Mixed Integer Programming Approaches to Treatment Planning for Brachytherapy – Application to Permanent Prostate Implants

EVA K. LEE *   eva.lee@isye.gatech.edu
*Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA, USA and
Radiation Oncology, Emory University School of Medicine, Atlanta, GA, USA*

MARCO ZAIDER
Brachytherapy Physics, Memorial Sloan Kettering Cancer Center, New York, NY, USA

Abstract. Mixed integer programming models and computational strategies developed for treatment planning optimization in brachytherapy are described. The problem involves the designation of optimal placement of radioactive sources (seeds) inside a tumor site. Two MIP models are described. The resulting MIP instances are difficult to solve, due in large part to dense constraint matrices with large disparities in the magnitudes of the nonzero entries. A matrix reduction and approximation scheme is presented as a computational strategy for dealing with the dense matrices. Penalty-based primal heuristic and branching strategies to assist in the solution process are also described. Numerical results are presented for 20 MIP instances associated with prostate cancer cases. Compared to currently used computer-aided planning methods, plans derived via the MIP approach use fewer seeds (20–30 fewer) and needles, and provide better coverage and conformity – measures commonly used to assess the quality of treatment plans. Good treatment plans are returned in 15 CPU minutes, suggesting that incorporation of this MIP-based optimization module into a real-time comprehensive treatment planning system is feasible.

Keywords: brachytherapy, treatment planning, mixed integer programming, optimization, prostate cancer

AMS subject classification: 90C10, 90C11, 90C90, 90-08

1. Introduction

In recent years, technical advances in medical devices have led to a resurgence in the use of radioactive implants as an alternative or supplement to external beam radiation for treating a variety of cancers. This treatment modality, known as brachytherapy, involves the placement of encapsulated radionuclides ("seeds") either within or near a tumor [5]. In the case of prostate cancer, seed implantation is typically performed with the aid of a transrectal ultrasound transducer attached to a template consisting of a plastic slab with a rectangular grid of holes in it. The transducer is inserted into the rectum and the template rests against the patient's perineum. A series of transverse images are taken through the prostate, and the ultrasound unit displays the template grid superimposed on the anatomy

---

* Corresponding author.

of the prostate. Needles inserted in the template at appropriate grid positions enable seed placement in the target at planned locations.

Despite the advances in devices that assist in accurate placement of seeds, deciding *where* to place the seeds remains a difficult problem. A treatment plan must be designed so that it achieves an appropriate radiation dose distribution to the target volume, while keeping the dose to surrounding normal tissues at a minimum.

Traditionally, to design a treatment plan, several days (or weeks) prior to implantation the patient undergoes a simulation ultrasound scan. Based on the resulting images, an iterative process is performed to find a pattern of needle positions and seed coordinates along each needle which will yield an acceptable dose distribution. Adjustments are typically guided by repeated visual inspection of isodose curves overlaid on the target contours. Since the process requires manual inspection at each iteration, the process is not only lengthy – sometimes taking up to 4 hours to complete – but it also means that only a small fraction of possible configurations can actually be examined. More importantly, by the time the implantation is performed several days later, the prostate volume will have changed in both shape and size, making the pre-plan invalid.

In recent years, computer-aided iterative approaches and automated methods have been developed to aid in brachytherapy treatment planning in the operating room [4,8,14–16,18]. In this paper, we describe a novel application of mixed integer programming to brachytherapy treatment planning and its application to the planning of permanent implants into the prostate. The treatment model, described in section 2, involves using 0/1 indicator variables to capture the placement or nonplacement of seeds in a prespecified three-dimensional grid of potential locations. In section 3, we analyze the computational issues and strategies related to the mixed integer programming instances. The MIP instances proved to be very difficult to solve to optimality. The numerical results presented in section 4 indicate that "good" solutions can be obtained via the MIP approach within 5–15 minutes. Section 5 provides some concluding remarks.

