Practical Implementation of Congestion Cluster Pricing Method

Yong Yoon† Marija Ilić† (*IEEE Fellow*) Ken Collison† José Arce††
† Energy Laboratory
Massachusetts Institute of Technology
Cambridge, MA 02139
†† Instituto de Energiá Eléctrica
Universidad Nacional de San Juan
(5400) San Juan, Argentina

Abstract—
The paper describes a practical approach to implementing the congestion cluster pricing method as a viable congestion management system (CMS) in the operation of electric power systems.

The congestion cluster pricing method is formulated as a stochastic optimization problem of the cluster design. In the formulation the performance of the pricing method is introduced as a measurable function of cluster design, based on the conceptual criteria necessary for an effective CMS. We define the search space from which a particular design may be selected. Following the formulation the stochastic elements to the optimization problem are discussed by developing suitable representation of various uncertainties in the system. It is shown that the complexity of the problem leads to the search based methods as the preferred option for solving for the optimal cluster design.

Finally, some reasonable approximations are suggested to solving the problem thus making it a *practical* approach to implementing the congestion cluster pricing method. A numerical example is given to illustrate the proposition.

I. Introduction

Competition and market mechanisms have been introduced into the electric power industry to maximize system efficiency. Here the well designed market structure replaces the strict regulation regime of the vertically integrated industry. In these markets trades in the spot market are frequently linked to short-term efficiency, while bilateral transactions are linked to long-term efficiency .[2]

The congestion management system (CMS)[1], plays a significant role in operating the energy market. At the time of writing, there are two schools of thoughts in implementing a market-based CMS. The first is bus-based CMS, an example of which is nodal pricing.[7] The second is cluster-based CMS, in which nodes belonging to the same cluster receive a single cluster-wide price. An example of this is cluster-based CMS.[8]

The cluster-based CMS is much more accommodating to implementing bilateral transactions by providing transparent information on the status of transmission (system) congestion. The uniform prices within clusters are another advantage of the cluster-based CMS. The disadvantages are related to the unfavorable increase in cost of dispatched generators in short term due to: (1) the cost from the cluster-wide prices in inter-cluster pricing and (2) the cost from the uplift charges in intra-cluster pricing.

The congestion cluster pricing method is quite suitable as a viable CMS as it reduces the the effect of disadvantages while preserving the effect of advantages of the implementation of the cluster-based CMS.[9] The key to the method is the novel approach proposed in [10] used to compute the sensitivity measures of injection.

[1]Congestion management system is the process of choosing which generators to dispatch in the presence of congestion.[4]

The implementation of the congestion cluster pricing method consists of two steps: (1) aggregation of individual nodes into clusters and (2) computation of cluster-wide prices. The number of clusters and the duration of fixed cluster boundaries are required to be specified ahead of time with respect to some heuristic measure of long term efficiency according to the need of the market and its participants. Details of the congestion cluster pricing method can be found in [8] and [9].

The minimum desired criteria for the congestion cluster pricing method can be summarized as

1. the transaction between any buses within the same cluster have little impact of power flows on the congested transmission lines
2. the energy cost computed after relieving inter-cluster congestion is relatively small
3. the additional energy cost necessary for relieving intra-cluster congestion is relatively small

The paper is organized as follows:
Section II shows the problem of the cluster design formulated as a stochastic optimization problem. The search based methods are introduced as the preferred option for solving for the optimal cluster design given the high complexity of the optimization problem in Section III. Section IV presents the numerical examples to illustrate the proposition, and Section V summarizes the conclusions of the paper.

II. Formulation of Cluster Design Problem

Throughout the paper the formulation of the problems is performed under the following two assumptions.

1. DC power flow
2. Quadratic generation cost This implies that under the perfectly competitive market condition, the optimal production decision for the given price is to generate based on the marginal cost given by,

$$\begin{aligned} MC_{G_i} &= \frac{dC_{G_i}}{dQ_{G_i}} \\ &= 2a_{G_i}Q_{G_i} \end{aligned} \quad (1)$$

First, we present the formulation of the aggregation step in the implementation of the congestion cluster pricing method as a stochastic optimization problem given by

$$\theta^{\star} = \arg\min_{\theta\in\Theta} \int_0^T J(\theta,t)dt \equiv \mathcal{E}\left[\int_0^T L(\theta,\xi(t),t)\right] \quad (2)$$

where Θ represents the search space from which various cluster design alternatives can be selected for aggregating individual nodes into clusters. In Eq. (2) the performance measure denoted by $J(\theta, t)$ is the expected value of the sample performance, $L(\theta, \xi(t), t)$ which is a function of the cluster design, θ, and the uncertainty, $\xi(t)$ in the system. Given that it is desired to keep the same cluster boundaries for a certain period of time, i.e. a season, T represents the duration of fixed boundaries. A slightly modified form of Eq. (2) is a little more useful as the minimum time scale at which the operation of the system takes place is typically one hour. Thus, the optimization problem of the interest is given by

$$\theta^{\star} = \arg\min_{\theta \in \Theta} \sum_{k=0}^{T} J(\theta, k) \equiv \mathcal{E}\left[\sum_{k=0}^{T} L(\theta, \xi[k], k)\right] \quad (3)$$

A. Modeling Uncertainties in the System

In Eq. (3) the uncertainty, $\xi(t)$ in the system can actually be broken into three parts: the uncertainty in load $\xi_{Q_{D_i}}(t)$, the uncertainty in generation bid, $\xi_{C_{G_i}}(t)$, and the uncertainty in status of equipment, i.e. generator or transmission line, $\xi_{Q_{G_i}}$ or ξ_{F_l} respectively.

We use a time series representation of each uncertainty.

