A decision support system for the electrical power districting problem

Paul K. Bergey[a,*], Cliff T. Ragsdale[b], Mangesh Hoskote[c]

[a] Department of Business Management, North Carolina State University, Raleigh, NC 27695, USA
[b] Department of Business Information Technology, Blacksburg, VA 24061, USA
[c] The World Bank, 1818 H Street NW, Room No. J9-027, Washington DC 20433, USA

Abstract

Many national electricity industries around the globe are being restructured from regulated monopolies to deregulated marketplaces with competitive business units. The business units responsible for transmission and distribution must be given physical property rights to certain parts of the power grid in order to provide reliable service and make effective business decisions. However, partitioning a physical power grid into economically viable districts (distribution companies) involves many considerations. We refer to this complex problem as the electrical power districting problem (EPDP). This research identifies the fundamental characteristics required to appropriately model and solve an EPDP. The proposed solution methodology is implemented as a decision support system (DSS) featuring a visualization tool that allows decision makers (DMs) to explore what we refer to as a ''soft efficient frontier.'' This DSS was found to effectively support DMs at The World Bank in solving an EPDP in the context of a case study for the Republic of Ghana.
© 2002 Elsevier Science B.V. All rights reserved.

1. Introduction

In 1990, the electricity industry in England and Wales was the first to introduce competition to the activities of power generation and supply. Initially, supply competition was made available to only 5000 large industrial consumers. However, in 1994, competition was extended to an additional 50,000 medium size consumers such as small factories and businesses [43]. In the summer of 1999, the United Kingdom (UK) implemented the first full-blown competitive electricity industry, where the full compliment of 26 million consumers was allowed to choose among suppliers of electrical power. Other countries, such as Australia, New Zealand, Bolivia, Canada, and the United States, are all at various stages of deregulation of their electricity sectors.

To facilitate deregulation of the electricity market, independent business units must be established to manage power transmission and distribution functions. In addition, these business units must be given physical property rights to certain parts of the power grid in order to provide reliable service and make

---

* Corresponding author. Department of Business Management, North Carolina State University, Raleigh, NC 27695, USA. Tel.: +1-919-515-6949; fax: +1-919-515-6943.

E-mail address: Paul_Bergey@ncsu.edu (P.K. Bergey).

effective business decisions. However, partitioning a physical power grid into economically viable districts involves many considerations. In this research, we refer to this interesting and complex problem as the electrical power districting problem (EPDP). We anticipate that many new instances of the EPDP will arise as deregulation of electricity markets takes place around the world.

Historically, researchers have shown a great deal of interest in applying management science techniques to develop political districts and align sales districts. This study is intended to synthesize the various approaches to districting problems and identify the necessary and fundamental characteristics involved in appropriately modeling an EPDP. To this end, this research has five objectives. First, to identify the issues relevant to the EPDP. Second, to investigate the similarities and differences of the EPDP with other districting problems published in the research literature. Third, to develop and recommend an appropriate solution methodology for the EPDP. Fourth, to demonstrate the effectiveness of our solution method for a specific instance of an EPDP in the Republic of Ghana. Last, to describe a decision support system (DSS) built to aid decision makers (DMs) at The World Bank in finding an acceptable solution to the EPDP in a case study of the Republic of Ghana.

1.1. Background on the electricity industry

Because of the prohibitively large start up capital required to establish a presence in the electricity industry, every electrical power system in the world was born as a natural monopoly. Electricity is simply a commodity, much like natural gas and sugar, which can be bought and sold in a free market. However, electricity has several unique features that distinguish it from other commodities. For instance, today's technology does not provide an economical means of storing bulk electricity. Thus, unlike natural gas and sugar, electricity must be produced in the right quantity at exactly the right time. In addition, electricity cannot be differentiated by its originator once it reaches the transmission and distribution grid. Finally, the flow of electricity along the grid obeys the laws of physics. Each of these natural attributes poses considerable difficulty in creating a competitive marketplace for electricity.

The actual production of electricity is a process termed "generation." Electricity can be generated in a variety of ways: hydroelectric involves the conversion of falling water into electricity, engines or turbines may be used to convert fossil or nuclear fuel into electricity, and wind or solar power may be captured and harnessed into electricity. Ultimately, the generation function involves the conversion of some form of energy into a bulk supply of electrical power.

Once electricity is generated, it is usually transported over great distances to the end users. The transportation process is referred to as the "transmission" function, which involves delivery of bulk electricity from the generators to bulk supply distribution points. Since the supply and demand for electricity must be balanced in real-time, the transmission function also involves the scheduling and dispatch of all bulk electricity for all of the generators connected to the network. At the bulk supply distribution points, the voltage of the electricity is reduced to levels that are practical for delivery to local substations. At the local substations, the voltage is once again reduced to consumption levels. The process of local delivery and voltage reduction is known as "distribution."

For most users of electricity, the functions of generation, transmission, and distribution have appeared seamless because these processes have historically been provided as a "bundled" service by the installed monopoly. Deregulation of electricity markets involves a partial "unbundling" of this vertical supply chain. Separation of these functions facilitates a competitive market for electricity via "open access" to the power grid. With open access, potentially profitable economic opportunities are available to new entrants into the electricity market. However, unbundling the supply chain also creates ambiguities regarding divisible property rights for an indivisible physical network.

For the generation function, the ability to inject electricity into the network and trade a certain amount over a period of time is required. This will allow the management of generating facilities to focus their attention upon their core activities and streamline their efforts for the efficient production of electricity. In a competitive electricity market, buyers may choose their suppliers and thus, each unit of electricity has a specified source and destination. This is referred to

as the "contract path" for the delivery of a specified amount of electricity.

The transmission and distribution functions are both basically a "wires business", which involves the coordination of delivery services across the physical network. The difficulty with unbundling these two functions is based upon transmission pricing issues. Because the actual path taken by electricity from source to destination is dependent upon current network conditions, the actual path may differ greatly from the contract path. Therefore, the true cost of delivering electricity is based upon current network conditions both at the transmission level and the distribution level. For this reason, it is likely that these services will be grouped together as deregulation takes place.

Researchers intimate with the deregulation of electricity markets have expressed concern that the process is taking place without a complete understanding of the impact of long-term decisions and that without the proper studies there is a risk of being "...locked into an inferior market design which will be costly to change" [6]. Furthermore, without proper research, there is also a risk that the full social benefits resulting from the deregulation process will not be realized [33].

1.2. Review of districting problem research

The research literature contains a variety of mathematical characterizations for the generalized districting problem as well as several suitable application areas. The application areas include, but are not limited to, political redistricting [3,5,14,18,32,35,45,46], sales territory alignment [9,17,30,42,48], and school redistricting [11]. We posit that the EPDP has similar constructs to the above applications and represents a huge open opportunity for researchers interested in districting problems to create substantial social and economic benefit. This section provides a review of the various applications of districting problems covered in the research literature, the manner in which the problem has been characterized, and the various solution techniques applied.

The most prevalent redistricting applications in the management science literature are political redistricting and sales territory alignment. The primary goal of the political redistricting problem is to provide geographically compact districts that are respectful of existing political units to the maximum extent possible. In addition, the districts must have populations with approximately equal voting potential. The motivation for using a computerized solution method is to reduce the effects of "Gerrymandering", which may occur when political incumbents bias the redistricting solution to accommodate their political agenda.

Automated political redistricting has been of great interest to politicians and researchers for the last four decades. The first mathematical characterization of the political districting problem was proposed by Vickery [46]. The redistricting problem has been characterized in the recent literature as a set-partitioning problem [3], a graph-partitioning problem [5,10], and most recently as an integer programming problem for redrawing congressional districts in the state of South Carolina [32].

Designing sales territories is another application area that can be viewed as a districting problem. The sales territory alignment problem is concerned with grouping a number of smaller geographic regions into clusters forming nonoverlapping sales territories that span a larger geographic region. Sales territories may need to be realigned whenever changing market conditions warrant, such as the introduction of a new product or variation in sales force size. Sales managers are motivated to carefully design an equitable districting plan, otherwise they risk low morale, poor performance, high turnover rate, and ultimately low productivity within the sales force. Furthermore, a balanced territorial design provides a consistent and effective basis for evaluating and comparing individual performance within the sales force.

