 - 2d((A \cap B),[1]) - 2d((A \cup B),[0]) \tag{5}$$

is the similarity measure between set A and set B.

proof. We prove that the eq. (5) satisfies the similarity definition. (S1) means the commutativity of set A and B , hence it is clear from (3) itself. From (S2),

$$s(A, A^c) = 4 - 2d((A \cap A^c),[1]) - 2d((A \cup A^c),[0])$$

then $2d((A \cap A^c),[1])$ and $2d((A \cup A^c),[0])$ are the maximum values of between A and arbitrary set. For arbitrary sets A, B , inequality of (S3) is proved by

$$s(A,B) = 4 - 2d((A \cap B),[1]) - 2d((A \cup B),[0])$$
$$\le 4 - 2d((D \cap D),[1]) - 2d((D \cup D),[0])$$
$$= s(D,D).$$

Inequality is satisfied from $d((A \cap B),[1]) \ge d((D \cap D),[1])$ and $d((A \cup B),[0]) \ge d((D \cup D),[0])$.

Finally, (S4) is $\forall A,B,C \in F(X), A \subset B \subset C$,

$$\begin{aligned} s(A.B) &= 4-2d((A\cap B),[1])-2d((A\cup B),[0]) \\ &= 4-2d(A,[1])-2d(B,[0]) \\ &\geq 4-2d(A,[1])-2d(C,[0]) \\ &= s(A,C), \end{aligned}$$

Similaritily, $s(B,C) \geq s(A,C)$ is obtained through

$d(B,[0]) \leq d(C,[0])$ and $d(B,[1]) \leq d(A,[1])$.

Therefore, we know that the proposed similarity measure (5) is satisfied by the similarity definition.

III. New Dissimilarity with Location Information

In the previous section we have derived the modified similarity measures which satisfying the definition of similarity. To the large scale power system, though that has similar measure values on locational marginal price(LMP) at the node, they can be located far away. Then it is not realistic to gather that even though they have similar valued measure. So we need to consider LMP as well as another characteristic values, that is location information. We can group for the point that is having similar characteristic values through FCM. Original FCM only consider a distance between a cluster and a center. So we make use of proposed similarity measure to take LMP into account. To apply to FCM with dissimilarity, $d_{ij} = || c_i - x_j ||$, it is required that similarity has to change dissimilarity. Hence we consider the reciprocal relation that is frequently used. The dissimilarity of the proposed improved similarity measure is represented as below:

$$S_1(A,B) = \frac{1}{4-2d((A\cap B),[1])-2d((A\cup B),[0])}$$

Consequently, we consider the combined dissimilarity measure as

$$s(A,B) = \omega_1 s_1(A,B) + \omega_2 s_2(A,B) \qquad (6)$$

where $S2(A,B)$ is the geometrical distance value, and ω_1, ω_2 are the weighting values.

IV. Numerical Example

In this paper, we are to conduct the simulation in real Korea power system by using proposed improved FCM approach. If the nodes have almost similar or equivalent LMP and adjacent location information, we use representative value. Consequently, this simulation is conducted by using information of representative 128 nodes of nearly 700 nodes. Each node has information of location and LMP. In following Fig 1, it shows real location of each node. And it shows the name of the each node, LMP, location information of 2 dimension planes in Table 1.