1. Modeling load uncertainty

   The time series model of load can be represented in a general version of a discrete time random walk by

$$Q_{D_i}[k+1] = f_{D_i}\left(\bar{Q}_{D_i}[k+1], Q_{D_i}[k]\right) + e_{Q_{D_i}}[k+1] \quad (4)$$

   where $e_{Q_{D_i}}[k]$ is normally distributed with zero mean and variance $\sigma^2_{Q_{D_i}}$ and is independent of $e_{Q_{D_i}}[l]$ for any $k \neq l$. The expected value of the demand at k is denoted with $\bar{Q}_{D_i}[k]$ while the projected demand at k computed through Eq. (4) is indicated with $Q_{D_i}[k]$. Typically $f(\cdot)$ is assumed to take on either linear or exponential form, and the parameter estimation is performed to complete this regression model.[1]

2. Modeling generation bid uncertainty

   The time series model of generation bid can be represented in a similar way by

$$MC_{G_i}[k+1] = 2a_{G_i}[k+1]Q_{G_i} + b_{G_i}[k+1] \quad (5)$$

   where typically the slope, a_{G_i} is assumed to be fixed, i.e. $a_{G_i}[k] = a_{G_i}$ and the intercept follows another linear regression model given by

$$b[k+1] = b[k] + e_{C_{G_i}}[k] \quad (6)$$

   where $e_{Q_{C_{G_i}}}[k]$ is again normally distributed with zero mean and variance $\sigma^2_{Q_{C_{G_i}}}$.[1]

3. Modeling equipment status uncertainty

   The time series model of equipment status can be represented using a conventional Markovian chain consisting two states as shown in Figure 1. The parameters for the failure rate and the repair rates are denoted by λ and μ respectively for each component in the figure.

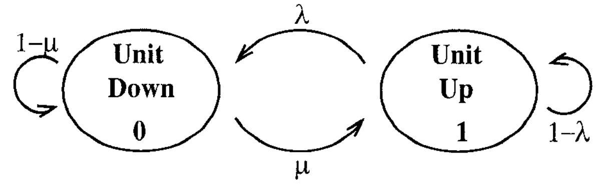

Fig. 1. Markovian Chain Modeling of Equipment Status

Using these parameters, the governing equations for transition probability for states 0 and 1 are given by

$$\pi_0[k] = -\frac{\lambda}{\lambda+\mu}\left[1-(\lambda+\mu)\right]^k + \frac{\lambda}{\lambda+\mu} \quad (7)$$

$$\pi_1[k] = \frac{\lambda}{\lambda+\mu}\left[1-(\lambda+\mu)\right]^k + \frac{\mu}{\lambda+\mu} \quad (8)$$

respectively given that the component is initially in the "up" state.[6]

B. Function for Sample Performance

In Eq. (3) the function describing sample performance, $L(\theta, \xi[t], k)$ determines how the superior designs are compared to the inferior ones; i.e. if $\mathcal{E}\left[L(\theta_i, \xi[k], k)\right] < \mathcal{E}\left[L(\theta_j, \xi[k], k)\right]$, then θ_i is a better cluster design than θ_j. Thus, the function is directly related to the various criteria for a good congestion cluster pricing method. The minimum desired criteria for the method are already discussed in the previous section and are listed here again for completeness:

1. the transaction between any buses within the same cluster have little impact of power flows on the congested transmission lines, $L_{D(i,j)}(\theta, \xi[k], k)$
2. the energy cost computed after relieving inter-cluster congestion is small, $L_{Q_{G_i}}(\theta, \xi[k], k)$
3. the additional energy cost necessary for relieving intra-cluster congestion is small, $L_{\Delta Q_{G_i}}(\theta, \xi[k], k)$

Limiting the sample performance to reflect only the measures of the above three criteria, we consider the overall sample performance function to be given as

$$L(\cdot) = \alpha_{D(i,j)} L_{D(i,j)}(\cdot) + \alpha_{Q_{G_i}} L_{Q_{G_i}}(\cdot) + \alpha_{\Delta Q_{G_i}} L_{\Delta Q_{G_i}}(\cdot) \quad (9)$$

where α's denote the relative importance factors of each criterion. Typically, the factors are selected such that $\alpha_{D(i,j)} L_{D(i,j)}(\cdot) \geq \alpha_{\Delta Q_{G_i}} L_{\Delta Q_{G_i}}(\cdot) \geq \alpha_{Q_{G_i}} L_{Q_{G_i}}(\cdot)$.

The congestion distribution factors (CDFs) proposed in [10] give good measure of the impact of transactions between buses to the congested lines. CDFs are novel distribution factors computed in such a way as to be independent of the slack bus. See [10] for details of the derivation.

The energy cost after relieving inter-cluster congestion is closely related to the computation of cluster-wide prices step in the implementation of the congestion cluster pricing method. As a matter of fact, the equations used for computing the energy cost and the cluster-wide prices are the same. Suppose the nodes $G_i, G_{i+1}, \cdots, G_{i+k}$ are in the cluster z_j. Then, at some t the new generation cost associated with the cluster z_j is given by

$$C_{z_j}(Q_{z_j}) = f_{z_j}(Q_{G_i}, Q_{G_{i+1}}, \cdots, Q_{G_{i+k}}) \quad (10)$$

where f_{z_j} is the monotonically increasing nonlinear function representing the least cost combination of Q_{G_i}'s in z_j for producing Q_{z_j}. The marginal cost of zone z_j, MC_{z_j}, can be used in order to compute $f_{z_j}(\cdot)$ where

$$MC_{z_j} = \begin{cases} \left(\frac{1}{2a_l} + \frac{1}{2a_{l+1}} + \cdots + \frac{1}{2a_{l+s}}\right)^{-1} Q_{z_j} & Q_{z_j} \in R_{I_1} \\ \left(\frac{1}{2a_m} + \frac{1}{2a_{m+1}} + \cdots + \frac{1}{2a_{m+t}}\right)^{-1} Q_{z_j} & Q_{z_j} \in R_{I_2} \\ \vdots & \\ \left(\frac{1}{2a_n} + \frac{1}{2a_{n+1}} + \cdots + \frac{1}{2a_{n+u}}\right)^{-1} Q_{z_j} & Q_{z_j} \in R_{I_k} \end{cases} \quad (11)$$

where R_{I_i}'s define the region of operating condition in cluster j with q number of generators are still below the generation limits. a_r's represent the coefficient of associated marginal cost of those generators below their generation limits.

