# Experimental Analysis of Algorithms for Bilateral-Contract Clearing Mechanisms Arising in Deregulated Power Industry

Chris Barrett[2], Doug Cook[1], Gregory Hicks[6], Vance Faber[5], Achla Marathe[2], Madhav Marathe[2], Aravind Srinivasan[3], Yoram J. Sussmann[4], and Heidi Thornquist[7]

[1] Department of Engineering, Colorado School of Mines, Golden CO 80401.
[2] D-2 and CCS-3, Current address: P.O. Box 1663, MS M997, Los Alamos National Laboratory, Los Alamos NM 87545. Email: `barrett,achla,marathe@lanl.gov`.
[3] Lucent Technologies, 600 Mountain Ave., Murray Hill, NJ 07974-0636. Part of this work was done while at the National University of Singapore. Email: `srin@research.bell-labs.com`.
[4] Lizardtech Inc. Email: `vance@lizardtech.com`.
[5] Email: `yorams@atl.fundtech.com`.
[6] Department of Mathematics, North Carolina State University, Raleigh. Email: `hicksgp@yahoo.com`.
[7] Department of Computational Science, Rice University, Houston, TX. Email: `heidi@caam.rice.edu`.

**Abstract.** We consider the bilateral contract satisfaction problem arising from electrical power networks due to the proposed deregulation of the electric utility industry in the USA. Given a network and a (multi)set of pairs of vertices (contracts) with associated demands, the goal is to find the maximum number of simultaneously satisfiable contracts. We study how four different algorithms perform in fairly realistic settings; we use an approximate electrical power network from Colorado. Our experiments show that three heuristics outperform a theoretically better algorithm. We also test the algorithms on four types of *scenarios* that are likely to occur in a deregulated marketplace. Our results show that the networks that are adequate in a regulated marketplace might be inadequate for satisfying all the bilateral contracts in a deregulated industry.

## 1 Introduction

The U.S. electric utility industry is undergoing major structural changes in an effort to make it more competitive [21, 11, 19, 12]. One major consequence of the deregulation will be to decouple the controllers of the network from the power producers, making it difficult to regulate the levels of power on the network; consumers as well as producers will eventually be able to negotiate prices to buy and sell electricity. See [18]. In practice, deregulation is complicated by the facts that all power companies will have to share the same power network in the short term, with the network's capacity being just about sufficient to meet the current demand. To overcome these problems, most U.S. states have set up an ISO (independent system operator): a non-profit governing body to arbitrate the use of the network. The basic questions facing ISOs are how to decide which contracts to deny (due to capacity constraints), and who is to bear the costs accrued when contracts are denied. Several criteria/policies have been proposed and/or are being legislated by the states as possible guidelines for the ISO to select a maximum-sized subset of contracts that can be cleared simultaneously [18]. These include: (a) Minimum Flow Denied: The ISO selects the subset of contracts that denies the *least* amount of proposed power flow. This proposal favors clearing bigger contracts first. (b) First-in First-out: The contract that comes first gets cleared first; this is the least discriminating to the contractors. (c) Maximum Consumers Served: This clears the smallest contracts first and favors the small buyers whose interests normally tend to go unheard.

There are three key issues in deciding policies that entail specific mechanisms for selecting a subset of contracts: *fairness* of a given policy to producers and consumers; the *computational complexity* of implementing a policy, and how *sound* a given policy is from an economic standpoint. (For instance, does the policy promote the optimal clearing price/network utilization etc.) Here we focus on evaluating the efficacy of a given policy with regard to its computational resource requirement and network resource utilization. It is intuitively clear that the underlying network, its capacity and topology, and the spatial locations of the bilateral contracts on the network, will play an important role in determining the efficacy of these policies. We do not discuss here the fairness and economics aspects of these policies: these are subjects of a companion paper. The work reported here was done as part of two inter-related projects at Los Alamos. The first project is to develop a mathematical and computational theory of simulations based on Sequential Dynamical Systems. The second aims at developing a comprehensive coupled simulation[1] of the *deregulated* electrical power industry [7]. See `http://www.lanl.gov/orgs/tsa/tsasa/` for additional details. To achieve this goal, we carry out the first experimental analysis of several algorithms for simultaneously clearing a maximal number of bilateral contracts. The algorithms were chosen according to provable performance, ability to serve as a proxy for some of the above-stated policies, and computational requirement. The algorithms are as follows; see § 3 for their specification. The ILP-Randomized rounding (RR) algorithm has a provable performance guarantee under certain conditions. The computational resource requirement was quite high, but the approach also provides us with an upper bound on any optimal solution and proved useful in comparing the performance of the algorithms. The Largest-First Heuristic (LF) is a proxy for the *Minimum Flow Denied* policy. The Smallest-First Heuristic (SF) serves as a proxy for the *Maximum Contracts Served* policy. The Random-Order Heuristic (RO) clears the contracts in the random order. This algorithm was chosen as a proxy for the *First-in First-out* policy. Such a policy is probably the most natural clearing mechanism and is currently in place at many exchanges.