2.   Mixed integer programming models

Our treatment model involves using 0/1 variables to record placement or nonplacement of seeds in a prespecified three-dimensional grid of potential locations. In the case of prostate cancer, the locations correspond to the projection of the holes in the template onto the region representing the prostate in each of the ultrasound images. If a seed is placed in a specific location, then it contributes a certain amount of radiation dosage to each point in the images. Thus, once the grid of potential seed locations is specified, the total dose level at each point can be modeled. Let x_j be a 0/1 indicator variable for recording placement or non placement of a seed in grid position j. Then the total radiation dose at point P is given by

$$\sum_j D\big(\|P - X_j\|\big)x_j, \tag{1}$$

where X_j is a vector corresponding to the coordinates of grid point j, $\| \cdot \|$ denotes the Euclidean norm, and $D(r)$ denotes the dose contribution of a seed to a point r units away. The target lower and upper bounds, L_P and U_P, for the radiation dose at point P can be incorporated with (1) to form constraints for the MIP model:

$$\sum_j D\big(\|P - X_j\|\big)x_j \geqslant L_P,$$
$$\sum_j D\big(\|P - X_j\|\big)x_j \leqslant U_P. \tag{2}$$

Of course, not all points P in the images are considered. The images are discretized at a granularity that is conducive both to modeling the problem accurately and to enabling computational approaches to be effective in obtaining solutions in a timely manner. For discretizations that provide accurate modeling, it is typically not possible to satisfy desired dose constraints at all points simultaneously. This is due in part to the proximity of diseased tissue to healthy tissue. Also, because of the inverse square factor, the dose level contribution of a seed to a point less than 0.3 units away, say, is typically larger than the target upper bound for the point. Hence, the preprocessing techniques commonly used in integer programming literature cannot be applied directly to these dose constraints (as this will result in assigning 0 to all seed positions). In this paper, we will focus on two models which address this difficulty and computational strategies pertaining to solving the resulting MIP instances.

2.1. Model 1

This model identifies a *maximum feasible subsystem* in the proposed linear system. By introducing additional 0/1 variables one can directly maximize the *number* of points satisfying the specified bounds. In this case, constraints (2) are replaced by

$$\sum_j D\big(\|P - X_j\|\big)x_j + N_P\big(1 - v_P^L\big) \geqslant L_P,$$
$$\sum_j D\big(\|P - X_j\|\big)x_j - M_P\big(1 - v_P^U\big) \leqslant U_P, \tag{3}$$

where v_P^L and v_P^U are 0/1 variables, and M_P and N_P are suitably chosen positive constants. If a solution is found such that $v_P^L = 1$, then the right hand side of the first inequality in (3) is zero; and hence, the lower bound for the dose level at point P is not violated. Similarly, if $v_P^U = 1$, the upper bound at point P is not violated. In order to find a solution that satisfies as many bound constraints as possible, it suffices to maximize the sum of these additional 0/1 variables; i.e., maximize $\sum_P (v_P^L + v_P^U)$. In practice, achieving the target dose levels for certain points may be more critical than achieving the target dose levels for certain other points. In this case, one could maximize a weighted sum: $\sum_P (\alpha_P v_P^L + \beta_P v_P^U)$, where the more critical points receive a relatively larger weight. Using a weighted sum was important for the prostate cancer cases to be discussed in section 3. Since there were significantly fewer urethra and rectum points compared to

the number of points representing the prostate, to increase the likelihood that the former points achieved the target dose levels, a large weight was placed on the associated 0/1 variables.

The role of the constants N_P and M_P in (3) is to ensure that there will be feasible solutions to the mathematical model. In theory, these constants should be chosen suitably large so that if v_P^L or v_P^U is zero, the associated constraint in (3) will not be violated regardless of how the 0/1 variables x_j are assigned. In practice, the choice is driven by computational considerations of the optimization algorithm being used and/or by decisions by the radiation oncologist. For a branch-and-bound algorithm, it is advantageous computationally to assign values that are as tight as possible. The medical expert can guide the selection of the constants by either assigning absolute extremes on acceptable radiation dose levels delivered to each point (note that $U_P + M_P$ is the absolute maximum dose level that will be delivered to point P under the constraints in (3), and $L_P - N_P$ is the absolute minimum), or by estimating the number of seeds needed for a given plan. In the latter case, if the number of seeds needed is estimated to be between k_1 and k_2 ($k_1 \leqslant k_2$), say, then the constant N_P can be taken to be L_P minus the sum of the smallest k_1 of the values $D(\|P - X_j\|)$, and the constant M_P can be taken to be the sum of the largest k_2 such values minus U_P. Selection in this fashion will ensure that no plan having between k_1 and k_2 seeds will be eliminated from consideration.