With $C_{z_j}(Q_{z_j})$, the generation costs (and/or cluster-wide prices) are computed by solving the optimization problem given as

$$Q^{\star}_{z_j} = \arg\min_{Q_{z_j}} \sum_{z_j} C_{z_j}(Q_{z_j}) \quad (12)$$

subject to the load flow constraint, i.e., total generation is equal to system load,

$$\sum_{z_j} Q_{z_j} = \sum_{D_i} Q_{D_i} \qquad : \lambda \quad (13)$$

the congestion interface flow limit constraints, i.e., the power flow on any line l *along only the congestion interfaces* is within the maximum rating of the line,

$$|F_l| = \left|\sum_{z_i} H_{lz_i} Q_{z_i} - \sum_{D_i} H_{lD_i} Q_{D_i}\right| \le F_l^{max} \quad : \mu_l \quad (14)$$

and the generation limit constraints, i.e., the dispatch amount in cluster z_j is within the sum of maximum rating of the corresponding generators within the cluster

$$0 \le Q_{z_j} \le \sum_{G_i \in z_j} Q_{G_i}^{max} \qquad : \eta_{z_j} \quad (15)$$

The computation of H_{lz_i} yields

$$H_{lz_i} = \frac{dF_l}{dQ_{G_i}} \frac{\partial Q_{G_i}}{\partial Q_{z_j}} + \frac{dF_l}{dQ_{G_{i+1}}} \frac{\partial Q_{G_{i+1}}}{\partial Q_{z_j}} + \cdots + \frac{dF_l}{dQ_{G_{i+k}}} \frac{\partial Q_{G_{i+k}}}{\partial Q_{z_j}} \quad (16)$$

with

$$\frac{dF_l}{dQ_{G_i}} = H_{lG_i} \quad (17)$$

and with

$$Q_{G_i} = \frac{1}{2a_i} \left(\frac{1}{2a_i} + \frac{1}{2a_{i+1}} + \cdots + \frac{1}{2a_{i+k}}\right)^{-1} Q_{z_j} \quad (18)$$

if $Q_{G_i} \in R_{I_i}$.

The solution to the optimization problem (12) then given by

$$\rho_{z_i} = \lambda + \sum_l \mu_l H_{lz_i} \quad (19)$$

where $\mu_l \neq 0$ if and only if $|F_l| = F_l^{max}$ and

$$Q_{G_i} = \begin{cases} Q_{G_i}^{max} & \rho_{z_i, G_i \in z_i} \ge p_{G_i}^{max} \\ \frac{\rho_{z_i}}{2a_{G_i}} & 0 \le \rho_{z_i, G_i \in z_i} \le p_{G_i}^{max} \\ 0 & \text{otherwise} \end{cases} \quad (20)$$

where $p_{G_i}^{max} = 2a_{G_i} Q_{G_i}^{max}$. The total energy cost after relieving inter-cluster congestion is then given by

$$TC_{Q_{G_i}} = \sum_{z_i} \rho_{z_i} Q_{z_i} \quad (21)$$

The computation of the energy cost after relieving intra-cluster congestion is similar to that after inter-cluster congestion. The optimization problem to be solved in order to determine the location marginal prices is given by

$$\Delta Q_{G_i} = \arg \min_{\Delta Q_{G_i}, G_i \in \mathcal{Z}} \sum_{G_i} C_{G_i}(\Delta Q_{G_i}) \quad (22)$$

where

ΔQ_{G_i} : the adjusted generation amount at node G_i

$\mathcal{Z}$: the subset of clusters experiencing intra-cluster congestion

subject to the load flow constraint

$$\sum_{G_i \in \mathcal{Z}} \Delta Q_{G_i} = 0 \quad (23)$$

the transmission line flow limit constraints, i.e., the power flow on any line l in the entire system is within the maximum rating of the line,

$$|F_l + \Delta F_l| = |H_{lG_i}(Q_{G_i} + \Delta Q_{G_i}) + H_{lD_i} Q_{D_i}| \le F_l^{max} \quad (24)$$

and the generation limit constraints, i.e., the dispatch amount at node $G_i \in \mathcal{Z}$ is within the maximum rating of the corresponding generator

$$0 \le Q_{G_i} + \Delta Q_{G_i} \le Q_{Gi}^{max} \quad (25)$$

The additional energy cost necessary for relieving intra-cluster congestion is then given by

$$TC_{\Delta Q_{G_i}} = \sum_{G_i \in \mathcal{Z}} a_{G_i} \Delta Q_{G_i} (Q_{G_i} + \Delta Q_{G_i}) \quad (26)$$

III. Practical Solution to the cluster Design Problem

After the formulation we discover quickly that the so-called *real value based methods* are unlikely to yield a good result for solving this particular optimization problem. The real variable based methods refer to the analytical approaches to finding the optimal solution which require a sequential improvement by examining the gradients of smooth trajectories in the system with respect to search space. The reason for the difficulty in applying the real variable based methods to the problem lies in the lack of the nice structure of the search space, Θ, such as continuity, differentiability, etc., which are essential for finding smooth trajectories and computing gradients. This leads one to believe that the *search based methods* are more suitable for the optimization problem in Eq. (3). The search based methods refer to the simulation supported approaches to finding the optimum which requires a ranking of all possible design after a thorough evaluation of performance of each design alternative.