There are a number of criteria that can be used to assign geographic regions to districts or sales people to customers. Some single alignment criteria methods seek to balance income or revenue potential [30], while other methods attempt to balance workload or effort [9]. Another common objective is to maximize overall expected profitability. The consideration of multiple objectives was first proposed by Zoltners [47].

Sales territory redistricting has a considerable history in the research literature. Similar to the political redistricting problem, a variety of solution techniques have been applied to this problem. A set-partitioning approach was investigated [42] as well as various

assignment methods [17,41,47]. For a thorough review of integer programming approaches, see Zoltners and Sinha [48]. Heuristic approaches that utilize incremental improvements have been proposed [9,16,29].

The relationship between a political district and a sales territory is fairly obvious. The political district is driven by the "one electorate–one vote" principle, which seeks to balance legislative power among districts. When assigning geographic regions to sales territories or sales people, most equitable solutions seek to balance income potential or effort. Both scenarios are similar to a bin-packing problem, where the objective is to minimize the total deviation of some criteria in each bin (district) from the ideal (global) bin mean. Typically, there are other nontrivial considerations distinguishing a bin-packing problem from a districting problem, such as the compactness of a districting plan or contiguity among the geographic regions allocated to a particular district.

1.3. The electrical power districting problem

The design of a competitive electricity market is driven by two mitigating factors. First, energy flows along a physical network according to the laws of physics, which requires a coordination of effort in order to achieve balance, reliability, and frequency control. Second, there is presently no economical way to store electricity, which means that it must be delivered in real-time on demand. Some of the major design considerations that result from the above requirements were proposed by Rassenti and Smith [37]:

(1) Coordination of the dispatch and delivery of electricity for a centralized network with decentralized suppliers and buyers.
(2) Financial instruments (futures and spot markets with bilateral contracts) which yield appropriate market signals for trading and making long-term investment decisions. For details on this issue, see Jamison [25].
(3) Defining divisible property rights for an indivisible common transmission network.
(4) Establishing pricing policies such as Zonal or Nodal pricing. For details on this issue, see Schweppe et al. [40].

(5) Facilitating competition at the local distribution level.

The EPDP has risen directly out of issue number 3 above. Between the electricity generator and the customer stretches a system of transmission lines that are interconnected to form a physical network. Once electricity is generated, it is transmitted with high voltage over large geographic regions to distribution nodes referred to as bulk supply points (BSPs). From the BSPs, the voltage is reduced and the electricity is distributed to local user groups such as residential neighborhoods, industrial users, or commercial developments. Due to the large costs of equipment and maintenance, transmission and distribution are likely to remain more economical if a large proportion of the power grid is maintained by a single entity in a given geographic area. Thus, the available customer base, which can be represented by BSPs, must be grouped into districts that can be effectively managed in a competitive market.

Similar to political redistricting and sales territory alignment problems, the EPDP is primarily concerned with creating groups of approximately equal revenue or profit potential. The motivation for this objective is to foster competition by attracting private investment. A good districting plan must also consider the compactness and contiguity of BSPs that comprise a district. Districts that are compact over a geographic region rather than disbursed will be more economical to maintain and thus more profitable. The requirement for contiguity is rather intuitive since product delivery takes place over physical wires connecting the BSPs. From a practical point of view, it is desirable to deliver electricity between BSPs within a common district without paying rents to another transmission enterprise. However, while contiguity appears to be a simple and intuitive requirement, there are a number of hidden complexities that arise when actual power flow is considered.

Perhaps the most difficult and sensitive issue regarding competitive design of the electricity market is establishing a pricing policy. There are two pricing configurations presently being explored in the research literature and implemented in emerging markets throughout the world: nodal pricing and zonal pricing [7,40]. Nodal pricing schemes establish a unit charge for electricity demanded at each BSP. Under

this scenario, unit prices of electricity may be based upon historical consumption and delivery costs. Another form of nodal pricing is locational marginal pricing (LMP). LMP is not based upon historical data but driven by the concepts of economic efficiency. The curious reader is referred to Schweppe et al. [40] and Jamison [25] for a detailed description of LMP methods. The zonal approach aggregates a number of BSPs into larger zones under the assumption that this would reduce the complexity of the pricing issue. In this case, a number of BSPs share a common unit price.

In our research, we have implemented a nodal pricing scheme based upon historical data provided by The World Bank. The research literature suggests that nodal pricing accounts for "loop-flow" phenomenon more accurately than zonal pricing [19,20]. Furthermore, because the nodal pricing that we implement is based upon historical information, it is less ambiguous and therefore more appealing to the DMs at The World Bank.

2. Mathematical framework

A districting plan is a partitioning of units (populations, sales regions, BSPs) into nonoverlapping districts (groups) that the are contiguous (adjacent) and geographically compact. A desirable districting plan optimizes one or more objectives such as balanced legislative power or revenue potential. By associating each BSP with a node, and each long distance transmission line between BSPs with an edge, the EPDP can be modeled as a graph-partitioning problem. This method was applied to the political districting problem where the weight on the node was equal to the corresponding population size [3,32].

Let us define $G (N, E)$ to be a graph G with nodes N defined to be the set of BSPs that comprise G, and edges E defined to be a pair-wise connection matrix corresponding to the set of long distance transmission lines that connect N. A *district* is a node induced subgraph $G' (N', E')$ of $G (N, E)$ that is contiguous. A district G' is contiguous if all nodes N' assigned to the district are connected by edges E'. Contiguity also implies that it is possible to reach any node in a particular district from any other node assigned to the same district without leaving the
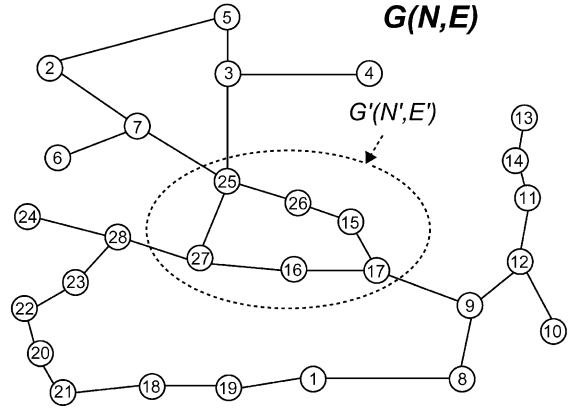

Fig. 1. A graph representation of an electrical transmission and distribution network.

district. A district is also referred to as a partition or a group belonging to $G (N, E)$ in later sections. See Fig. 1.

2.1. Balanced revenue objective

As mentioned earlier, one objective DMs may wish to pursue in solving an EPDP involves creating a districting plan with approximately equal revenue or profit potential in each district. A districting plan where n nodes (bulk supply points) are assigned to K districts can be described by the solution vector $x_{ik} \in \{0,1\}$ where $x_{ik} = 1$ if node i is assigned to district $k \in \{1,...,K\}$ and equal to zero otherwise. To obtain K transmission districts of approximately equal earning potential, we can minimize the total deviation of the revenue in each district from a target value. This can be modeled as follows:

$$\text{Minimize}: \quad z_1 = \sum_{k=1}^{K}(u_k + v_k) \tag{1}$$

$$\text{where}: \quad \bar{r} = \frac{1}{K}\sum_{i=1}^{n} r_i \tag{2}$$

$$R_k = \sum_{i=1}^{n} r_i x_{ik}, \quad \forall k = 1 \ldots K \tag{3}$$

$$R_k - \bar{r} = u_k - v_k, \qquad \forall k = 1 \ldots K \qquad (4)$$

$$0 \leq u_k, v_k \leq \delta \bar{r}, \qquad \forall k = 1 \ldots K \qquad (5)$$

$x_{ik} = 1$ if node i is assigned to district k, and 0 otherwise; R_k = sum total of revenue potential in district j; K = number of districts (partitions); r_i = amount of revenue potential contained in node i; n = number of nodes in the network; 100δ, $(0 \leq \delta \leq 1)$, is the maximum allowable percentage deviation of the actual revenue in a district from the target.