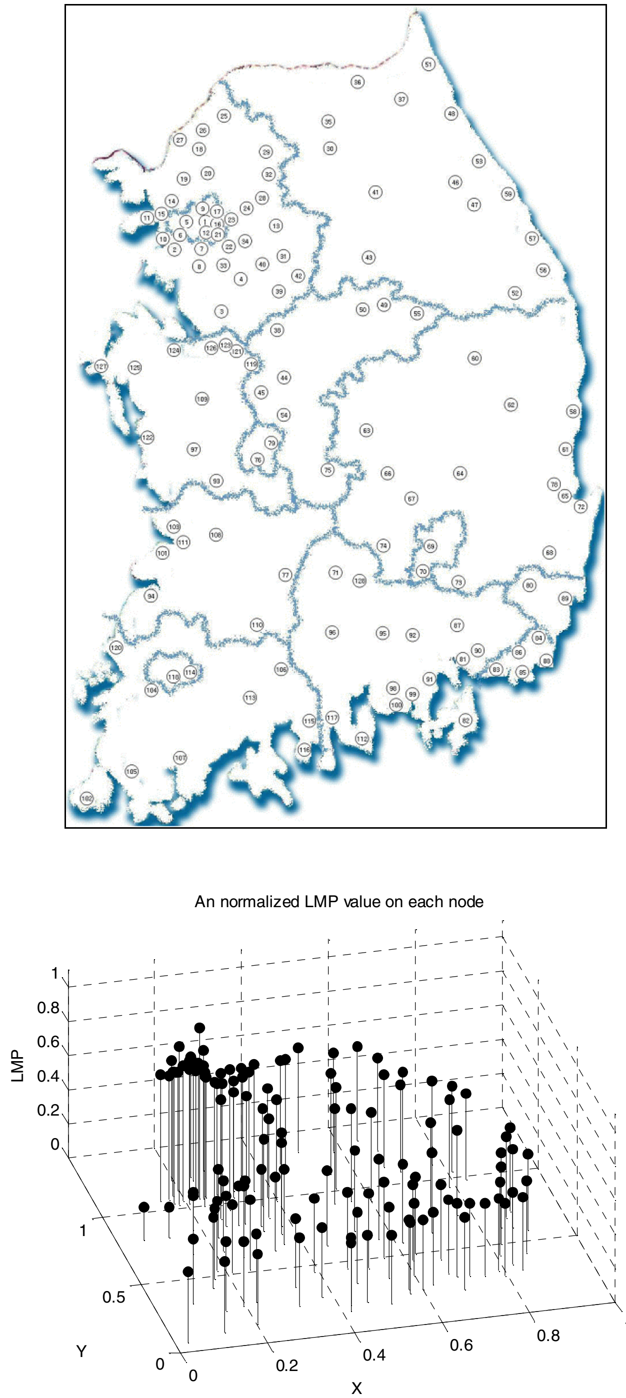

Fig. 1. Real location and an normalized LMP values at each node

Table I Locational marginal prices and per unit locations at each node

Num	Name	LMP	Position	
			X	Y
1	Hwasung S	78.31	7.8	35.2

2	Ansan	73.05	6	34.2
3	Pengtak TP3	66.73	8.5	30
4	D Osan	64.3	9.5	31.8
5	West Seoul M2	63.5	7	35.5
6	D Anyaung	62.18	6.5	35
7	Uiwang	62.09	6.8	34.2
8	West Suwon	61.95	7.8	33.2
9	YeongdeungpoTR	61.49	7.8	35.7
10	Sinsiheung M2	60.36	5.8	34.8
⋮				
119	Jochiwon	26.35	11	27
120	Yeonggang #6	29.49	3	11
121	Chunan	19.89	10	28
122	Borung ST#4	28.21	5	22.5
123	Onyang	17.94	9.5	28
124	Dangjin	15.95	6	28
125	Sinseosan TR	14	4	27
126	Asan M2	13.81	8.5	28
127	Teaan #6	15.26	2	27
128	Habchun #1G	34.37	17	14.5

Each node has a particular LMP due to the physical characteristics of the electricity transmission network. Namely, LMP include the cost of network congestion and transmission loss as well as a production cost. Electricity price (LMP) is distributed from 13.81 to 78.31 as above table. The location information of 128 nodes are represented through 2-dimensional plane at which plane is assumed to be flat.

It should be simulated by following the proposed dissimilarity measure on real Korea power system.

$$s(A,B)=\omega_1 s_1(A,B)+\omega_2 s_2(A,B)$$

At first, we partitioned the 128 nodes to the 2 cluster. And the result is illustrated in Fig. 2. 128 nodes are located in 2 dimensional spaces. It is represented as the location information in x-y plane. Also, nodes have a particular LMP. We can obtain the result of with only location information. However realistically, there is little point in only the consideration of location information due to having a special LMP on each node. Hence we will consider the location information and LMP simultaneously to satisfy on the user's request.