In order to apply the search based methods for the problem in Eq. (3) we first examine the search space, Θ. Suppose that the system is composed of N_{TR} transmission lines and N_B buses; N_G generators and N_D loads, and that the maximum number of clusters allowed is limited to N_z. Since once the maximum number of clusters are fixed, it is always possible to devise a cluster design to perform better than or at least equal to any existing design by allowing one more cluster in terms of $L(\theta, \xi[k], k)$ defined earlier [9], we start with the search space of size, $|\Theta|$ given by

$$|\Theta| = N_z^{(N_B - N_z)} \tag{27}$$

A typical electric power system consists of hundreds to thousands of buses, so conservatively let $N_B = 100$. Even if the number of clusters allowed is less than 10, assume $N_z = 5$, the number of designs to be considered in the search based methods is given by

$$|\Theta| = 5^{(100-5)} \approx 2.5 \times 10^{66} \tag{28}$$

which is typical by combinatorial standards.

Even though a further reduction in the size of Θ may be possible depending on the topology of system, it is clear from examining the size of the search space that a brute force application of the search based methods is not likely to be a good approach for any reasonable simulation time. Therefore, it is necessary to exploit any structural characteristics of the search space linked to the sample performance function.

One such characteristic is the first cut cluster design based on CDFs. Even though no analytical justification on the effective measures is available, there are a few empirical results which suggest that the size of the search space can be reduced significantly by designing clusters based on CDFs with little concerns for carelessly excluding good designs from the remaining search space.[8] This is especially true if the sample performance function,

$$L(\cdot) = \alpha_{D(i,j)} L_{D(i,j)}(\cdot) + \alpha_{Q_{G_i}} L_{Q_{G_i}}(\cdot) + \alpha_{\Delta Q_{G_i}} L_{\Delta Q_{G_i}}(\cdot) \tag{29}$$

is such that [9]

$$\alpha_{D(i,j)} L_{D(i,j)}(\cdot) \gg \alpha_{\Delta Q_{G_i}} L_{\Delta Q_{G_i}}(\cdot) \gg \alpha_{Q_{G_i}} L_{Q_{G_i}}(\cdot) \tag{30}$$

A practical approach to the clustering design, thus starts with the system operator identifying the potentially critical lines, some of which may be congested at the same time or at different times. Typically, the number of critical lines, N_{TR}^c, is less than five, so again conservatively let $N_{TR}^c = 3$. For each of the three transmission lines, corresponding CDFs are computed. Then, based on the relative values of CDFs the system is divided into clusters as described in [10]. Since there are multiple critical lines, the clusters defined for each line must be superposed on top of each other, and the intersections of the clusters constitutes the first cut design. The empirical results show that for a system of $N_B = 100$, three critical lines result in around 20 clusters. Given that the desired number of clusters is five, the search space of the problem is reduced from 2.52×10^{66} to 3.05×10^{10}.

Although the size of Θ is reduced by the orders of magnitude, the problem is still not manageable from the optimization point of view. Suppose 10,000 samples are selected randomly from Θ and serve as the sample set, Θ' for applying the search based method. The probability of the optimum solution from Θ being contained in this sample space is given by

$$\text{Prob}(\theta^\star \in \Theta') = 1 - \left(1 - \frac{1}{3.05 \times 10^{10}}\right)^{10{,}000} = 3.28 \times 10^{-7} \tag{31}$$

which is less than unlikely.

Still the sample size must be further reduced to a manageable size before applying any search based method to Eq. 3. Fortunately many of 3.05×10^{10} are infeasible as geographically distant clusters after the first cut design cannot be combined to be included in the sample set. Some more topological characteristics allows further reduction of the size of the sample set. Even though a generalization of exploiting the topological characteristics of the system may be made based on the recent development in various graph partitioning methods, we employ more heuristic approach to reducing the sample set. For instance, there are some rules of thumb, such as not allowing clustering near the critical lines, that significantly limit the possible designs to be included in the sample set. We claim without an analytical proof that the heuristic approach by an experienced system operator allows for the sample set containing around 1,000 design from which at least 50 designs belong in the top 100 designs of the original search space for $N_B \approx 100$, $N_{TR} \approx 200$ and $N_z \approx 10$. Thus, by and large the complexity of finding the optimal solution to the problem in Eq. (3) is reduced from the search space of $|\Theta| \approx 2.5 \times 10^{66}$ to the sample space of $|\Theta'| \approx 1,000$.

A. Application of Ordinal Optimization Method

Here we examine the optimization problem in Eq. (3) from the perspective of the ordinal optimization (OO) methodhas been proven to be very effective in dealing with the search based methods [3].

The basic idea of the OO method is the softening of the objective of finding the optimum to finding any design belonging to the "good enough" subset. For example, the good enough subset can be defined as the top-n% of the design space. The softening of the objective allows for working only in the much reduced selected subset with the expectation for a reasonable number of designs belonging to the good enough set at a high confidence. If the performance of each design is measured without any noise, the original optimization problem is transformed into the problem of selecting the design with the smallest evaluated performance belonging to the selected subset.[3] When the performance estimate is noisy, it becomes necessary to include more than one design in order to secure with higher confidence a certain degree of matching, or alignment, between the selected subset and the good enough subset.[5]

Suppose that the size of the search space containing all possible designs is in the order of 10^{10} as in the case with our problem. By goal softening principle we limit our goals to picking any of the top 5% designs. Consider a set consisting of 1,000 random samples from the search space. Then, the probability of retaining at least one of the top 5% designs in this sample space is given by

$$\mathrm{Prob}\,(G \cap \Theta' \neq \emptyset) = 1 - (1 - 0.05)^{1{,}000} \approx 1 \qquad (32)$$

where G and Θ' denote the set of the top 5% designs and the sample space respectively.