2.2. Model 2

An alternative model involves using continuous variables to capture the deviations of the dose level at a given point from its target bounds and minimizing a weighted sum of the deviations. In this case, the constraints (2) are replaced by constraints of the form

$$
\begin{aligned}
\sum_j D\big(\|P - X_j\|\big)x_j + y_P^L &\geqslant L_P, \\
\sum_j D\big(\|P - X_j\|\big)x_j - y_P^U &\leqslant U_P,
\end{aligned}
\tag{4}
$$

where y_P^L and y_P^U are nonnegative continuous variables. The objective for this model takes the form: minimize $\sum_P (\alpha_P y_P^L + \beta_P y_P^U)$, where α_P and β_P are nonnegative weights selected according to the relative importance of satisfying the associated bounds. For example, weights associated with an upper bound on the radiation dose for points in a neighboring healthy organ may be given a relatively larger magnitude than weights associated with an upper bound on the dose level for points in the diseased organ.

2.3. Model variations and other side constraints

Both models allow the incorporation of alternative seed types. There are a variety of radioactive sources that are used for brachytherapy, including palladium-103, iodine-125, cesium-137, iridium-192, and gold-198, each of which has its own set of exposure rate constants. (Pd-103 or I-125 are commonly used for treating prostate cancer.) At

this time however, a single seed type is used in a given treatment plan. This fact is, in part, due to the difficulty of designing treatment plans with multiple seed types as well as identifying multiple seed types in post-dosimetry analysis. The allowance of multiple seed types can easily be incorporated into the MIP framework – one need only modify the total dose level expression (1) as

$$\sum_j \sum_i D_i\big(\|P - X_j\|\big)x_{ij}. \tag{5}$$

Here, x_{ij} is the indicator variable for placement or nonplacement of a seed of type i in grid location j, and $D_i(r)$ denotes the dose level contribution of a seed of type i to a point r units away. In this case, a constraint restricting the number of seeds implanted at grid point j is also needed: $\sum_i x_{ij} \leqslant 1$.

Besides the basic dosimetric constraints, the models also include dose-volume constraints to ensure sufficient coverage to the tumor volume. Other physical constraints desired by the clinicians are also incorporated into our MIP models. One could incorporate constraints to control the percentage of each tissue structure satisfying specified target bounds. Alternatively, one could – if desired – constrain the total number of seeds and/or needles used. Note also that one can ensure that target dose bounds at specific points are satisfied by fixing the associated "feasibility" variables ($v_P^L, v_P^U, y_P^L, y_P^U$) to appropriate values.

3.   Computational strategies

We note that unlike most of the industrial applications in which the MIP instances contain sparsely populated nonzero entries in the constraint matrices, the resulting MIP instances for treatment optimization have mostly dense matrices. Furthermore, the magnitudes of the coefficients range from the order of tens to tens of thousands. Below we describe some specialized strategies that have shown to be effective in improving the tractability of the resulting instances.

3.1. Matrix reduction and approximation scheme

Motivated by the dense constraint matrices and range in the magnitudes in the nonzero entries, a matrix reduction and perturbation approach was investigated. The reduction scheme partitions the constraint matrix into two submatrices, based on the magnitude of the coefficients. The right-hand-side is perturbed to compensate for the change in the matrix coefficients. Specifically, we are interested in a dense MIP instance of the following form:

$$\begin{aligned}
Ax - Ny &\geqslant L, \\
Ax + Mz &\leqslant U, \\
x \in Z_+^n, \quad y &\in \Re_+^p, \quad z \in \Re_+^q,
\end{aligned} \tag{S}$$

where A is an $m \times n$ nonnegative dense matrix, and N and M are $m \times p$ and $m \times q$ nonnegative (sparse) matrices, respectively. Let $P^S = \text{conv}\{(x, y, z): Ax - Ny \geqslant L,$

$Ax + Mz \leqslant U, \, x \in Z_+^n, \, y \in \mathfrak{R}_+^p, \, z \in \mathfrak{R}_+^q\}$. For the sake of presentation, we assume that y and z are continuous variables. However, the method described below works also when they are restricted to assume integer values.