[1] This coupled simulation has a market component and a physical network component.

We used a coarse representation of the Colorado electrical power network (see § 4) to qualitatively compare the four algorithms discussed above in fairly realistic settings. The realistic networks differ from random networks and structured networks in the following ways: (i) Realistic networks typically have a very low average degree. In fact, in our case the average degree of the network is no more than 3. (ii) Realistic networks are not very uniform. One typically sees one or two large clusters (downtown and neighboring areas) and small clusters spread out throughout. (iii) For most empirical studies with random networks, the edge weights are chosen independently and uniformly at random from a given interval. However, realistic networks typically have very specific kinds of capacities since they are constructed with particular design goal.

From our preliminary analysis, it appears that although the simple heuristic algorithms do not have worst-case guarantees, they outperform the theoretically better randomized rounding algorithm. We tested the algorithms on four carefully chosen scenarios. Each scenario was designed to test the algorithms and the resulting solutions in a deregulated setting. The empirical results show that networks that are capable of satisfying all demand in a regulated marketplace can often be inadequate for satisfying all (or even a acceptable fraction) of the bilateral contracts in a deregulated market. Our results also confirm intuitive observations: e.g., the number of contracts satisfied crucially depends on the scenario and the algorithm.

As far as we are aware, this is the first study to investigate the efficacy of various policies for contract satisfaction in a deregulated power industry. Since it was done in fairly realistic settings, the qualitative results obtained here have implications for policy makers. To compare the algorithms in a quantitative and (semi-)rigorous way, we employ statistical tools and experimental designs. Although many of the basic tools are standard in statistics, the use of formal statistical methods in experimental algorithmics for analyzing/comparing the performance of heuristics has not been investigated to the best of our knowledge. We believe that such statistical methods should be investigated further by the experimental algorithmics community for deriving more quantitative conclusions when theoretical proofs are hard or not very insightful. Our results can also be applied in other settings, such as bandwidth-trading on the Internet. See, e.g., [2]. Finally, to our knowledge, previous researchers have not considered the effect of the underlying network on the problems; this is an important parameter especially in a free-market scenario.

## 2 Problem Definitions

We briefly define the optimization problems studied here. We are given an undirected network (the power network) $G = (V, E)$ with capacities $c_e$ for each edge $e$ and a set of source-sink node pairs $(s_i, t_i)$, $1 \le i \le k$. Each pair $(s_i, t_i)$ has: (i) an integral *demand* reflecting the amount of power that $s_i$ agrees to supply to $t_i$ and (ii) a negotiated *cost* of sending unit commodity from $s_i$ to $t_i$. As is traditional in the power literature, we will refer to the source-sink pairs along with the associated demands as *a set of contracts*. In general, a source or sink may have multiple associated contracts. We find the following notation convenient to describe the problems. Denote the set of nodes by $N$. The contracts are defined by a relation $R \subseteq (N \times N \times \Re \times \Re)$ so that tuple $(v, w, \alpha, \beta) \in R$ denotes a contract between source $v$ and sink $w$ for $\alpha$ units of commodity at a cost of $\beta$ per unit of the commodity. For $A = (v, w, \alpha, \beta) \in R$ we denote $\mathit{source}(A) = v$, $\mathit{sink}(A) = w$, $\mathit{flow}(A) = \alpha$ and $\mathit{cost}(A) = \beta$. Corresponding to the power network, we construct a digraph $H = (V \cup S \cup T \cup \{s, t\}, E')$ with source $s$, sink node $t$, capacities $u : E' \to \Re$ and costs $c' : E' \to \Re$ as follows. For all $A \in R$, define new vertices $v_A$ and $w_A$. Let $S = \{v_A \mid A \in R\}$ and $T = \{w_A \mid A \in R\}$. Each edge $\{x, y\}$ from $G$ is present in $H$ as the two arcs $(x, y)$ and $(y, x)$ that have the same capacity as $\{x, y\}$ has in $G$, and with cost 0. In addition, for all $A = (v, w, \alpha, \beta) \in R$, we introduce: (i) arcs $(v_A, v)$ and $(w, w_A)$ with infinite capacity and zero cost; (ii) arc $(s, v_A)$ with capacity $\mathit{flow}(A) = \alpha$ and cost 0; and (iii) arc $(w_A, t)$ with capacity $\mathit{flow}(A) = \alpha$ and cost equaling $\mathit{cost}(A)$. By this construction, we can assume without loss of generality that each node can participate in exactly one contract. A *flow* is simply an assignment of values to the edges in a graph, where the value of an edge is the amount of flow traveling on that edge. The value of the flow is defined as the amount of flow coming out of $s$ (or equivalently the amount of flow coming in to $t$). A generic *feasible flow* $f = (f_{x,y} \geq 0 : (x, y) \in E')$ in $H$ is any non-negative flow that: (a) respects the arc capacities, (b) has $s$ as the only source of flow and $t$ as the only sink. Note that for a given $A \in R$, in general it is not necessary that $f_{s,v_A} = f_{w_A,t}$. For a given contract $A \in R$, $A$ is said to be *satisfied* if the feasible flow $f$ in $H$ has the additional property that for $A = (v, w, \alpha, \beta)$, $f_{s,v_A} = f_{w_A,t} = \alpha$; i.e., the contractual obligation of $\alpha$ units of commodity is shipped out of $v$ and the same amount is received at $w$. Given a power network $G(V, E)$, a contract set $R$ is *feasible* (or *satisfied*) if there exists a feasible flow $f$ in the digraph $H$ that satisfies every contract $A \in R$. Let $B = \mathit{supply}(s) = \mathit{demand}(t) = \sum_{A \in R} \mathit{flow}(A)$.