Similar to the idea of taking an exit poll from a limited number of electoral votes, if the designs in the sample space are chosen completely random, then we may assume that Θ' of the size 1,000 will more or less include 50 designs that belong to G. We can thus reduce the problem from finding any of designs that belongs to the set of top 5% designs from the search space of size 10^{10} to finding any designs that belong to top 50 designs from the sample space of size 1,000. The reduction of complexity is, indeed, quite considerable.

Let G' denote the set consisting the top 50 design contained in the sample space, Θ'. Now consider the selected subset consisting s designs chosen randomly from Θ'. We are interested in necessary s such that the alignment probability defined as $\mathrm{Prob}\,(|G \cap S| \geq k) \geq \mathcal{P}_A$ where k and $\mathcal{P}_A$ are defined depending on the purpose. For example, let $k = 3$ and $\mathcal{P}_A = 90\%$. For the parameters given the equation for computing the alignment probability is given as [3]

$$\mathrm{Prob}\,(|G \cap S| \geq 3) = \sum_{i=3}^{50} \frac{\begin{bmatrix} 50 \\ i \end{bmatrix} \begin{bmatrix} 1{,}000 - 50 \\ |S| - i \end{bmatrix}}{\begin{bmatrix} 1{,}000 \\ |S| \end{bmatrix}} \geq 0.90 \qquad (33)$$

Using Eq. (33) we deduce that the selected subset requires to have at least 102 designs in order to have at least 3 of them belong to the top 50 designs of the sample space.

To summarize the application of the OO method allows for a considerable savings of computational time in obtaining an acceptable solution to optimization problem through search based methods while the method itself involves only the following simple steps [5]

1. selecting the sample set of size N, $|\Theta'| = N$
2. defining the goals: # of good designs, g, # of good design alignment in the selected subset, k and the probability of alignment, $\mathcal{P}_A$
3. determining the subset size, s and selection rules that meets the goals
4. constructing the selected subset, S
5. comparing the designs in the selected subset

Before describing the method for accurate comparison of designs, we point out that the goal stated at the beginning is not a very impressive one since given that the size of the search space is in the order of 10^{10}, the top 5% design include the designs that is as far as 5×10^8 away from the true optimum.

For the optimization problem at hand, however, the top 50 designs in the sample space consisting of 1,000 are much better representatives than the top 5% of the entire search space. As discussed earlier this is because the designs in the sample space are not picked randomly but through a rigorous testing of the performance based on the first criterion for good cluster design. It is stated earlier that if the clusters are defined based on the CDFs, and if the importance of each criterion is defined such that first criterion is weighed orders of magnitude higher than the other two, then the designs based on the CDFs are ranked much closer to the true optimum than the rest of the possible designs. It may not be possible to accurately quantify how much better are the top 50 designs in the sample space to the the top 5% of the entire search space. However, it would not be surprising to find that the sample space contains at least 50 of the top 100 cluster designs from the search space if the clusters are defined based on CDFs respect to the critical transmission lines identified by an experienced operator relying on many heuristic tools.

B. Fairly Accurate Comparison of Designs in the Selected Subset

The ranking of each design alternative requires evaluating the sample performance. According to three criteria for good cluster design, $L(\theta, \xi[k], k)$ is defined as a function consisting a linear combination of three parts, namely $L_{D(i,j)}(\cdot)$, $L_{Q_{G_i}}(\cdot)$, and $L_{\Delta Q_{G_i}}(\cdot)$ as shown in Eq. (9). Assume that the relative weights, α, are chosen so that

$$\alpha_{D(i,j)} L_{D(i,j)}(\cdot) \gg \alpha_{\Delta Q_{G_i}} L_{\Delta Q_{G_i}}(\cdot) \gg \alpha_{Q_{G_i}} L_{Q_{G_i}}(\cdot) \qquad (34)$$

Then we claim without proof that only $L_{Q_{G_i}}(\cdot)$ and $L_{\Delta Q_{G_i}}(\cdot)$ are relevant for evaluating the designs in the selected subset. The reason for this is because when the designs are chosen to be included in the selected subset, $L_{D(i,j)}(\cdot)$ is already used for comparison purposes. The designs in the selected set are assumed to have about the same $L_{\Delta Q_{G_i}}(\cdot)$ compared to the others in the same set for otherwise the selected subset can be further reduced due to Ineq. (34).

Consider the modified sample performance, $L'(\theta, \xi[k], k)$.

We write $\sum_{k=0}^{T} L'(\cdot)$ as

$$\sum_{k=0}^{T} L'(\theta,\xi[k],k) = \sum_{k=0}^{T}\left[\min_{Q_{z_j}[k]}\sum_{z_j} C_{z_j}(Q_{z_j}[k],k) + \min_{\Delta Q_{G_i}, G_i\in\mathcal{Z}[k]}\sum_{G_i} C_{G_i}(\Delta Q_{G_i}[k],k)\right] \quad (35)$$

subject to the load flow constraints at each hour k

$$\sum_{z_j} Q_{z_j}[k] = \sum_{D_i} Q_{D_i}[k] \quad (36)$$

$$\sum_{G_i\in\mathcal{Z}[k]} \Delta Q_{G_i}[k] = 0 \quad (37)$$

the transmission line flow limit constraints[2]

$$|F_{l'}[k]| = \left|\sum_{z_i} H_{l'z_i}[k]Q_{z_i}[k] - \sum_{D_i} H_{l'D_i}Q_{D_i}[k]\right| \le F_{l'}^{max}[k] \quad (38)$$

$$|F_l[k]+\Delta F_l[k]| = \left|H_{lG_i}[k]\left(Q_{G_i}[k]+\Delta Q_{G_i}[k]\right) + H_{lD_i}Q_{D_i}[k]\right| \le F_l^{max}[k] \quad (39)$$

and the generation limit constraints

$$0 \le Q_{z_j}[k] \le \sum_{G_i\in z_j} Q_{G_i}^{max}[k] \quad (40)$$

$$0 \le Q_{G_i}[k] + \Delta Q_{G_i}[k] \le Q_{G_i}^{max}[k] \quad (41)$$

Under the formulation presented above the uncertainty in the system is incorporated by considering