Note that the above formulation does not explicitly model the difficult task of ensuring contiguity within each district. Rather, our definition of a district (given above) *requires* that districts be contiguous. So the above model assumes that nodes assigned to a particular district are contiguous. A later section (and Appendix C) addresses how our solution methodology maintains contiguity in each district.

2.2. Compactness objective

Another objective (discussed earlier) of interest to DMs solving EPDPs is the geographical compactness of a districting plan. To measure the compactness of a district k, we use the total Euclidean distance from the centroid of district k to each node assigned to district k. The districting plan that minimizes the total compactness is then found by solving the following problem:

$$\text{Minimize}: \quad z_2 = \sum_{k=1}^{K} D_k \qquad (6)$$

where :

$$D_k = \sum_{i=1}^{n} x_{ik}$$
$$\cdot \sqrt{(C_x(k) - N_x(i))^2 + (C_y(k) - N_y(i))^2},$$
$$k = 1 \ldots K \qquad (7)$$

$$C_x(k) = \frac{1}{m_k} \sum_{i=1}^{n} x_{ik} \cdot N_x(i), \qquad k = 1 \ldots K \qquad (8)$$

$$C_y(k) = \frac{1}{m_k} \sum_{i=1}^{n} x_{ik} \cdot N_y(i), \qquad k = 1 \ldots K \qquad (9)$$

$$m_k = \sum_{i=1}^{n} x_{ik}, \qquad k = 1 \ldots K \qquad (10)$$

D_k = compactness for district k; K = number of districts (partitions); $C_x(k)$ = x—coordinate for the Centroid of district k; $C_y(k)$ = y—coordinate for the Centroid of district k; $N_x(i)$ = x—coordinate for node i; $N_y(i)$ = y—coordinate for node i; n = number of nodes in the network; m_k = number of nodes in district k.

2.3. Multi-criteria decision making

Because single criteria optimization methods often fail to adequately model the complexity of problems faced in today's rapidly changing business environment, more and more DMs are interested in using decision aids that support multi-criteria decision making (MCDM).

The mathematical foundation for multi-criteria decision making (MCDM) was developed over a century ago [36]. A typical characteristic of a MCDM problem is the absence of a unique global optimum. Rather, multiple solutions to the problem often exist that are superior to (dominate) the others in the solution space.

When solving a MCDM problem, it is generally assumed that DMs prefer or desire to obtain Pareto optimal (or nondominated) solutions. For the criterion vector $F(x) = (f_1(x), f_2(x), \ldots, f_q(x))$, a solution vector $x^* \in R^n$ is said to *dominate* $x \in R^n$ if $f_i(x^*)$ is at least as good as $f_i(x)$ for all i and $f_i(x^*)$ is better than $f_i(x)$ for at least one i. A solution vector $x^* \in R^n$ is *Pareto optimal* if there exists no other solution vector $x \in R^n$ that dominates x^*.

From a practical perspective, when using population-based heuristic search techniques, it is often difficult or impossible to know if another (yet to be observed) solution *exists* that dominates a particular known observed solution. As a result, in this work, we relax the definition of Pareto optimality to refer to known (observed) solution vectors rather than to all possible solution vectors that theoretically exist in R^n. That is, for a given population of solution vectors, we say that x^* is (currently) non-

dominated if there exists no other solution x in the population that dominates x^*. Similar relaxed definitions of Pareto optimality have been used [12,13,15,23,24,39,44].

2.4. The multi-criteria EPDP model

In this study, $f_1(x)$ is defined by Eq. (1) as the total revenue deviation of a districting plan and $f_2(x)$ is defined by Eq. (6) as the total Euclidean distance (compactness) of a districting plan. We propose the following model for simultaneously minimizing the $q = 2$ components of criterion vector $F(x)$ for the EPDP.

$$\text{Minimize}: \quad F(x) = (f_1(x), f_2(x)) \tag{11}$$

subject to equations Eqs. (1)–(10).

2.5. Solving the model

As mentioned earlier, one of the most difficult aspects of solving an EPDP is the issue of maintaining contiguity as various districting plans are created. To address this issue, we developed a custom genetic algorithm (GA) designed for solving EPDPs that automatically maintains contiguity in all of the districting plans it generates. The details regarding the design and operation of this GA are not central to the theme of this paper. However, for the curious reader, the custom crossover and mutation operators used in our GA are described in detail in Appendix D.

This GA was implemented as the search engine for our DSS using the Visual Basic for Applications (VBA) programming language in Microsoft Excel. The ubiquity of Excel, combined with the power of its VBA programming language and inherent data management, analysis, and visualization tools, makes it an ideal platform for creating a DSS that is easy for DMs at The World Bank to use and distribute. See Fig. 2.

To use our DSS, the DM simply selects the functions to be optimized and specifies the exogenous variables such as the maximum allowable percentage revenue deviation δ in Eq. (5), the number of desired districts (K) and the number of desired solutions (population size). In addition, the DM has the ability to control the optimization run time parameters by choosing from the available termination criteria: a

time limit (in minutes), a computational effort limit (in function evaluations), or a performance limit (in function improvement with respect to cumulative run time). The DSS then generates and displays a series of Pareto optimal solutions for the DM's consideration and exploration. For a thorough description of the computational performance aspects of our custom GA, the reader is referred to Ref. [4].

3. The Ghana case study

The Republic of Ghana, situated on the Gold Coast of Africa, is one of the most developed African countries south of the Sahara. Like all existing electricity infrastructures in the world today, the Ghana power sector began as a noncompetitive national monopoly. With the objective of developing an economically stable, customer-oriented industrial culture in Ghana, the Government of Ghana (GoG) is undertaking an industry reform program for their emerging competitive electricity market. The GoG has enlisted the services of The World Bank in hopes of bringing their objectives to fruition. To this end, the authors engaged in a joint research effort with The World Bank to develop a solution methodology and DSS to assist with restructuring the electric power sector in Ghana, with the intentions of applying this methodology to other countries. Specifically, the focus of the DSS is to assist DMs with developing a restructuring plan comprised of districts (groupings of BSPs) such that independent distribution enterprises may operate as reliable and economically viable electricity service distributors.

Under the present system, there are two monopoly organizations installed as the national service providers of electricity in the Republic of Ghana, the Electricity Corporation of Ghana (ECG) and the Northern Electricity Department (NED). Together, these organizations share the nine geographic regions that span Ghana. The ECG is the bigger and older of the two with five regions, while the remaining four are cared for by the NED.

The national electricity transmission grid of Ghana is provided in Fig. 3. The grid consists of high voltage transmission connections and 28 BSPs. Recall that BSPs represent bulk distribution junctions points where high voltage electricity is reduced for local
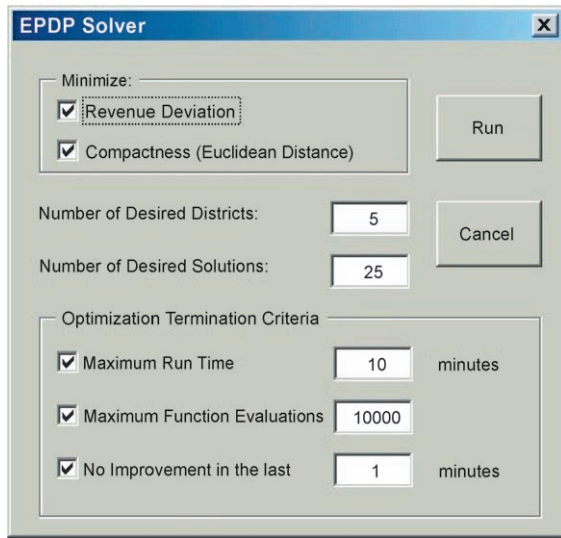
Fig. 2. EPDP Solver (custom genetic algorithm).

distribution services. Fig. 4 provides a legend of the BSPs included in this study. The information consists of the index number and name corresponding to each

BSP shown in Fig. 3. Also, the expected annual revenue potential is provided for each BSP in millions of dollars along with the standardized physical coordinates for each BSP on a two dimensional map.