In practice, it is typically the case that $R$ does not form a feasible set. As a result we have two possible alternative methods of relaxing the constraints: (i) relax the notion of feasibility of a contract and (ii) try and find a subset of contracts that are feasible. Combining these two alternatives we define the following types of "relaxed feasible" subsets of $R$. We will concern ourselves with only one variant here. A contract set $R' \subseteq R$ is a *0/1-contract satisfaction* feasible set if, $\forall A = (v, w, \alpha, \beta) \in R'$, $f_{s,v_A} = f_{w_A,t} = \alpha$.

**Definition 1.** *Given a graph $G(V, E)$ and a contract set $R$, the (0/1-*VERSION*, *MAX-FEASIBLE FLOW*) problem is to find a feasible flow $f$ in $H$ such that $\sum_{A \in R'} f(A)$ is maximized where $R'$ forms a* 0/1-contract satisfaction *feasible set of contracts. In the related (0/1-*VERSION*, *MAX-#CONTRACTS*) problem, we aim to find a feasible flow $f$ in $H$ such that $|R'|$ is maximized, where $R'$ forms a* 0/1-contract satisfaction *feasible set of contracts.*

Though such electric flow problems have some similarities with those from other practical situations, there are many basic differences such as reliability, non-distinguishability between the power produced by different generators, short life-time due to inadequate storage, line effects etc. [22]. The variants of flow problems related to power transmission studied here are intuitively harder than traditional multi-commodity flow problems, since we *cannot distinguish between* the flow "commodities" (power produced by different generators). As a result, current solution techniques used to solve single/multi-commodity flow problems are not directly applicable to the problems considered here.

## 3 Description and Discussion of Algorithms

We work on the (0/1-Version, Max-#Contracts) problem here. Let $n$ and $m$ respectively denote the number of vertices and edges in the network $G$. In [6], it was shown that (0/1-Version, Max-#Contracts) is *NP*-hard; also, unless $NP \subseteq ZPP$, it cannot be approximated to within a factor of $m^{1/2-\epsilon}$ for any fixed $\epsilon > 0$, in polynomial time. Thus, we need to consider good heuristics/approximation algorithms. First, there are three simple heuristics. The Smallest-First Heuristic considers the contracts in non-decreasing order of their demands. When a contract is considered, we accept it if it can be feasibly added to the current set of chosen contracts, and reject it otherwise. The Largest-First Heuristic is the same, except that the contracts are ordered in non-increasing order of demands. In the Random-Order heuristic, the contracts are considered in a random order.