In Fig 2 and 3, it shows the clustering result of nodes through weighting values of LMP and location information. w_1 and w_2 are 0.3 and 0.7 in Fig 2 respectively. The weighting value of LMP is w_1 and that of location information is w_2. Namely, it focuses on location information. In contrast Fig. 3 put LMP more than location information. It is considered by weighting values w_1 and w_2 as 0.7 and 0.3 respectively. If we take result into account to the actual, there is realistic that the total area is divided into the southern area locating a lot of power plants and the northwestern area location a lot of loads. And the more we consider about weighting value of LMP, the more it tend to be divided northwestern area and the other.

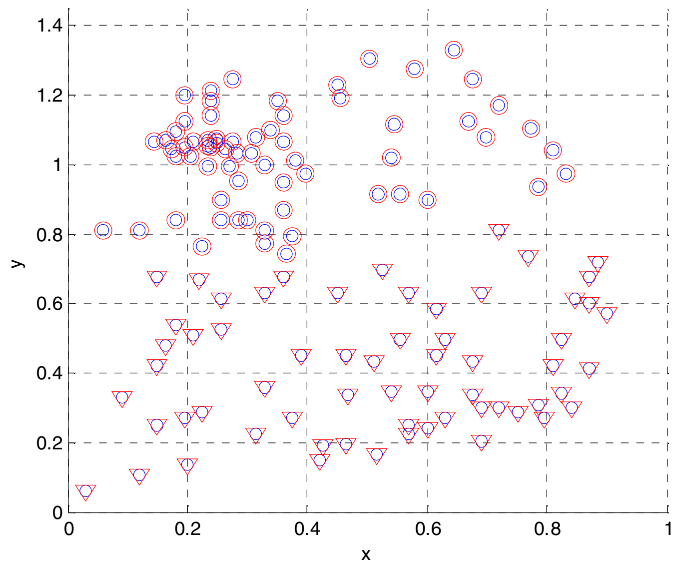

Fig. 2. Clustering by an improved FCM(w_1=0.3,w_2=0.7:location is focused)

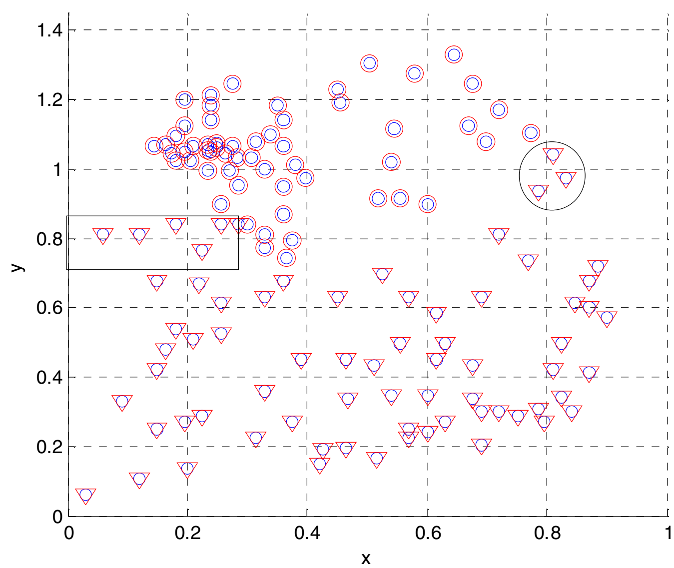

Fig. 3. Clustering by an improved FCM(w_1=0.7,w_2=0.3 :price is focused)

Next, we check on the result of clustering by a improved FCM when the number of cluster changes 2 to 4. We equally partition the 128 nodes to the 4 cluster and the result is illustrated. Also, each node is located in x-y plane. The result should be found by using the former simulation process that change weighting value.