3.1. A decision support system

According to Marakas [31,p,4], the role of a DSS "…is to provide support to the DM on the *structurable* portions of the decision, thus freeing the DM to focus his or her cognitive resources on the truly *unstructurable* portions of the problem—those portions that, given the limits of technology to execute the complex problem-solving strategies contained in human memory, are better left for resolution by the human DM(s)." As a result, generating a set of Pareto optimal solutions for a DM is only the first step to solving an EPDP. Evaluating preferences among power district configurations involves a great deal of (structurable) numerical analysis as well as (unstructurable) subjective assessment of each plan. While a DSS may be capable of locating a variety of non-dominated solutions automatically, the DM will often
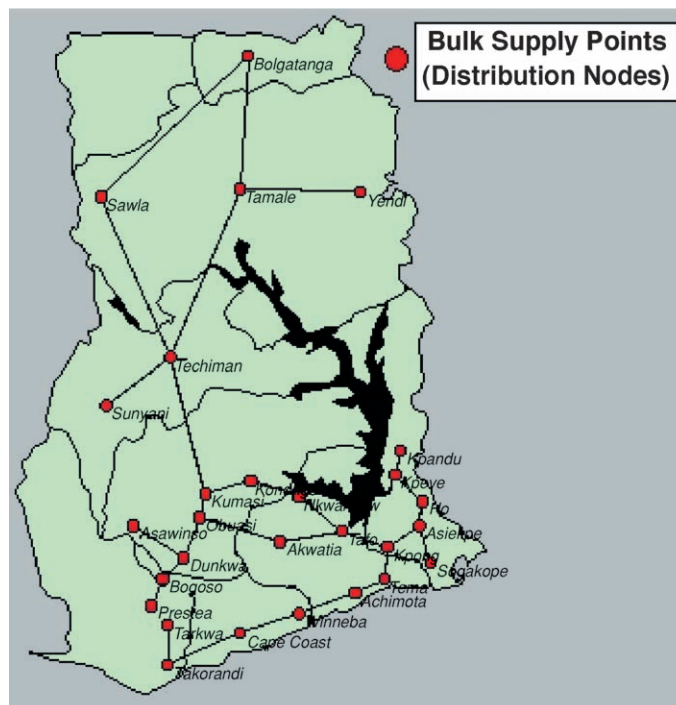

Fig. 3. The national electricity power grid of Ghana.

Index	BSP	Exp Rev($M)	X	Y
1	Achimota	$ 30.518	21	8
2	Sawla	$ 0.006	6	34.5
3	Tamale	$ 1.943	17	35
4	Yendi	$ 0.007	23	35
5	Bolgatanga	$ 0.873	17	45
6	Sunyani	$ 1.978	6	20.5
7	Techiman	$ 1.127	8.5	23
8	Tema	$ 12.369	23	9
9	Kpong	$ 2.511	23	12
10	Sogakope	$ 0.294	27	11.5
11	Ho	$ 0.403	26	15
12	Asiekpe	$ 0.010	26.5	13
13	Kpandu	$ 0.616	25	18
14	Kpeve	$ 0.208	25	16
15	Nkwankaw	$ 1.585	7.5	20
16	Akwatia	$ 0.571	17	11
17	Tafo	$ 0.985	19.3	13
18	Cape Coast	$ 1.654	14	5
19	Winneba	$ 1.053	18	6.5
20	Tarkwa	$ 3.104	9	6
21	Takorandi	$ 5.112	10	3
22	Prestea	$ 1.610	8	7
23	Bogoso	$ 1.479	8.5	8
24	Asawinso	$ 0.623	6.5	13
25	Kumasi	$ 11.284	11	16
26	Konongo	$ 0.434	14	15
27	Obuasi	$ 1.111	11	12.5
28	Dunkwa	$ 0.163	10	11.5

Fig. 4. Bulk supply point information.

prefer to make small perturbations to an efficient solution, possibly causing the preferred solution to be sub-optimal (in a Pareto sense) or even violate one or more constraints. This is particularly likely to occur if a decision involves more than one DM, which is typically the case for an EPDP.

For the Ghana EPDP, the DMs at The World Bank considered the following criteria in evaluating the attractiveness of alternative districting plans:

Revenue Potential
Spread of Geographical Area
Customer Spread
Clarity of Demarcation

The primary criterion in the evaluation of a districting plan is total absolute revenue deviation (Eq. (1)), which is more easily quantifiable than some of the others. The geographic and customer spread are modeled in our procedure using total Euclidean distance as the measure of compactness (Eq. (6)). Finally, and perhaps the most difficult, is the criterion described as "Clarity of Demarcation." By this term, the DMs at The World Bank are referring to the allocation of districts with respect to existing political units, proximity to natural geographic boundaries such as lakes and rivers, and consistency with other economic influences. The nature of this criterion makes it difficult to model accurately (i.e. it is unstructurable). As a result, solution vectors that are on the Pareto optimal "efficient frontier" formed by the explicitly modeled (i.e. structurable) criteria may be dominated or infeasible when considering the DM's utility function for all the criteria impacting the problem. It is this unstructurable component of the decision problem that motivates us to develop a DSS to solve the EPDP.

We believe that the true underlying frontier (adjusted for the utility of the DMs) is best exposed by considering alternative solutions near the projected or interpolated efficient frontier. Further, we submit that a DSS that allows DMs to explore this "soft efficient frontier" of solution vectors will be most effective in modeling MCDM problems.

3.2. A DSS for the EPDP

Traditionally, MCDM problems tend to group members of the solution set into two primary categories, the nondominated (optimal) set and the dominated (sub-optimal) set. The Pareto ranking technique offers a further distinction between the members of the dominated set. It is a technique which consists of ranking the members in the dominated set according to a modified definition of Pareto dominance. See Fig. 5. First, the nondominated solutions are given a rank of one and then removed from the solution set. In the absence of the nondominated set, another layer of solutions is "exposed", which become nondominated (by the remaining solutions). The subset of "exposed" solutions is given a rank of two and then removed from the solution set. In their absence, another layer of solutions is "exposed", which are given a
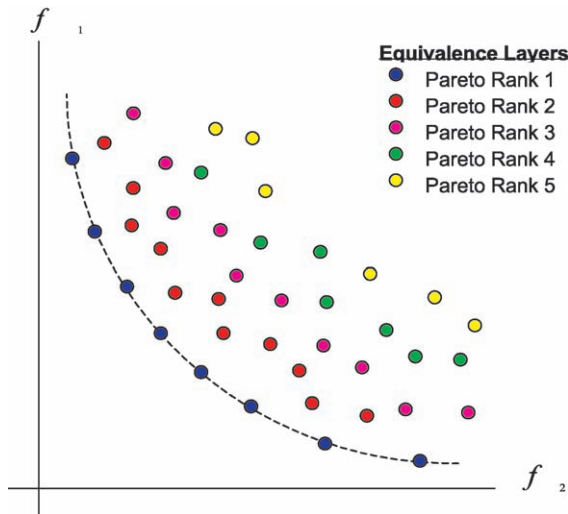
Fig. 5. Pareto ranking the dominated subset.

rank of three. The process is continued until all of the solutions in the solution set have been ranked. The different layers of nondominated solutions identified by the Pareto ranking technique are also referred to as equivalence class layers. The term is fitting because within each layer, none of the solutions dominate any of the other solutions in the same layer, thus they are equivalent. Furthermore, between layers of solutions, there exists at least one solution in the lower equivalence class, which dominates any solution in a higher equivalence class. In this regard, the layers are distinct.