We next briefly discuss an approximation algorithm of [6]. This has proven performance only when all source vertices $s_i$ are the same; however, this algorithm extends naturally to the multi-source case which we work on. An integer linear programming (ILP) formulation for the problem is considered in [6]. We have indicator variables $x_i$ for the contract between $s_i$ and $t_i$, and variables $z_{i,e}$ for each $(s_i, t_i)$ pair and each edge $e$. The intended meaning of $x_i$ is that the total flow for $(s_i, t_i)$ is $d_i x_i$; the meaning of $z_{i,e}$ is that the flow due to the contract between $(s_i, t_i)$ on edge $e$ is $z_{i,e}$. We write the obvious flow and capacity constraints. Crucially, we also add the valid constraint $z_{i,e} \le c_e x_i$ for all $i$ and $e$. In the integral version of the problem, we will have $x_i \in \{0, 1\}$, and the $z_{i,e}$ as non-negative reals. We relax the condition "$x_i \in \{0, 1\}$" to "$x_i \in [0, 1]$" and solve the resultant LP; let $y^*$ be the LP's optimal objective function value. We perform the following rounding steps using a carefully chosen parameter $\lambda > 1$. (a) Independently for each $i$, set a random variable $Y_i$ to 1 with probability $x_i/\lambda$, and $Y_i := 0$ with probability $1 - x_i/\lambda$. (b) If $Y_i = 1$, we will choose to satisfy $(1 - \epsilon)$ of $(s_i, t_i)$'s contract: for all $e \in E$, set $z_{i,e} := z_{i,e}(1 - \epsilon)/x_i$. (c) If $Y_i = 0$, we choose to have no flow for $(s_i, t_i)$: i.e., we will reset all the $z_{i,e}$ to 0. A deterministic version of this result based on *pessimistic estimators*, is also provided in [6]; see [6] for further details.

**Theorem 1.** **([6])** *Given a network $G$ and a contract set $R$, we can find an approximation algorithm for* (0/1-Version, Max-#Contracts) *when all source vertices are the same, as follows. Let $OPT$ be the optimum value of the problem, and $m$ be the number of edges in $G$. Then, for any given $\epsilon > 0$, we can in polynomial time find a subset of contracts $R'$ with total weight $\Omega(OPT \cdot \min\{(OPT/m)^{(1-\epsilon)/\epsilon}, 1\})$ such that for all $i \in R'$, the flow is at least $(1 - \epsilon)d_i$.*

## 4 Experimental Setup and Methodology

To test our algorithms experimentally, we used a network corresponding to a subset of a real power network along with contracts that we generated using different scenarios. The network we used (as shown in Figure 1) is based on the power grid in Colorado and was derived from data obtained from PSCo's (Public Service Company of Colorado) Draft Integrated Resources Plan listing of power stations and major substations. We restricted our attention to major trunks only. The plant capacities were derived from the PSCo data. To generate maximum aggregate demands for the sink nodes we used data from the US Census bureau on the number of household per county. The demands were generated by assigning counties to specific sink nodes in the network.

All the test cases were generated from the basic model. The general approach we used was to fix the edge capacities and generate source-sink contracts combinations using the capacities and aggregate demands in the basic model as upper bounds. To ensure that the test cases we generated corresponded to (1) difficult problems, i.e. infeasible sets of contracts, and (2) problems that might reasonably arise in reality, we developed several scenarios that included an element of randomness.

The current implementation is still based upon a network which should be feasible only if the total source capacity is greater than the total sink capacity and the only requirement is that the total sink capacity be satisfied regardless of which source provides the power. **Scenarios 1, 2, 3** and **4** are based around the network with total generating capacity, **6249 MW**, and reduced sink capacities near **4400MW** combined.

*Scenario 1:* The source capacity of the network was reduced until the maximum flow in the network was slightly less than the demand.

*Scenario 2:* For this scenario, we took the basic network and increased the sink capacity while the source capacity remained fixed.

*Scenario 3:* The edge capacities were adjusted, reduced in most cases, to limit the network to a maximum flow of slightly more than 4400MW given its source and sink distribution. Here, if the load is allowed to be fulfilled from any source (as is normally done with centralized control), the network and the edge capacities are enough to handle a total of 4400MW. However, if we insist that a particular source needs to serve a particular sink (as is done in bilateral contract satisfaction), then the capacities may not be enough to handle the same load of 4400MW.

*Scenario 4:* For this scenario, we took the network of Scenario 3 and biased the selection of source nodes towards the lower valued source units.

**Methodology:** We worked with the four scenarios and ran all four algorithms for each. For the three greedy heuristics the implementations are fairly straightforward, and we used public-domain network flow codes. Implementing the randomized rounding procedure requires extra care. The pessimistic estimator ap-