1. Load uncertainty
   substitute Eq. (4) into Q_{D_i} in Eqs. (36), (38) and (39)
2. Generation bid uncertainty
   substitute Eqs. (5) and (6) into $\frac{dC_{G_i}}{dQ_{G_i}}$ in Eq. (35)
3. Equipment status uncertainty
   substitute 0 for $F_l^{max}[k]$ (or $Q_{G_i}^{max}[k]$) if the transmission line l (or the generator G_i) is in the "down" state

With Eqs. (35) - (41) we can rewrite the cluster design problem as the stochastic optimization problem given by

$$\theta^\star = \arg\min_{\theta\in\Theta}\sum_{k=0}^{T} J'(\theta,k) \equiv \mathcal{E}\left[\sum_{k=0}^{T} L'(\theta,\xi[k],k)\right] \quad (42)$$

The expectation in Eq. (42) can be evaluated using the search based method (the Monte Carlo method) by

$$\mathcal{E}\left[\sum_{k=0}^{T} L'(\theta,\xi[k],k)\right] = \frac{1}{N_{\text{iter}}}\sum_{i=1}^{N_{\text{iter}}}\sum_{k=0}^{T} L'(\theta,\xi_i[k],k) \quad (43)$$

where ξ_i represents the ith sample of the uncertainty.[3]

It is recognized that because of the uncertainty is modeled using either a general version of a discrete time random walk or the transient Markovian chain, the number of probabilistic states that need to be evaluated grow exponentially with time k in order to compute Eq. (35). This is quite limiting in applying the search based method. Therefore, some modifications are necessary in order to simplify the optimization to be manageable. One such modification is to work with the steady state probability rather than the transient probability.

[2] $lprime$ denotes the lines only on the congestion cluster interfaces.

B.1 Steady State Approximation of Uncertainty

For representing the uncertainty in load and the uncertainty in generation bid through steady state probability, the models described in [11] is useful. First, for modeling the load the identified are the several basic load patterns: typically peak load pattern, normal load pattern and off-peak pattern as shown in Figure 2, and the range of system

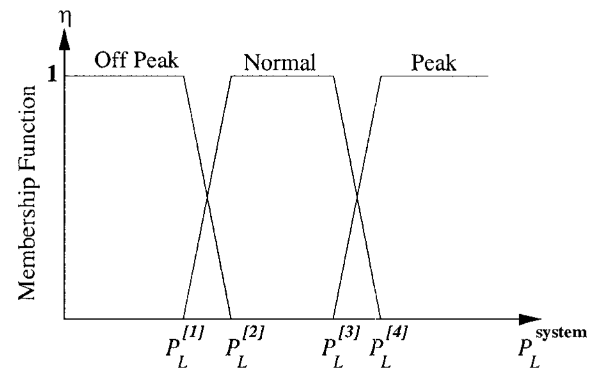

Fig. 2. Membership Functions for Individual Load Pattern

load levels given in discretized steps of hMW starting from Q_D^0 MW, i.e. $Q_D^{tot}(k) = Q_D^0 = kh$ as shown in Figure 3. Then, in the model if the total system load is larger than

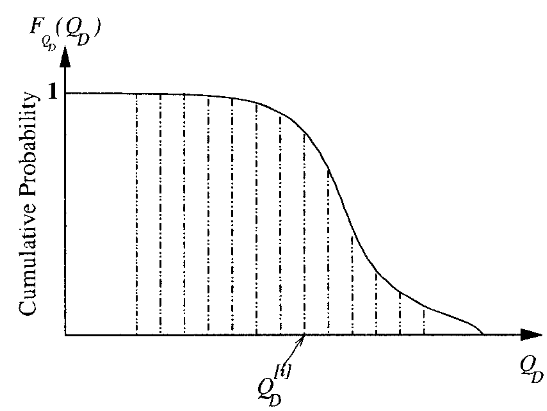

Fig. 3. Discretized Load Duration Curve for the Range of System Load

$Q_{[4]}$, the load distribution follows that of the peak ; if system load falls between $Q_{[2]}$ and $Q_{[3]}$, it follows the normal load distribution; and if system load is less than $Q_{[1]}$, off-peak load pattern is used to depict the load distribution. If the system load is either between $Q_{[1]}$ and $Q_{[2]}$ or $Q_{[3]}$ and $Q_{[4]}$ the appropriate patterns are meshed to create typical individual load pattern. This process can be, employing fuzzy logic, summarized as

$$Q_D^{[k]} = \left(\eta^{[N]} \frac{Q_D^{[N]}}{1^T Q_D^{[N]}} + \eta^{[OP]} \frac{Q_D^{[OP]}}{1^T Q_D^{[OP]}} + \eta^{[PK]} \frac{Q_D^{[PK]}}{1^T Q_D^{[PK]}} \right) \quad (44)$$

where η denotes the membership function. Similar approach is taken for modeling the generation bid. The details for modeling generation bid using the steady state probability is referred to [11].