Traditionally, MCDM problems represent objective function values for a set of solution vectors as a list or a table. Our DSS renders the solution set using a visualization tool, which we refer to as a Pareto rank scatter plot (PRSP). The PRSP is useful in helping the DMs judge alternative districting plans
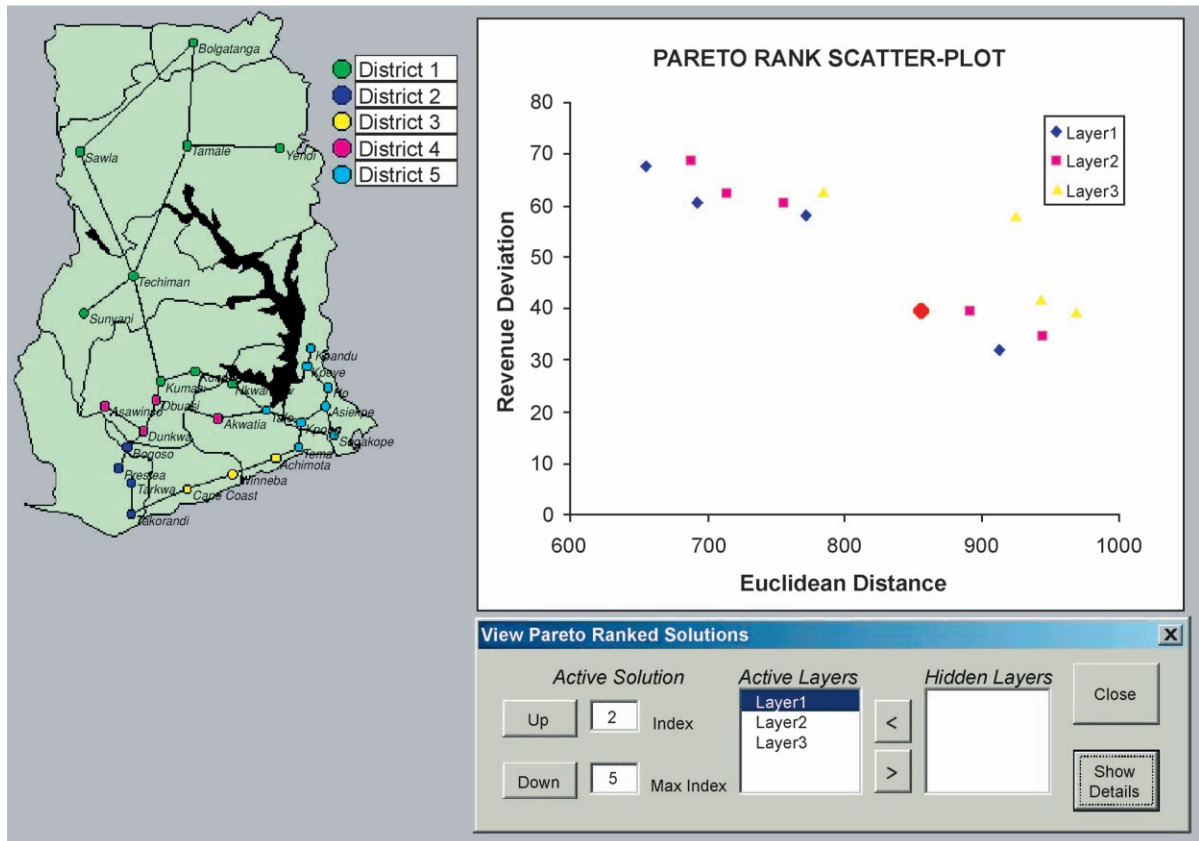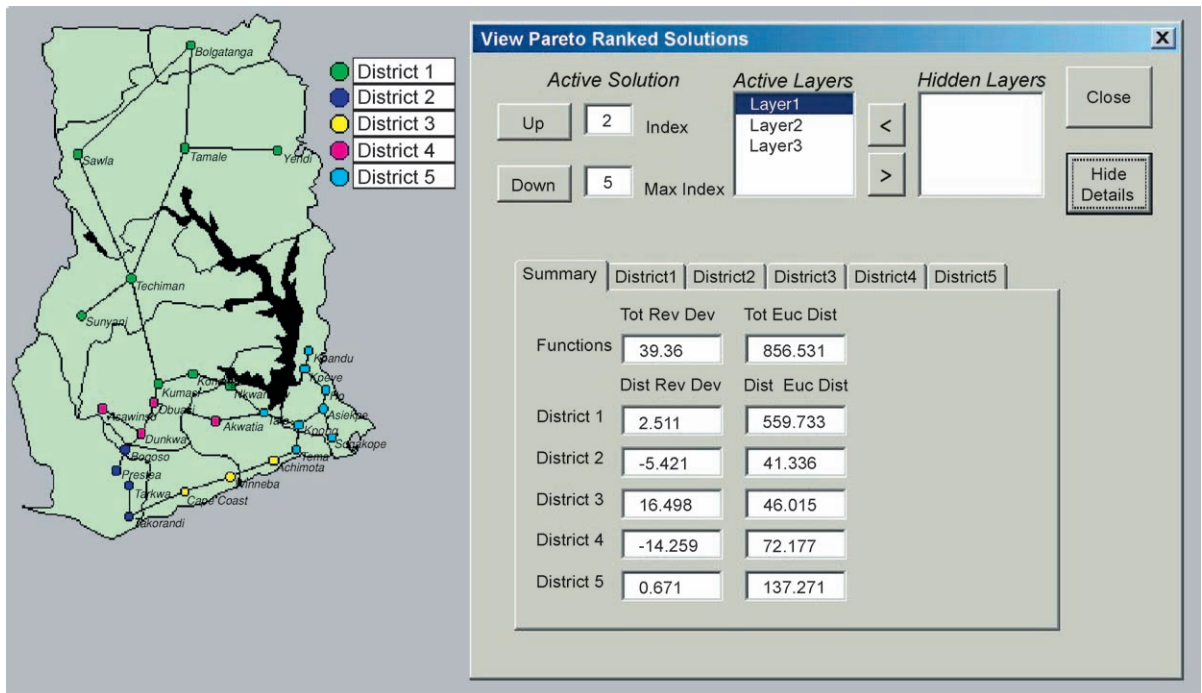

Fig. 6. Layer 1—Solution index 2.

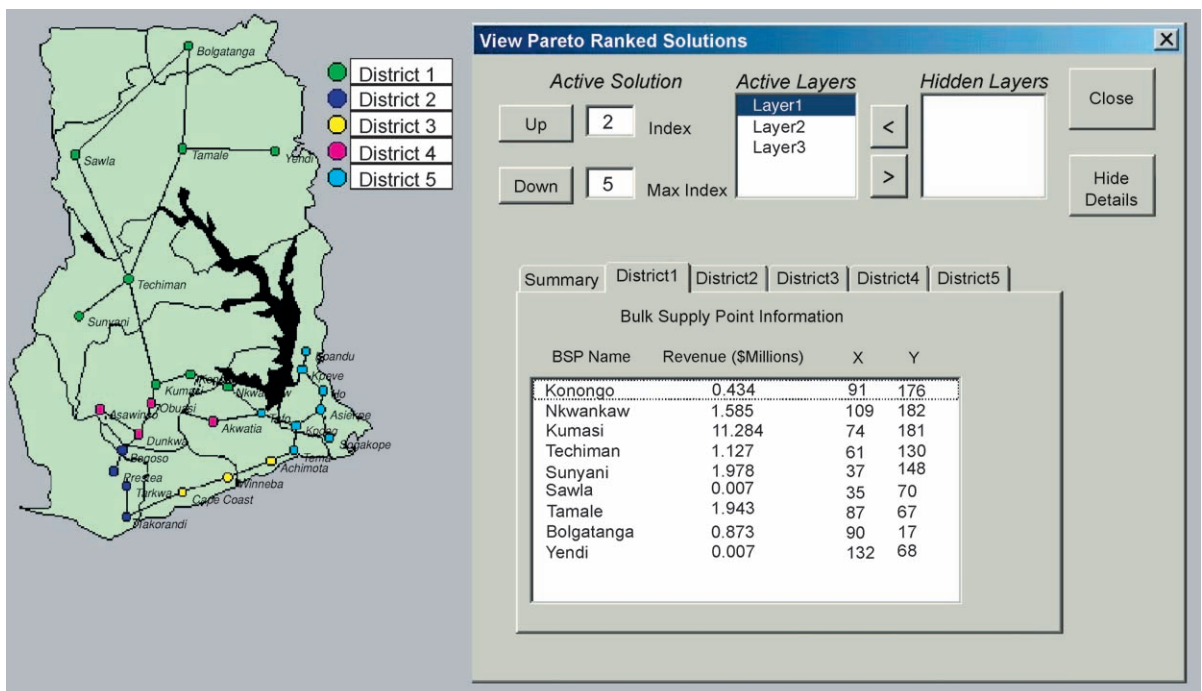Fig. 7. Criterion data aggregated at the district level.