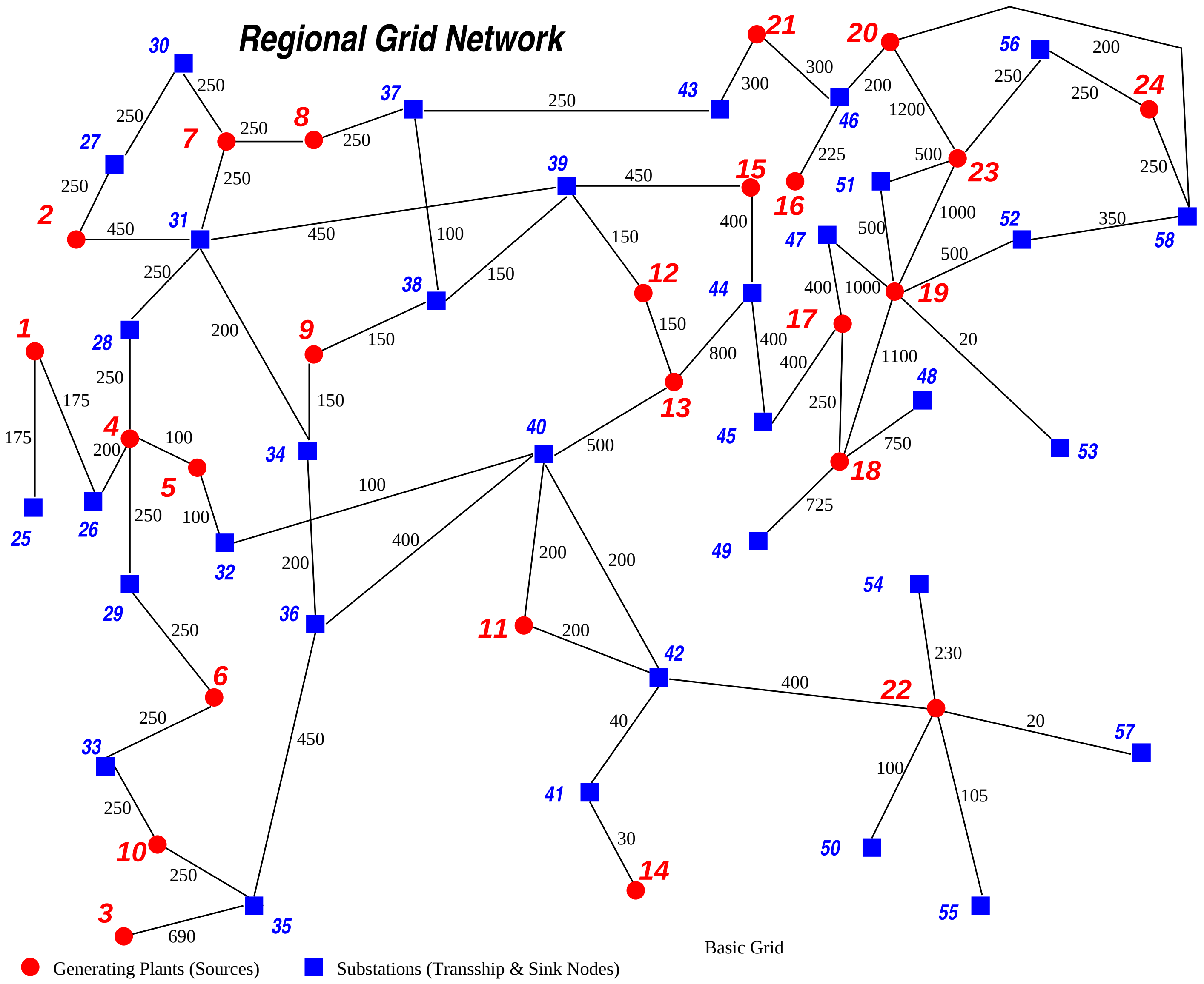


**Fig. 1.** This shows the network with nodes numbered as they are referenced in all scenarios and edge capacities labeled at values used for Scenarios 1 & 2. The placement of the nodes and edges are what is probably the final form. The least number of edges cross and the nodes in the upper right are spread out a little bit maintaining the general feel of the distribution while allowing easier reading.

proach of [6] works with very low probabilities, and requires significant, repeated re-scaling in practice. Thus we focus on the randomized version of the algorithm of [6]; five representative values of $\epsilon$ varying from .1 to .5 were chosen. We believe that satisfying a contract partially so that a contract is assigned less than .5 of the required demand is not very realistic. For each scenario, and for each of the 5 values of $\epsilon$, the programs implementing the algorithms under inspection produced 30 files from which the following information could be extracted: the running times and solutions of all four algorithms, and the LP upper bound on the ILP. The number 30 was chosen to ensure that a statistically "large" sample of each measure would be provided in order to make valid statistical inference. More attention is given to the quality-of-solution measure of an algorithm (rather than the running-time measure), since from a social standpoint, contract satisfaction may leave little room for finding solutions that are far from optimal. For a given algorithm $\mathcal{A}$, let $Value_{\mathcal{A}}$ denote the number of contracts that can be satisfied by $\mathcal{A}$. $p_{\mathcal{A}} = \frac{Value_{\mathcal{A}}}{\lfloor y^* \rfloor}$ provides a measure of the quality of the algorithm's solution. The value $\lfloor y^* \rfloor$ provides an upper bound on the objective function value. The objective of our experiments was to find out which, if any, of the algorithms discussed here performs better than the others, in terms of quality of solution and running time for different contract scenarios.