Definition 3.1. For a chosen $\delta > 0$, split the matrix A as $A = A^1 + A^2$, where

$$a_{ij}^1 = \begin{cases} a_{ij} & \text{if } a_{ij} \geqslant \delta, \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad a_{ij}^2 = \begin{cases} a_{ij} & \text{if } a_{ij} < \delta, \\ 0 & \text{otherwise.} \end{cases}$$

Let A_i denote the ith row of matrix A, and let $\bar{x}$ and $\hat{x}$ solve the following linear programs, respectively:

$$\max_{\{i: \, A_i^2 \neq 0\}} \max \{A_i^2 x : x \in P^S\} \quad \text{and} \quad \min_{\{i: \, A_i^2 \neq 0\}} \min \{A_i^2 x : x \in P^S\}.$$

Consider the following two systems:

(i) The system

$$\begin{aligned} A^1 x - Ny &\geqslant L - A^2 \bar{x}, \\ A^1 x + Mz &\leqslant U - A^2 \hat{x}, \\ x \in Z_+^n, \quad y &\in \mathfrak{R}_+^p, \quad z \in \mathfrak{R}_+^q \end{aligned}$$

is called a δ-*reduction* for (S). It is easy to check that if (x, y, z) is feasible for (S), then it is feasible for its δ-reduction. The converse does not hold. Clearly, if $x_j \in [\alpha_j, \beta_j]$, $j = 1, \ldots, n$, we can approximate $\bar{x}_j$ and $\hat{x}_j$ by β_j and α_j, respectively.

(ii) Let $\varepsilon \in Z_+^n$. The system

$$\begin{aligned} A^1 x - Ny &\geqslant L - A^2 \varepsilon, \\ A^1 x + Mz &\leqslant U - A^2 \varepsilon, \\ x \in Z_+^n, \quad y &\in \mathfrak{R}_+^p, \quad z \in \mathfrak{R}_+^q \end{aligned}$$

is called a δ-*reduction-ε-approximation* for (S) if $A^2 \hat{x} \leqslant A^2 \varepsilon \leqslant A^2 \bar{x}$.

We caution that applying these schemes to the MIP instances is difficult. In particular, the selection of δ and ε is empirical and problem dependent, since the coefficients in each row of the dosimetric constraint matrix vary greatly. In some rows, the coefficients are distributed in the range from tens to hundreds, whereas in others there are various coefficients (fewer than 10%) which have magnitudes in the hundreds of thousands and tens of thousands, and the rest range from tens to thousands. A comprehensive development and detailed analysis of the reduction and approximation schemes and their effect on helping to solve these and other dense MIP instances can be found in [9]. Here we describe an implementation using the dosimetric constraints for model 1:

$$\sum_{j=1}^n D\big(\|P - X_j\|\big)x_j + N_P\big(1 - v_P^L\big) \geqslant L_P,$$

$$\sum_{j=1}^{n} D\big(\|P - X_j\|\big)x_j - M_P\big(1 - v_P^U\big) \leqslant U_P.$$

To select δ, one pass is made through the constraint matrix to evaluate the distribution of the nonzeros in each row. For row i, we calculate the average of the largest 5% of the nonzero coefficients, $ave^{\max}$. Initialize $\mathcal{K} = \emptyset$. The set $\mathcal{K}$ will be populated with the set of indices selected in a nondecreasing manner, starting from the smallest nonzero coefficient. We continue to place an index into $\mathcal{K}$ until $ave\{D(\|P - X_j\|): j \in \mathcal{K}\}$ is approximately equal to $\gamma * ave^{\max}$, or until the cardinality of the set $\mathcal{K}$ exceeds 50% of the total number of nonzeros in the given row. We then assign $\delta_i = \max\{D(\|P - X_j\|): j \in \mathcal{K}\}$, and $\delta = \max\{\delta_i\}$. For a row with coefficients exceeding a magnitude of 10, 000, we set $\gamma = 0.5\%$. For all other rows, we increase this value gradually with the amount of decrease in the magnitude of the coefficients.

Assuming there are m rows for each of the two classes of dosimetric constraints, the complexity of the search includes $\mathcal{O}((n + 1)\log(n + 1))$ operations for sorting the nonzero coefficients for each row, and $\mathcal{O}(m(n + 1)\log(n + 1))$ operations to set up the δ-reduction system.

3.2. Penalty-based adaptive primal heuristic procedure

The heuristic procedure is an LP-based primal heuristic in which at each iteration, some binary variables are set to 1 and the corresponding linear program is resolved. The procedure terminates when the linear program returns an integer feasible solution or when it is infeasible. In the former case, reduced-cost fixing is performed at the root node, as well as locally on each of the branch-and-bound nodes.