Fig. 8. Layer 1—Solution index 1—District 1.

relative to others in the solution space. The PRSP organizes the solutions into a "soft efficient frontier" comprised of equivalence class layers. Each solution in an equivalence class layer is displayed with a marker corresponding to the legend in the PRSP. Our DSS also renders a map of Ghana showing a visual representation of any specific districting plan. The PRSP synchronizes the active solution (districting plan) shown in the map with all of the alternative solutions in the set shown in the PRSP. The active solution is indicated by a large red dot in the PRSP regardless of which layer it belongs. The objective of this tool is to visually coordinate the two different levels of data aggregation (i.e. view one specific solution vs. view the set of all solutions). The DM may navigate through the "soft frontier" using the form shown directly below the PRSP in Fig. 6 to compare alternatives and consider the details of various trade-offs.

The DM can also use the controls to hide and/or display any of the equivalence class layers if the solution space becomes congested. The details of any plan can be viewed by clicking on the "Show Details" button. As shown in Fig. 7, this causes the form to expand to provide information that is aggregated at the district level for whatever solution is currently displayed on the map.

The information contained on the summary page allows the DM to determine the contribution of each individual district to the overall value of the objective functions. For example, the top two text boxes (labeled "Functions") correspond to the objective function values of the districting plan currently rendered on the map. Recall that the revenue deviation function is the sum of absolute values, however, the values corresponding to each district are provided in nonabsolute terms. This allows the DM to determine if a given district is above or below the ideal target value for the overall districting plan. A negative value indicates the extent to which the district is below the target value while a positive value indicates the extent to which the district is above the target value. For the total Euclidean distance objective function, the corresponding district values indicate the sum of Euclidean distances (compactness) for each district. Fig. 8 shows the detail information of the BSPs aggregated at the district level for District 1.

4. Conclusions

With competitive electricity markets emerging throughout the world, it is necessary to develop effective solution methods (systems and algorithms) for designing power districts. This research represents an initial attempt to characterize the EPDP. Specifically, we identified similarities and differences of EPDPs with other districting problems, and developed a DSS that is effective in finding acceptable solutions. The motivation underlying this research was to develop a useful decision support tool for DMs at The World Bank, with the intent of applying the tool to other countries as deregulation takes place around the world.

In order to enact a reform policy designed to introduce competition into the electric utility industry, each nation will need to address the EPDP for their network in some manner. It is significant to note that developing nations can often experiment with more innovative approaches than are being considered in developed countries due to fewer existing barriers to reform. The opportunity for researchers to make a real impact in these nations is clearly at hand.

We believe that one of the reasons the solution method presented in this research is effective is because it is independent of the utility function the DM brings to the problem. When the Pareto set can be easily visualized (using the DSS), the DM is able to choose a final alternative (from the GA population) without the rigor and uncertainty of utility assessment. Furthermore, because the GA has thoroughly explored the search space, the DM is able to easily manipulate efficient districting plans and move forward with a high level of confidence that the important trade-offs relevant to the decision have been considered.

Acknowledgements

Appendix A. Biologically inspired algorithms

Evolutionary algorithms (EAs) represent a powerful, general purpose optimization paradigm where the

computational process mimics Darwin's theory of biological evolution [21,27,38]. In a nutshell, most EAs start with a set of chromosomes (numeric vectors) representing possible solutions to a problem. The individual components (numeric values) within a chromosome are referred to as genes. New chromosomes are created by crossover (the probabilistic exchange of values between vectors) or mutation (the random alteration of values within a vector). Chromosomes are then evaluated according to a fitness (or objective) function with the fittest surviving into the next generation. The result is a gene pool that evolves over time via these genetic operators to produce better and better solutions to a problem.

In order to solve partitioning problems effectively, it is necessary to design operators specifically for the task. This appendix is intended to describe in detail the operators that have been tailored specifically for solving the EPDP.

Appendix B. Toward a DNA based evolutionary algorithm

In this appendix, it is our objective to provide rationale for introducing analogies that link the DNA structure to the general graph model. We do not pretend that our presentation of the biological composition of DNA is complete. For a comprehensive review of the biology of DNA the reader is referred to Alberts et al. [2].

Science has revealed that DNA works the same in all forms of life. Without DNA life would not exist—not plants, nor animals, not even bacteria. All living tissue in any life form is comprised of DNA. This molecule which is so fundamental to life is composed of a very simple four-code alphabet referred to as DNA nucleotides (bases). They are adenine (A), thymine (T), cytosine (C), and guanine (G). Together, these four bases create all forms of life depending upon the manner in which they are bonded and sequenced. Surprisingly, the rules, which they follow to accomplish this seemingly miraculous task, are rather simple. There are only four possible conditions in which the DNA bases can bond: A–T, T–A, C–G, and G–C. Bonded DNA bases are referred to as base pairs which are then linked together (sequenced) to form a DNA strand. DNA strands are extremely long and herein lies the complexity of living organisms. The human description has been estimated to be approximately 3 billion base pairs long.

The process by which new sequences in the DNA strand are explored is referred to as DNA replication. An interesting aspect of the recombination process that occurs in all sexually reproducing organisms is that it too follows the simple bonding rules, which allows the DNA to replicate. Thus, while new gene sequences are evolved over time, the fact remains that A–T, T–A, C–G, and G–C still holds before and after the DNA strand has been reproduced in the new host organism.

Fig. 9 provides a visual example of the replication process in DNA. The process begins with two parents contributing a DNA strand comprised of base pairs. DNA division is the process of separating the base pairs in the parent strands from each other. One half of the divided strand from each parent retains their sequential alignment, while the other half breaks the DNA bases apart completely. The separated DNA bases temporarily float in a "soup" while awaiting to reassemble. The contiguous divided strands contributed by each parent is then randomly segmented into smaller sequentially aligned pieces and then recombined to form a new child strand. It is significant to note that each half of a DNA strand contains all of the required information to reconstruct the original host organism. Thus, when the new child strand reassembles the base pairs in the DNA soup, each segment attempts to reconstruct its original parent DNA strand (host organism). The crossover operator that we implement in our GA mimics the DNA replication process described above.