## 5 Results and Analysis

### 5.1 Statistical Analysis

We use a statistical technique known as *analysis of variance* (ANOVA) to test whether differences in the sample means of algorithms and scenarios reflect differences in the means of the statistical populations that they came from or are just sampling fluctuations. This will help us identify which algorithm and scenarios perform the best. See ([9], [10]) for details.

**Quality of Solution:** We first describe the experiment for the quality of solution, i.e., $p_{\mathcal{A}}$. We use a two-factor ANOVA model since our experiment involves two factors: the algorithms $\mathcal{A}_i$, $i = 1, 2, 3$ and 4, and scenarios $\mathcal{S}_j$, $j = 1, 2, 3$ and 4. Following classical statistics terminology, we will sometimes refer to algorithms as *treatments* and the scenarios as *blocks*. We will use $\mathcal{A}$ to denote the set of algorithms and $\mathcal{S}$ to denote the set of scenarios. For each algorithm-scenario pair we have *30 observations (or replicates)*. When testing the efficacy of the algorithms, we use 4 algorithms, each having 120 observations (30 for each scenario) from the corresponding population. The design of experiment used here is a *fixed-effect complete randomized block*. The model for randomized block design includes constants for measuring the scenario effect (block effect), the algorithm effect (treatment effect) and a possible interaction between the scenarios and the algorithms. An appropriate mathematical model for testing the above hypothesis is given by: $X_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \varepsilon_{ijk}$, where $X_{ijk}$ is the measurement ($p_{\mathcal{A}}$) for the $kth$ sample within the $ith$ algorithm and the $jth$ scenario. $\tau_i$ is the algorithm effect. $\beta_j$ is the scenario effect. $(\tau\beta)_{ij}$ captures the interaction present between the algorithms and the scenarios. $\varepsilon_{ijk}$ is the random error. We use S-Plus [16] software to run two-factor ANOVA to test the following three different hypotheses; as is standard, when discussing any one of these hypotheses, we refer to it as the *null hypothesis*.

1. Are the means given by the 4 different algorithms equal? That is, the hypothesis $H_0$ here is that $\tau_1 = \tau_2 = \tau_3 = \tau_4$.
2. Are the means given by the 4 different scenarios equal? The hypothesis $H_0$ here is that $\beta_1 = \beta_2 = \beta_3 = \beta_4$.
3. Is there any interaction between the two factors? $H_0$ here is that $(\tau\beta)_{ij} = 0$.

The results of two-factor ANOVA are shown in Table 1 and Table 2. In Table 1, the term $\overline{X}_{11\cdot}$, for instance, means that we fix $i = 1$ and $j = 1$, and take the average of $X_{ijk}$ over all $k$. Similarly for other such terms. In the following discussion, we explain the meaning of each column. $DF$ refers to the degrees of freedom, $SS$ refers to the sum of squared deviations from the mean. $MS$ refers to the mean square error, which is the sum of squares divided by the degrees of freedom. The $p$-value gives the smallest level of significance at which the null hypothesis can be rejected. The lower the $p$-value, the lesser the agreement between the data and the null hypothesis. Finally the $F$**-test** is as follows. To test the null hypothesis that the population means are equal, ANOVA compares two estimates of $\sigma^2$. The first estimate is based on the variability of each population mean around the grand mean. The second is based on the variability of the observations in each population around the mean of that population. If the null hypothesis is true, the two estimates of $\sigma^2$ should be essentially the same. Otherwise, if the populations have different means, the variability of the population mean around the grand mean will be much higher than the variability within the population. The null hypothesis in the $F$-test will be accepted if the two estimates of $\sigma^2$ are almost equal.

In a two-factor fixed-effect ANOVA, three separate $F$**-tests** are performed: two tests for the factors and the third for the interaction term. If the $F$-ratio is close to 1, the null hypothesis is true. If it is considerably larger–implying that the variance between means is larger than the variance within a population–the null hypothesis is rejected. The $F$-distribution table should be checked to see if the $F$-ratio is significantly large. The results in Table 2 show that all the above three null hypothesis are rejected at any significance level. This implies that the performance (measured by $p_{\mathcal{A}}$) of at least one of the algorithms is significantly different from the other algorithms. Also, different scenarios make a difference in the performance. Finally, the scenarios and the algorithms interact in a significant way. The interaction implies that the performance of the algorithms are different for different scenarios. More details on what caused the rejection of null hypotheses and a discussion of running times are omitted here due to the lack of space. Interested readers can request a copy of the more complete version of this work by contacting one of the authors.