Again focusing on model 1, let x^{LP} be an optimal solution of some linear program relaxation at a branch-and-bound node. (For simplicity of notation, the variables v_P^L and v_P^U are included as part of x^{LP}.) At the start of the heuristic procedure, penalties, p_j, for all variables are set to zero. Let $\mathcal{U} = \{j: x_j^{\text{LP}} = 1\}$, and $\mathcal{F} = \{j: 0 < x_j^{\text{LP}} < 1\}$. The procedure works by first setting $x_j = 1$ for all $j \in \mathcal{U}$. For each $j \in \mathcal{F}$ corresponding to a grid point j with coordinates X_j, the penalty on x_j is updated according to the formula

$$p_j = \sum_{\substack{k:x_k^{\text{LP}}=1 \\ k \text{ a grid point}}} \frac{1}{\|X_k - X_j\|}.$$

For all other $j \in \mathcal{F}$ (i.e., j corresponding to v_P^L and v_P^U) the penalties remain at value zero. Let $x^{\max} = \max\{x_j^{\text{LP}}: j \in \mathcal{F}\}$, and $\varepsilon > 0$. In nondecreasing order of p_j's, the variables in $\mathcal{F}$ are set to 1 if $x_j^{\text{LP}} \geqslant x^{\max} - \varepsilon$. Since penalties for fractional v_P^U and v_P^L variables are always set to zero, these variables are always considered first for setting to 1. For every binary variable that is set to 1, logical implication (probing) [3,13] is performed to avoid conflicts in variable fixing. The value ε is chosen dynamically at each iteration so that about 10% of the fractional variables are set to 1, a strategy which appears to work well empirically for our MIP instances.

3.3. *Penalty branching strategy*

Branching variables are selected based on pseudo-costs as well as penalties. Let $\varepsilon > 0$ be given, and let $\mathcal{K} = \{j\colon \varepsilon < x_j^{\mathrm{LP}} < 1 - \varepsilon\}$. One can control the size of $\mathcal{K}$ by choosing ε so that $|\mathcal{K}|$ reaches a certain percentage with respect to $|\mathcal{F}|$. For each $k \in \mathcal{K}$ the degradations U_k and D_k in the objective value when branching with x_k set to 1 and 0, respectively, are calculated, and the penalties p_k are computed in the same manner as described in the previous section. The branching variable is chosen as that with the maximum penalty-weighted degradation, which is computed as $\max_{\{k \in \mathcal{K}\}}\{D_k + U_k/(p_k + 1)\}$. The values U_k and D_k can be calculated exactly by solving the respective linear programs, or can be approximated by performing only a fixed number of simplex iterations. The approximation strategy helps to control the required computational effort. We report results based on performing 50 simplex iterations using the steepest-edge strategy.

4.   Numerical results

We present results of our MIP approach using data from twenty prostate cancer cases. In each case, iodine-125 was used as the radioactive source, and four separate categories of points, corresponding to distinct anatomical structures, were specified and used for setting up the MIP instances. *Contour points* defined the boundary of the diseased organ in each of the slices; the regions enclosed by each boundary were populated with uniformly spaced points, termed *uniformity points*; and points representing the positions of the *urethra* and *rectum* in each slice were also specified. For the 20 cases considered, the average numbers of points in each category were: uniformity 1305, contour 461, urethra 28, and rectum 59. The lower and upper dose bounds for each point type were specified as multiples of the target prescription dose. These are tabulated in table 1.

Numerical tests were performed using two distinct models. Model 1 utilized constraints (3) and the associated objective $\max \sum(\alpha_P v_P^L + \beta_P v_P^U)$; and model 2 utilized constraints (4) and the objective $\min \sum(\alpha_P y_P^L + \beta_P y_P^U)$. Various combinations of objective function weights for each of the two models were tested. One of the goals of the numerical work involves assessing the effects of using the different models and model parameters.

For both models, it is advantageous to place relatively large weights on the objective function variables associated with urethra, rectum, and contour points. For the results reported herein, the objective function weights for the variables associated with uniformity points were set equal to 1; those associated with contour points were set equal

Table 1
Lower and upper bound specifications as multiples of target prescription dose.