Appendix C. Enforcing contiguity with DNA bonds

In most GA implementations, constraints are enforced via penalty functions. When the size of the constraint set is large, the penalty function method often results in the production of a substantial number of infeasible chromosomes. Thus, a great deal of computational effort is wasted in the process. Enforcing contiguity in a districting plan requires an expo-
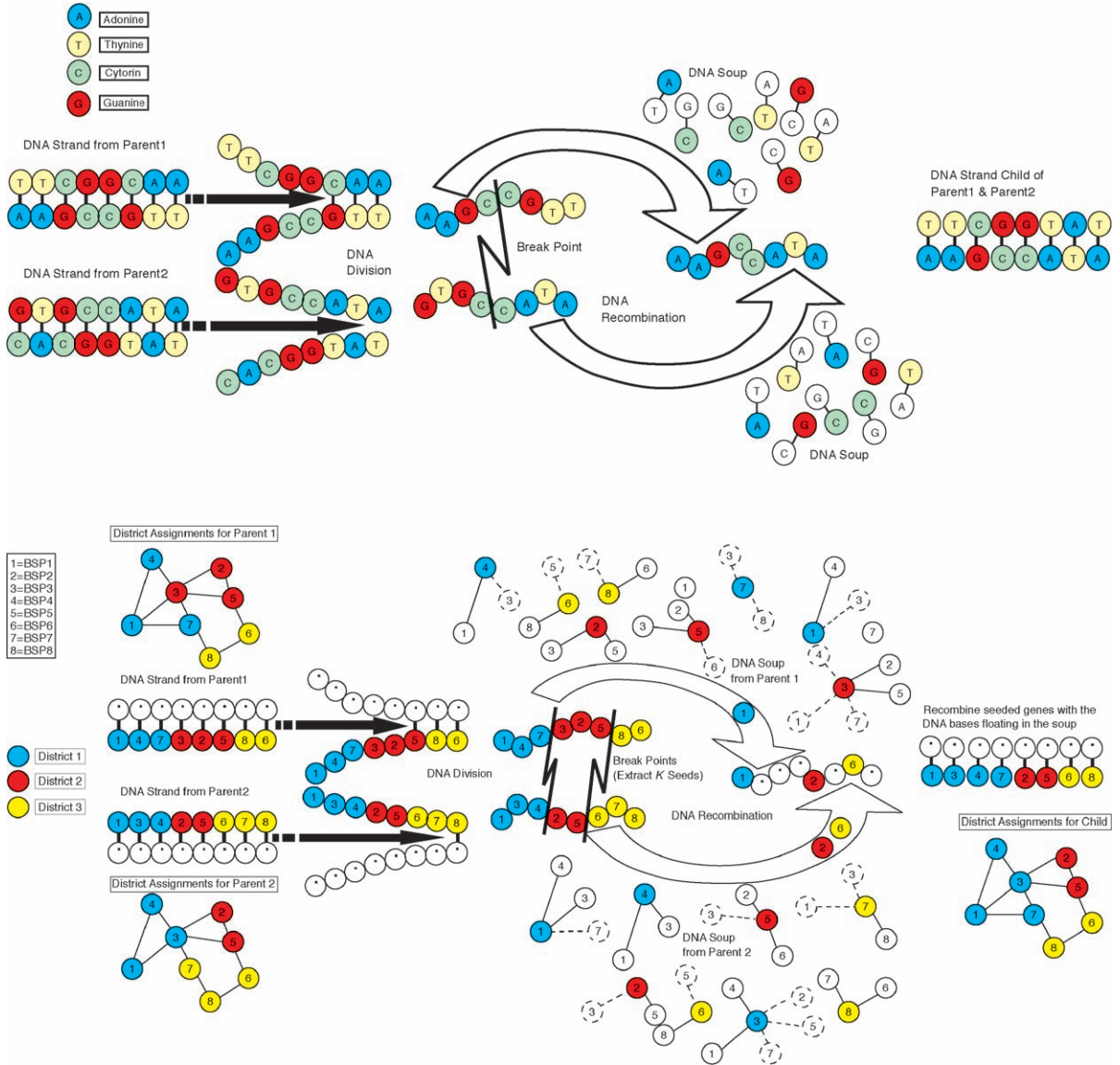
Fig. 9. (top) The DNA replication process.
Fig. 11. (bottom) Enforcing contiguity during DNA replication.

nential number of constraints in the model with respect to the number of nodes in the graph [32]. An alternative to including the constraint set in the model formulation would be to enforce contiguity strictly in the search algorithm's problem representation.

In our DNA object model, we create an additional layer of abstraction beyond the traditional GA model

where a collection of DNA bases reside within a gene object. In the model, a DNA base is analogous to a specific node in the graph, while a gene is analogous to a node induced subgraph—a district. It is the unique collection of DNA bases that gives the gene (district) a set of unique properties or characteristics. A base pair is analogous to the set of links belonging to a node in

the graph. The links connecting the nodes in the graph are used as a surrogate for the rules, which allow DNA bases to bond and sequence into a strand. In our algorithm, DNA bases are limited to bonding with their neighbors during the replication process, which results in only feasible solutions to highly constrained graphical network problems without the use of inefficient penalty functions. Fig. 10 depicts the object model hierarchical relationships.

Fig. 11 provides a visual representation of the DNA replication process in our GA. As described above, the process begins with two parents contributing complete DNA strands (chromosomes). Each DNA strand represents an assignment of DNA bases to a gene, which corresponds to an assignment of BSPs to a district. Parent 1 has the following assignments: District 1={BSP1, BSP4, BSP7}, District 2={BSP2, BSP3, BSP5}, District 3={BSP6, BSP8}. Parent 2 has the following assignments: District 1={BSP1, BSP3, BSP4}, District 2={BSP2, BSP5}, District 3={BSP6, BSP7, BSP8}. DNA division in our algorithm is represented by the instantiation of

new DNA bases in the DNA soup containing identical information to their parent strands. Thus, each DNA base is aware of its current district assignment, its connected neighbors (solid line) and its unconnected neighbors (dotted line).

During recombination, a new child chromosome containing K gene objects is instantiated. Each gene in the child strand is randomly seeded with a DNA base from the available DNA objects in the DNA soup. The assignment of the DNA base to a gene in the child strand corresponds to its district assignment in the parent gene. Once K unique seeds have been randomly selected and assigned, the recombination process allows each seed to "attempt" to reconstruct its original parent gene. When a DNA base is assigned, it is removed from the DNA soup of both parent strands. Since the process is random, the complete set of DNA bases may not be available to reconstruct the parent gene entirely. DNA bases that become "stranded" due to the randomness of assignments are placed in a temporary location until all possible DNA nucleotides have been assigned based upon the seed selection.
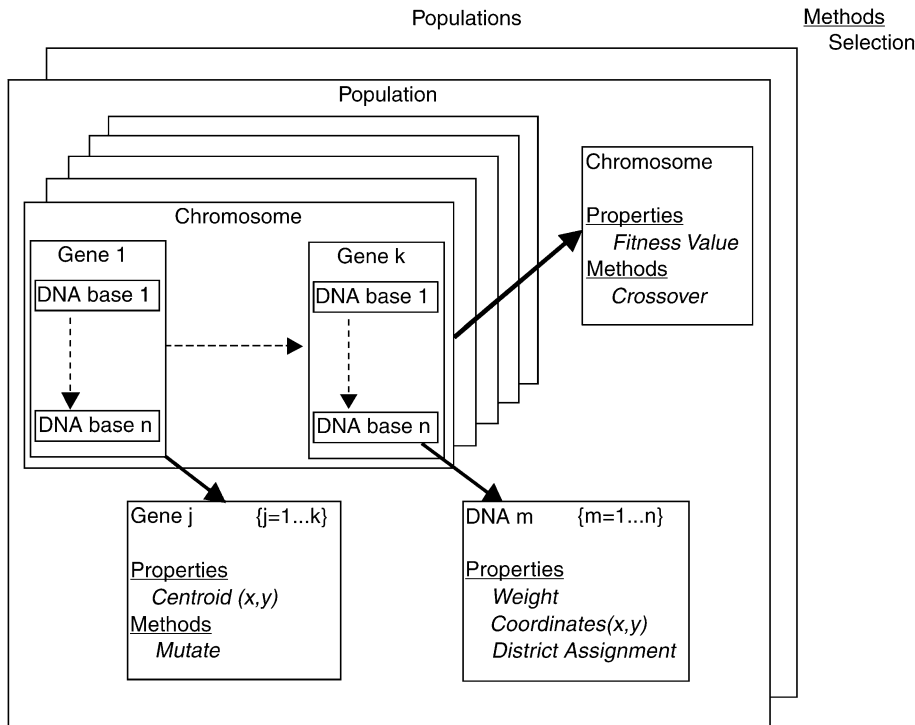

Fig. 10. DNA object model.

The remaining nodes in the temporary location are then randomly assigned back to a feasible gene (not necessarily the original parent gene).

Note that the "Break Points" for each parent strand in Fig. 11 are located between distinct genes, where the genes can contain a variable number of DNA bases. The recombination step in this example takes place as follows: BSP1 is selected from parent 1 as the seed for district 1. BSP2 is selected from parent 2 as the seed for district 2. BSP6 is selected from parent 2 as the seed for district 3. Each gene in the child strand is now ready to "attempt" to reconstruct back to its original configuration. BSP1 in gene 1 examines the soup from parent 1 and adds the available neighbors BSP4 and BSP7. BSP2 in gene 2 examines the soup from parent 2 and adds the available neighbor BSP5. BSP6 in gene 3 examines the soup from parent 2 and adds the available neighbor BSP8. The random process has stranded BSP3 in the DNA soup because all available neighbors have been assigned to completely reconstructed districts. Recall that when a DNA base is assigned, it is removed from the DNA soup of both parents. BSP3 is placed in a temporary location and then randomly assigned to a feasible district—in this case District 1. The new child strand of DNA consists of the following assignments: District 1={BSP1, BSP3, BSP4, BSP7}, District 2={BSP2, BSP5}, District 3={BSP6, BSP8}. Note that it strongly resembles each parent, yet differs slightly due to random recombination.