| Performance Measure: Quality of Solution (in %) | | | | | |
|---|---|---|---|---|---|
| | RR | SF | LF | RO | Scenario Means |
| Scenario 1 | $\overline{X}_{11\cdot}$=48.68 | $\overline{X}_{21\cdot}$=99.73 | $\overline{X}_{31\cdot}$=97.97 | $\overline{X}_{41\cdot}$=97.78 | $\overline{X}_{\cdot 1\cdot}$=86.02 |
| Scenario 2 | $\overline{X}_{12\cdot}$=46.91 | $\overline{X}_{22\cdot}$=99.56 | $\overline{X}_{32\cdot}$=98.38 | $\overline{X}_{42\cdot}$=98.93 | $\overline{X}_{\cdot 2\cdot}$=85.94 |
| Scenario 3 | $\overline{X}_{13\cdot}$=45.69 | $\overline{X}_{23\cdot}$=99.25 | $\overline{X}_{33\cdot}$=97.10 | $\overline{X}_{43\cdot}$=98.82 | $\overline{X}_{\cdot 3\cdot}$=85.22 |
| Scenario 4 | $\overline{X}_{14\cdot}$=46.99 | $\overline{X}_{24\cdot}$=98.03 | $\overline{X}_{34\cdot}$=88.65 | $\overline{X}_{44\cdot}$=93.41 | $\overline{X}_{\cdot 4\cdot}$=81.77 |
| Algo. Means | $\overline{X}_{1\cdot\cdot}$=47.07 | $\overline{X}_{2\cdot\cdot}$=99.14 | $\overline{X}_{3\cdot\cdot}$=95.51 | $\overline{X}_{4\cdot\cdot}$=97.24 | $\overline{X}_{\cdots}$ = 84.74 |

**Table 1.** The Mean Values of the Quality of Solution.

| Source | $DF$ | $SS$ | $MS$ | $F$-**test** | $p$-**value** |
|---|---|---|---|---|---|
| Scenario (Block) | 3 | 0.14 | 0.05 | 43.38 | 0 |
| Algorithm (Treatment) | 3 | 22.78 | 7.59 | 6792.60 | 0 |
| Scenario:Algorithm | 9 | 0.12 | 0.01 | 15.90 | 0 |
| Residuals | 464 | 0.40 | .0008 | | |
| Total | 479 | 23.45 | | | |

**Table 2.** Results of Two-Factor ANOVA for Quality of Solution

### 5.2 General Conclusions

1. Although there exist instances where the three heuristics produce solutions as large as $\Omega(n)$ times the optimal fractional solution, most of our tests show that we could find integral solutions fairly close to optimal.

2. Our experiments show that different scenarios make a *significant* difference in the type of solutions obtained. For example, the quality of solution obtained using the fourth scenario is significantly worse than the first three scenarios. The sensitivity to the scenarios poses interesting questions for infrastructure investment. The market will have to decide the cost that needs to be paid for expecting the necessary quality of service. It also brings forth the equity-benefit question: namely, who should pay for the infrastructure improvements?

3. It is possible that for certain scenarios, the underlying network is incapable of supporting even an acceptable fraction of the bilateral contracts. This observation although fairly intuitive provides an extremely important message, namely, networks that were adequate to service customers in a completely regulated power market *might* not be adequate in deregulated markets. This makes the question of evicting the bilateral contracts more important.

4. One expects a trade-off between the number of contracts satisfied and the value of $\epsilon$, for the randomized rounding algorithm: as $\epsilon$ increases, and the demand condition is more relaxed, a higher number of contracts should get satisfied. But our experiments show that the change in the number of contracts satisfied for different values of $\epsilon$ is insignificant. Also, $\lambda = 2$ gave the best solutions in our experiments.

5. In practical situations, the Random-Order heuristic would be the best to use since it performs very close to the optimal in terms of quality of solution and has very low running time. Furthermore, though the Smallest-First heuristic does even better on many of our experiments, Random-Order is a natural proxy to model contracts arriving in an unforeseen way. Also, since the heuristics deliver solutions very close to the LP upper bound, we see that this LP bound is tight and useful. To further evaluate the randomized rounding algorithm, we need to implement its deterministic version [6], which is a non-trivial task.