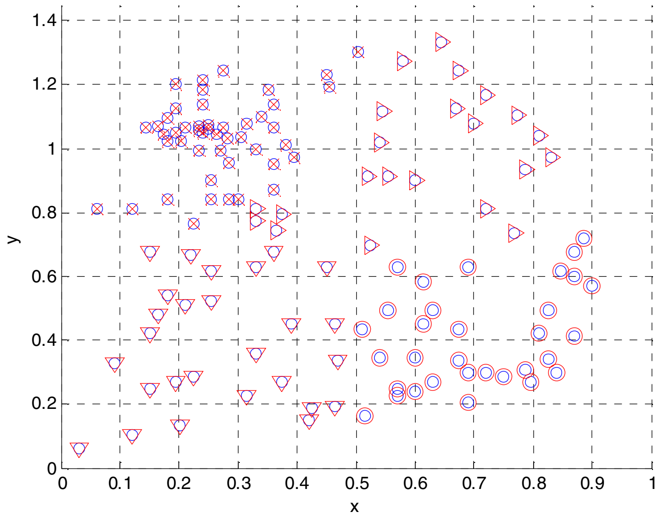

Fig. 4. Clustering by an improved FCM(w_1=0.2,w_2=0.8:location is focused)

In Fig. 4, it shows the result dividing nodes into 4 cluster by using the proposed improved FCM approach. Weighting value w_1 and w_2 are 0.2 and 0.8 respectively. In other words, it focuses on location information. And Fig. 5 is the result of the same weighting value. In contrast Fig. 6 considers LMP before location information. It is considered by weighting values w_1 and w_2 as 0.65 and 0.35 respectively.

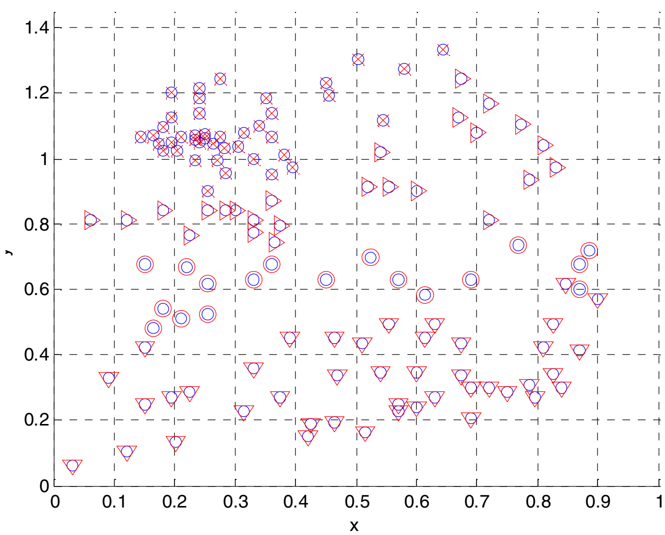

Fig. 5. Clustering by an improved FCM(w_1=0.5,w_2=0.5)

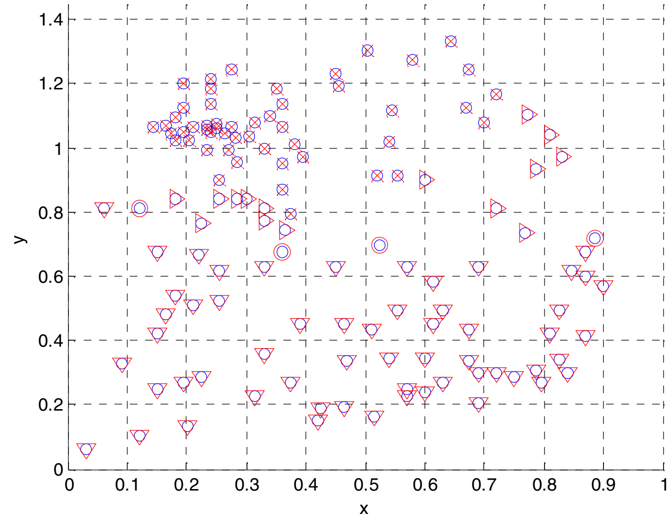

Fig. 6. Clustering by an improved FCM(w_1=0.65,w_2=0.35 :price is focused)

Results above represent that a partition on 4 clusters is more changeful than a partition on 2 clusters. And if the weighting value on LMP increases, that is similar with a result of 2 cluster focusing into LMP. Consequently when considering LMP more than location information, it is reasonable to assign a singularity area instead of increasing the number of cluster.