Appendix D. The mutation operator

The mutation operator uses a series of neighborhood moves that exchanges a single node with a neighboring district. Thus, the mutation operator is limited to nodes that exist on the boundary between two partitions. It can be viewed as a hill-climbing operator that is designed to optimize the DNA strand of the child chromosome by enumerating all feasible solutions that are exactly one move away from the current solution and only the best solution is retained.

References

[2] B. Alberts, R. Lewis, W. Roberts, Molecular Biology of the Cell, 3rd ed., Garland Publishing, 1994, ISBN 0-8153-1619-4.

[3] M. Altman, Is automation the answer?—The computational complexity of automated redistricting, Rutgers Computer & Technology Law Journal 23 (1) (1995) 81–142.

[4] P.K. Bergey, C.T. Ragsdale, M. Hoskote, A Mimetic Algorithm for the Electrical Power Districting Problem. Working Paper, 2002.

[5] M.H. Browdy, Simulated annealing: an improved computer model for political redistricting, Yale Law & Policy Review 8 (163) (1990).

[6] H. Chao, H.G. Huntington, Introduction: economic and technological principles in designing power markets, in: H. Chao, H.G. Huntington (Eds.), Designing Competitive Electricity Markets, Kluwer Academic Publishing, 1998, pp. 1–10.

[7] H. Chao, S. Peck, A market mechanism for electric power transmission, Journal of Regulatory Economics 10 (1) (1996) 25–29.

[9] C. Easingwood, Heuristic approach to selecting sales regions and territories, Operations Research Quarterly 24 (4) (1973) 527–534.

[10] E. El-Farzi, G. Mitra, Solution of set-covering and set-partitioning problems using assignment relaxations, Journal of Operations Research Society 43 (483) (1992).

[11] J.A. Ferland, G. Guenette, Decision support system for the school districting problem, Operations Research 38 (1) (1990) 15–21.

[12] C.M. Fonseca, P.J. Flemming, An overview of evolutionary algorithms in multi-objective optimization, Evolutionary Computation 3 (1) (1995) 1–16.

[13] C.M. Fonseca, P.J. Flemming, Multi-objective optimization and multiple constraint handling with evolutionary algorithms: Part I. A unified formulation, IEEE Transactions on Systems, Man and Cybernetics. Part A, Systems and Humans 28 (1) (1998) 26–37.

[14] R.S. Garfinkel, G.L. Nemhauser, Optimal political districting by implicit enumeration techniques, Management Science 16 (8) (1970) B495–B508.

[15] D.E. Goldberg, Genetic Algorithms in Search, Optimization and Machine Learning, Reading, MA Addison-Wesley, 1989.

[16] M.S. Heschel, Effective sales territory development, Journal of Marketing 41 (2) (1977) 39–43.

[17] S.W. Hess, S.A. Samuels, Experiences with a sales districting model: criteria and implementation, Management Science 18 (4) (1971) 41–54.

[18] S.W. Hess, J.B. Weaver, H.J. Siegreldt, J.N. Whelan, P.A. Zitlau, A nonpartisan political redistricting by computer, Operations Research 13 (1965) 998–1008.

[19] W.W. Hogan, A market power model with strategic interaction in electricity networks, Energy Journal 18 (4) (1997) 107–141.

[20] W.W. Hogan, Nodes and zones in electricity markets: seeking simplified congestion pricing, in: H. Chao, H.G. Huntington (Eds.), Designing Competitive Electricity Markets, Kluwer Academic Publishing, 1998, pp. 33–62.

[21] J.H. Holland, Adaptation in Natural and Artificial Systems, University of Michigan Press, Ann Arbor, 1975, MI Reissued by MIT Press, Cambridge, MA.

[23] J. Horn, N. Nafpliotis, Multi-objective optimization using the niched pareto genetic algorithm, IlliGAL Report No. 93005, University of Illinois, Urbana–Champaign, 1993.

[24] H. Ishibuchi, T. Murata, A multi-objective genetic local search algorithm and its application to flowshop scheduling, IEEE Transactions on Systems, Man and Cybernetics. Part C, Applications and Reviews 28 (3) (1998) 392–403.

[25] M.A. Jamison, Industry Structure and Pricing, Kluwer Academic Publishing, 2000.

[27] J.R. Koza, Genetic programming: a paradigm for genetically breeding populations of computer programs to solve problems, Stanford University, Computer Science Department Technical Report STAN-CS-90-1314, 1990.

[29] L.M. Lodish, 'Vaguely right' approach to sales force allocations, Harvard Business Review, (January–February 1974) 119–124.

[30] L.M. Lodish, Sales territory alignment to maximize profit, Journal of Marketing Research 12 (February 1975) 30–36.

[31] G. Marakas, Decision Support Systems in the 21st Century, Prentice Hall, Upper Saddle River, NJ, 1999.

[32] A. Mehrotra, E.L. Johnson, G.L. Nemhauser, An optimization based heuristic for political districting, Management Science 44 (8) (1998) 1100–1114.

[33] G.M. Morgan, The role of research and new technology in a restructured networked energy system, in: H. Chao, H.G. Huntington (Eds.), Designing Competitive Electricity Markets, Kluwer Academic Publishing, 1998, pp. 141–158.

[35] S.S. Nagel, Simplified bipartisan computer redistricting, The Stanford Law Review 17 (863) (1965).

[36] V. Pareto, Cours d'Economie Politique, Rouge, Lausanne, Switzerland, 1896.

[37] S.J. Rassenti, V.L. Smith, Deregulating electric power: market design issues and experiments, in: H. Chao, H.G. Huntington (Eds.), Designing Competitive Electricity Markets, Kluwer Academic Publishing, 1998, pp. 105–120.

[38] I. Rechenberg, Evolutionsstrategie: optimierung technischer systeme nach prinzipen der biologischen evolution, Fromman-Holzboog Verlag, Stuttgart, 1973 (2nd Edition, 1993).

[39] J.L. Ringuest, S.B. Graves, A sampling-based method for generating nondominated solutions in stochastic MOMP problems, European Journal of Operational Research 126 (2000) 651–661.

[40] F.C. Schweppe, R.D. Caramanis, R.R. Tabors, Spot Pricing of Electricity, Kluwer Academic Publishing, 1998.

[41] M. Segal, D.B. Weinbergey, Turfing, Operations Research 25 (3) (1977) 367–386.

[42] R.J. Shanker, R.E. Turner, A.A. Zoltners, Sales territory design: an integrated approach, Management Science 22 (3) (1975) 309–320.

[43] R. Slark, Electricity: trading in the new market, Energy World 258 (1998) 8–10.

[44] N. Srinivas, K. Deb, Multiobjective optimization using nondominated sorting in genetic algorithms, Evolutionary Computation 2 (3) (1994) 221–248.

[45] R.G. Torricelli, J.I. Porter, Toward the 1980 census: the reapportionment of New Jersey's congressional districts, Rutgers Journal of Computers, Technology and the Law 7 (135).

[46] W. Vickrey, On the prevention of gerrymandering, Political Science Quarterly 76 (105) (1961).

[47] A.A. Zoltners, A unified approach to sales territory alignment, Sales Management: New Developments from Behavioral and Decision Model Research, Marketing Science Institute, Cambridge, MA, 1979, pp. 360–376.

[48] A.A. Zoltners, P. Sinha, Integer programming models for sales resource allocation, Management Science 26 (3) (1980) 242–260.

Dr. Paul K. Bergey is an Assistant Professor in the College of Management at North Carolina State University.

Dr. Cliff T. Ragsdale is a Professor in the Pamplin School of Business at Virginia Tech and is author of the textbook *Spreadsheet Modeling and Decision Analysis: A Practical Introduction to Management Science* published by South-Western College Publishing.