For modeling the uncertainty in status of equipment, the model presented through Eqs. (7) and (8) is used directly by considering the same probabilities as $k \to \infty$, i.e. the steady state approximation. The resulting probability is given by

$$\pi_0[\infty] = \frac{\lambda}{\lambda + \mu} \quad (45)$$

$$\pi_1[\infty] = \frac{\mu}{\lambda + \mu} \quad (46)$$

Using the probabilities given in Eqs. (45) and (46), the probability for different system status can be derived. For example, the probability corresponding to having three transmission line failures is given as

$$\text{Prob(3 line outage)} = \begin{bmatrix} N_{TR} \\ 3 \end{bmatrix} \pi_0^3[\infty] \pi_1^{N_{TR}-3}[\infty] \quad (47)$$

IV. Example

We illustrate the approach described in the paper using the American Electric Power System (AEP) 118 Bus Test System. The system consists of 118 buses: 54 generators and 64 loads, and 186 transmission lines interconnecting the entire system; i.e. $N_B = 118$ ($N_G = 54$ and $N_D = 64$), and $N_{TR} = 186$.

The congestion cluster pricing method is to be implemented on the 118 bus system for the cluster boundaries defined at $k = 0$ for a season consisting of 90 days ($T = 2160$ (hours)). The maximum number of clusters allowed is limited to , i.e. $N_z = 15$. Thus, the maximum size of the search space is 1.37×10^{121} computed by

$$\begin{aligned} |\Theta| &= 15^{(118-15)} \\ &= 1.37 \times 10^{121} \end{aligned} \quad (48)$$

which is an astronomical figure.

For each load in the system, three types of load patterns are assigned: peak, off-peak and normal as shown in Figure 4. There is no uncertainty in generation. The transmission lines in the system may experience outages with the failure rate of $\lambda = 5 \times 10^{-4}$ and the repair rate of $\mu = 0.5$. For example, the probability associated with no transmission line failure is given by

$$\begin{aligned} \text{Prob(no line outage)} &= \begin{bmatrix} N_{TR} \\ 0 \end{bmatrix} \pi_0^0[\infty] \pi_1^{N_{TR}}[\infty] \\ &= 83\% \end{aligned} \quad (49)$$

Based on the system parameters it is determined that there are four critical lines (lines likely to be congested) in

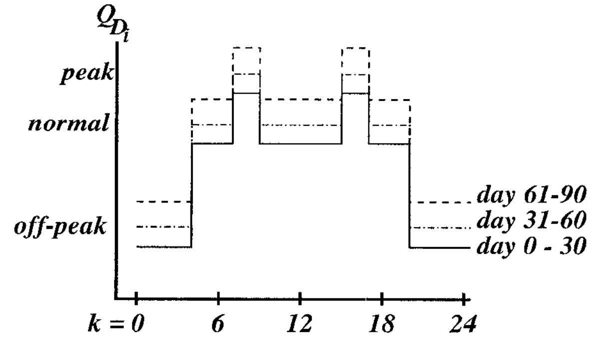

Fig. 4. Demand pattern for load i in the system

the system, namely the transmission lines between buses 30 and 38, between buses 59 and 63, between buses 70 and 71, and between buses 94 and 100.

The first cut cluster design is performed for each of these critical lines based on CDFs. The clusters are then superposed on top of each other to create the clusters over the entire season. The resulting number of clusters after the superposition is found to be 18. Therefore, the maximum size of the sample space is reduced to a measly 3,375 computed by

$$\begin{aligned} |\Theta'| &= 15^{(18-15)} \\ &= 3,375 \end{aligned} \quad (50)$$

The size of the actual sample space is even smaller once the clearly inferior cluster designs (or infeasible cluster designs) are eliminated from the initial sample space resulting in $|\Theta'| \approx 300$. From this sample space, 30 cluster designs are picked randomly to form a selected subset. The alignment probability for at least 3 matches in the selected subset of 30 designs for the top 50 designs is then, approximately 91% computed by

$$\text{Prob}\left(|G \cap S| \geq 3\right) = \sum_{i=3}^{30} \frac{\begin{bmatrix} 50 \\ i \end{bmatrix} \begin{bmatrix} 300-50 \\ 30-i \end{bmatrix}}{\begin{bmatrix} 300 \\ 30 \end{bmatrix}} = 90.91\% \quad (51)$$

Finally, the performance function is estimated for each of these 30 designs in the selected subset, S. The uncertainty in the status of transmission line is not considered at this estimation step. Table I summaries the estimated sample performance. As shown in the table, three cluster designs with the smallest evaluated performance are θ_9, θ_{13} and θ_{32}. Tables II and III describe how the clusters are defined for the first two designs.

For θ_{13} we incorporate the uncertainty in status of transmission lines into the estimation of the sample performance. It turns out that the probability associated with multiple line outages is very small; i.e. less than 1.6%. Thus, we consider only the single line outages. The newly estimated sample performance is given as

$$\mathcal{E}\left[\sum_{k=0}^{T} L'(\theta_9, \xi[k], k)\right] = 183.012 \quad (52)$$

Design	θ_1	θ_2	θ_3	θ_4	θ_5	θ_6
$L'(\cdot)$	189.283	185.457	188.195	189.283	185.678	187.644
Design	θ_7	θ_8	θ_9	θ_{10}	θ_{11}	θ_{12}
$L'(\cdot)$	185.841	187.121	184.424	185.407	185.436	185.709
Design	θ_{13}	θ_{14}	θ_{15}	θ_{16}	θ_{17}	θ_{18}
$L'(\cdot)$	184.205	185.430	187.070	187.425	185.733	187.447
Design	θ_{19}	θ_{20}	θ_{21}	θ_{22}	θ_{23}	θ_{24}
$L'(\cdot)$	185.736	185.687	188.195	184.481	184.434	184.478
Design	θ_{25}	θ_{26}	θ_{27}	θ_{28}	θ_{29}	θ_{30}
$L'(\cdot)$	184.440	184.470	187.277	184.727	185.687	184.440
Design	θ_{31}	θ_{32}	θ_{33}	θ_{34}	θ_{35}	
$L'(\cdot)$	187.978	184.243	186.929	185.457	188.195	