**Acknowledgments:** All authors except Aravind Srinivasan were supported by the Department of Energy under Contract W-7405-ENG-36. We thank Darrell Morgensen and Andy White for their support. Work by Doug Cook, Vance Faber, Gregory Hicks, Yoram Sussmann and Heidi Thornquist was done when they were visiting Los Alamos National Laboratory. We thank Dr. Martin Wildberger (EPRI) for introducing us to the problems considered herein, and for the extensive discussions. We thank Mike Fugate, Dean Risenger (Los Alamos National Laboratory) and Sean Hallgren (U.C. Berkeley) for helping us with the experimental evaluation. We also thank Massoud Amin (EPRI), Terence Kelly (University of Michigan), R. Ravi (CMU) and the members of the Power Simulation project (ElectriSim) at Los Alamos National Laboratory: Dale Henderson, Jonathan Dowell, Martin Drozda, Verne Loose, Doug Roberts, Frank Wolak (Stanford University) and Gerald Sheble (Iowa State University) for fruitful discussions. Finally, thanks to the WAE 2001 referees for their helpful suggestions.

## References

1. Ahuja, R.K., Magnanti, T.L., Orlin, J.B.: Network flows: theory, algorithms, and applications. Prentice Hall, Englewood Cliffs, New Jersey. (1993)
2. Example sites for information on bandwidth trading: `http://www.lightrade.com/dowjones2.html` `http://www.williamscommunications.com/prodserv/network/webcasts/bandwidth.html`
3. Barrett, C., Hunt, H.B., Marathe, M., Ravi, S.S., Rosenkrantz, D., Stearns, R.: Dichotomy Results for Sequential Dynamical Systems. To appear in Proc. MFCS (2001)
4. Cardell, J.B., Hitt, C.C., Hogan, W.W.: Market Power and Strategic Interaction in Electricity Networks. Resource and Energy Economics Vol. 19 (1997) 109-137.
5. California Public Utilities Commission. Order Instituting Rulemaking on the Commission's Proposed Policies Governing Restructuring California's Electric Services Industry and Reforming Regulation, Decision 95-12-063 (1995)
6. Cook, D., Faber, V., Marathe, M., Srinivasan, A., Sussmann, Y.J.: Low Bandwidth Routing and Electrical Power Networks. Proc. 25th International Colloquium on Automata, Languages and Programming (ICALP) Aalborg, Denmark, LNCS 1443, Springer Verlag, (1998) 604-615
7. Dowell, L.J., Drozda, M., Henderson, D.B., Loose, V., Marathe, M., Roberts, D.: ELISIMS: Comprehensive Detailed Simulation of the Electric Power In-

dustry. Technical Report LA-UR-98-1739, Los Alamos National Laboratory (1998)
8. The Changing Structure of the Electric Power Industry: Selected Issues. DOE 0562(98), Energy Information Administration, US Department of Energy, Washington, D.C. (1998)
9. Fryer, H.C.: Concepts and Methods of Experimental Statistics. Allyn and Bacon Inc. (1968)
10. Glass, G.V., Hopkins, K.D.: 1996. Statistical Methods in Education and Psychology. third edition (1996)
11. EPRI-Workshop: Underlying Technical Issues in Electricity Deregulation. Technical report forthcoming, Electric Power Research Institute (EPRI), April 25-27 (1997)
12. The U.S. Federal Energy Regulatory Commission. Notice of Proposed Rulemaking ("NOPRA"). Dockets No. RM95-8-000, RM94-7-001, RM95-9-000, RM96-11-000, April 24 (1996)
13. Garey, M.R., Johnson, D.S.: Computers and Intractability. A Guide to the Theory of NP-Completeness. Freeman, San Francisco CA (1979)
14. Hogan, W.: Contract networks for electric power transmission. J. Regulatory Economics. (1992) 211-242
15. Kolliopoulos, S.G., Stein, C.: Improved approximation algorithms for unsplittable flow problems. In Proc. IEEE Symposium on Foundations of Computer Science. (1997) 426-435
16. S-Plus5, "Guide to Statistics", MathSoft Inc. September (1988)
17. Srinivasan, A.: Approximation Algorithms via Randomized Rounding: A Survey. Lectures on Approximation and Randomized Algorithms (M. Karoński and H. J. Prömel, eds.), Series in Advanced Topics in Mathematics. Polish Scientific Publishers PWN, Warsaw.(1999) 9-71
18. Wildberger, A. M.: Issues associated with real time pricing. Unpublished Technical report, Electric Power Research Institute (EPRI) (1997)
19. Websites: http://www.magicnet.net/∼metzler/page2d.html and http://www.enron.com/choice/dereg.fset.html (also see references therein). http://www.eia.doe.gov/ (Energy Information Administration)
20. Wolak, F.: An Empirical Analysis of the Impact of Hedge Contracts on Bidding Behavior in a Competitive Electricity Market, `http://www.stanford.edu/∼wolak`. (2000)
21. The Wall Street Journal, August 04 2000, A2, August 14 2000 A6 and August 21 2000, November 27 2000, December 13, 14 2000.
22. Wood, A.J., Wollenberg, B.F.: Power Generation, Operation and Control. John Wiley and Sons. (1996)