V. Conclusion

In this paper, we have introduced a similarity measure and constructed the modified dissimilarity measure using the proposed similarity measure. Also, we have the proposed improved FCM approach for clustering LMP with location information. Accordingly, we have conducted the clustering in real Korea power system to verify the usefulness of the proposed improved FCM approach. Only if the nodes have almost similar or equivalent LMP and adjacent location information, representative value is chosen. Hence we have used representative 128 nodes of approximately 700 nodes. We analyzed the result clustering as 2 cluster and 4 cluster on nodes having LMP and location information. From the results, we can check the coherency between the degree of dissimilarity level and the number of clusters.

VI. References

[1] W. Li and A. Bose, "A coherency based rescheduling method for dynamic security", IEEE Transactions on Power Systems, Vol. 13, No. 3, 810-815, 1998.

[2] S.K. Joo, C.C. Liu, L.E. Jones, and J.W. choe, "Coherency and aggregation techniques incorporating rotor and voltage dynamics", IEEE Transactions on Power Systems, Vol. 19, No. 2, 1068-1075, 2004.

[3] A.M. Gallai and R.J. Thomas, "Coherency Identification for large electric power systems", IEEE Transactions on Circuits and Systems, Vol. CAS-29, No. 11, 777-782, 1982.
[4] F.F. Wu. N. Narasmithamurthi, "Coherency Identification for power system dynamic equivalents", IEEE Transactions on Circuits and Systems, Vol. CAS-30, No. 3, 140-147, 1983.
[5] Locational marginal prices, [Online].Available:http://www.iso-ne.com
[6] Liu Xuecheng, "Entropy, distance measure and similarity measure of fuzzy sets and their relations," Fuzzy Sets and Systems, 52, 305-318, 1992.
[7] J. L. Fan, W. X. Xie, "Distance measure and induced fuzzy entropy," Fuzzy Set and Systems, 104, 305-314, 1999.
[8] J. L. Fan, Y. L. Ma, and W. X. Xie, "On some properties of distance measures," Fuzzy Set and Systems, 117, 355-361, 2001.
[9] S.H. Lee, S.P. Cheon, and Jinho Kim, "Measure of certainty with fuzzy entropy function", LNAI, Vol. 4114, 134-139, 2006.
[10] S.H. Lee, J.M. Kim, and Y.K. Choi, "Similarity measure construction using fuzzy entropy and distance measure", LNAI Vol.4114, 952-958, 2006.
[11] J.C. Bezdek, *Fuzzy Mathematics in Pattern Classification*, Ph.D Thesis, Applied Math. Center, Cornell University, Ithaca, 1973.
[12] J.S.R. Jang, C.T. Sun, and E. Mizutani, *Neuro-Fuzzy and Soft Computing*, Prentice Hall, 1997.

VII. BIOGRAPHIES

Se-Hwan Jang He received his B.S degrees from Pusan National University of Korea in 2006. And he is working on M.S degree at Pusan National University. His research interest includes optimal operation and planning of power system resources, deregulation and restructuring of power industry, electricity market analysis, and intelligent systems applications to Power systems.

Sang-Hyuk Lee He received his Ph. D degree from Seoul National University in 1998 and he is at Changwon National University as an assistant professor. His area of interest are robust control, game theory and fuzzy theory

Jin-Ho Kim He received his Ph. D degree from Seoul National University in 2001 and he is at Pusan National University as an assistant professor. His research interest includes optimal operation and planning of power system resources, deregulation and restructuring of power industry, electricity market analysis, and distribution and renewable energy sources.

June Ho Kim He received his B.S., M.S. and Ph. D. degrees from Seoul National University of Korea in 1978, 1980, 1987 respectively, all in Electrical Engineering. He is currently a Professor at Pusan National University. His areas of interest are intelligent systems applications to Power systems. Dr. Park has been a member of the IEEE Power Engineering Society, Control System Society, and System, Man and Cybernetics Society.