TABLE I

ESTIMATED SAMPLE PERFORMANCE FOR $\theta_i \in S$

Cluster #	Bus #
1	1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,117
2	25,26,27,28,29,31,32,114,115
3	16,17,18,19,30,113
4	20,21,22,23,24
5	37,38,39,40
6	33,34,35,36
7	79,80,98,99,100,101,102,103,104,105 106,107,108,109,110,111,112
8	43,44,45,46,47,48,49
9	41,42
10	50,51,52,53,54,55,56,57,58
11	59,60,61,62,66,67
12	63,64,65
13	77,78,82,83,84,85,86,87,88,89,90,91 92,93,94,95,96,97
14	68,69,70,74,75,76,81,116,118
15	71,72,73

TABLE II

INDIVIDUAL BUS CLUSTER AFFILIATION FOR θ_9

Cluster #	Bus #
1	1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,117
2	27,28,29,31,32,114,115
3	16,17,18,19,25,26,30,113
4	20,21,22,23,24
5	37,38,39,40,41,42
6	33,34,35,36,43,44
7	103,104,105,106,107,108,109,110,111,112
8	79,80,98,99,100,101,102
9	77,78,82,83,84,85,86,87,88,89,90,91,92,93 94,95,96,97
10	50,51,52,53,54,55,56,57,58
11	59,60,61,62,66,67
12	63,64,65
13	45,46,47,48,49
14	68,69,70,74,75,76,81,116,118
15	71,72,73

TABLE III

INDIVIDUAL BUS CLUSTER AFFILIATION FOR θ_{13}

As expected some slight correction is made to the earlier estimation of the sample performance.[3]

V. CONCLUSION

In this paper a practical approach to implementing the congestion cluster pricing method have been introduced.

We have presented the formulation for the implementation of the congestion cluster pricing method as a stochastic optimization problem in which the minimum desired criteria for the method is translated into the performance function.

After introducing the uncertainty in the system, we have discussed some heuristic techniques to finding the solution to the newly formulated optimization problem. Given that the search based method is preferred for solving the problem, these heuristic techniques are particularly important because of the high degree of the stochastic nature and because of the large size of the search space. The ordinal optimization (OO) principles are used to provide some justifications to the techniques.

The OO principles are believed to be quite useful for exploring the presented optimization problem further. The natural next step may be employing the OO method to evaluate the sample performances only for ranking various cluster design alternatives rather than for calculating the actual performance measure for a particular design. The Monte Carlo formulation given for evaluating the sample performance may then be solved directly without much concern for the large number of iteration. The coefficient of variation needs to be computed from the ordinal optimization perspective if the Monte Carlo method is used so that some confidence bound can be estimated for the accurate alignment probability analysis. Finally, some sophisticated numerical techniques such as the importance sampling may also be worthwhile exploring as many uncertainties in the system presented in the paper have very low probabilities but high impact.

[3]This system exhibits a somewhat degenerate feature of the reduced system-wide generation cost with some of the lines taken out. This implies that the system operator may reduce the system congestion by cleverly controlling the existing resources.

Acknowledgment

The authors greatly appreciate the financial support provided for this research by the members of the MIT Consortium on *New Concepts and software for Competitive Power Systems: Operations and Management* and by the members of the EPRI/DoD CIN/SI Research Consortium on *Modeling and Diagnosis Methods for Large-Scale Complex Networks*.

References

[1] Allen, E., "Stochastic Unit Commitment in a Deregulated Electric Utility Industry." Ph.D. thesis, Massachusetts Institute of Technology, 1998

[2] Faruqui, A., Malko, J. R., eds. "Customer Choice: Finding Value in Retail Electricity Markets", Public Utilities Reports, Vienna, VA, p xviii, 365. 1999.

[3] Ho, Y.-C., "An Explanation of Ordinal Optimization: Soft Computing for Hard Problem", *Information Sciences* 113, 1999. pp. 169-192.

[4] Ilić, M., Galiana, F., Fink, L., "Power systems Restructuring. Engineering and Economics", Kluwer Academic Publishers. 1998

[5] Lau, T.W.E., Ho, Y.C., "Universal Alignment Probabilities and Subset Selection for Ordinal Optimization", *Journal of Optimization Theory & Applications*, vol.93, no.3, June, 1997. pp.455-89

[6] Salehfar, H., Li, T., "Stochastic Petri Nets for Reliability Assessment of Power Generating Systems with Operating Considerations", *Proceedings of the IEEE PES Winter Meeting*, New York, NY, 1999.

[7] Schweppe, F , Caramanis, M., Tabors, R., Bohn, R., "Spot Pricing of Electricity", Kluwer Academics, 1988.

[8] Yoon, Y., Arce, J., Collison, K., Ilić, M., "Implementation of Cluster-based Congestion Management Systems", 4th Intl. Conf. on Power System Operation and Planning (ICPSOP'2000), Acra, Ghana, Aug. 2000.

[9] Yoon, Y., Collison, K., Arce, J., Ilić, M., "Congestion Management System Methods: Comparison on the 118 Bus System", submitted for publication at 32nd Annual North America Power Symposium (NAPS 2000), Ontario, Canada, Oct. 2000.

[10] Yu, C-N, Ilić, M., "Congestion Clusters-Based Markets for Transmission Management", *Proceedings of the IEEE PES Winter Meeting*, New York, NY, 1999.

[11] Yu, C-N., Ilić, M.D., "An Algorithm for Implementing Transmission Rights in a Competitive Power Industry", *Proceedings of the IEEE PES Winter Meeting*, Singapore, 2000.