	Rectum	Urethra	Contour	Uniformity
Lower bound	0	0.9	1.0	1.0
Upper bound	0.78	1.1	1.5	1.6

to the ratio of the number of uniformity points to the number of contour points; and those associated with urethra and rectum points were set equal to the number of uniformity points. Selecting such large weights for the urethra and rectum points essentially ensures that the dose contribution to these points will lie within the specified bounds. The heavy weight for the contour points assists in achieving prescription isodose curves that conform well with the boundary of the diseased prostate.

Although the problem size is only moderate, even solving the initial linear programming relaxation is memory-taxing, often resulting in a process having a total size of 400 MB (including text, data, and stack), and total resident memory approaching 400 MB. Computational experience with these instances has demonstrated that they are extremely difficult to solve to optimality, requiring strenuous computational effort to improve the objective value by a marginal amount. Even obtaining good feasible solutions which are clinically acceptable is difficult. All the numerical tests were performed on a cluster of 550 MHz Pentium III Xeon PCs running RedHat Linux 7.1.

Table 2 shows the problem statistics for models 1 and 2. Here, *Rows* and *Cols* indicate the number of rows and columns, respectively, in the constraint matrix; and *0/1 vars* indicates the number of 0/1 variables.

To illustrate the level of difficulty, in table 3 we provide a solution profile for case Pt 1 using the MIP solver of CPLEX V7.1 (with pseudo-cost branching, which appears to be the best among the possible options). We note that none of the instances were

Table 2
Problem statistics.

Pt	Model 1			Model 2		
	Rows	Cols	0/1 vars	Rows	Cols	0/1 vars
1	4398	4568	4568	4398	4568	170
2	4546	4738	4738	4546	4738	192
3	3030	3128	3128	3030	3128	98
4	2774	2921	2921	2774	2921	147
5	5732	5957	5957	5732	5957	225
6	5728	5978	5978	5728	5978	250
7	2538	2658	2658	2538	2658	120
8	3506	3695	3695	3506	3695	189
9	2616	2777	2777	2616	2777	161
10	1680	1758	1758	1680	1758	78
11	5628	5848	5848	5628	5848	220
12	3484	3644	3644	3484	3644	160
13	3700	3833	3833	3700	3833	133
14	4220	4436	4436	4220	4436	216
15	2234	2330	2330	2234	2330	96
16	3823	3949	3949	3823	3949	126
17	4222	4362	4362	4222	4362	140
18	2612	2747	2747	2612	2747	135
19	2400	2484	2484	2400	2484	84
20	2298	2406	2406	2298	2406	108

Table 3
Solution statistics for Pt 1 running on CPLEX V7.1.

CPU secs elapsed	Best IP obj.	Best LP obj.	bbnodes searched
	Model 1		
67.0	–	1888013.3	0
261.70	–	1888013.14	21
1536.15	–	1888013.01	139
10290.29	–	1888011.80	16558
20800.35	–	1888011.01	50676
30535.66	–	1888009.66	90596
41889.83	–	1888009.10	130541
52761.01	–	1888008.02	170577
67060.84	–	1888007.45	222861
Cuts added 13029			
	Model 2		
106.60	1047338492.9	3.5015e+07	61
5008.38	440437196.1	7.3056e+07	2241
10037.33	108100907.2	8.0022e+07	6241
15185.81	93550763.5	8.3096e+07	15001
20357.77	93550763.5	8.4342e+07	25001
32736.21	93550763.5	8.5909e+07	50001
45911.94	93550763.5	8.6919e+07	77321
46884.43	93550763.5	8.6987e+07	79341

solved to optimality using CPLEX V7.1. For model 1 instances, great computational effort was exerted, only to yield marginal improvement in the objective value. Instances for model 2 were slightly more manageable, although the objective value improvement eventually stalled (e.g., after 80,000 nodes for PT 1). The columns *CPU secs elapsed*, *Best IP obj.*, *Best LP obj.*, and *bbnodes searched* record, respectively, the time elapsed within the solution process, the incumbent objective value corresponding to the best integer feasible solution, the corresponding best LP value from the remaining branch-and-bound nodes, and the number of nodes solved. For model 1, we report the solution process up to 67060.84 CPU seconds and for model 2, the solution process is observed up to 46884.43 CPU seconds.

The numerical work reported in the remainder of the paper is based on a specialized branch-and-bound MIP solver which is built on top of a general-purpose mixed integer research code (MIPSOL) [7], using CPLEX V7.1 as the intermediate LP solver. The general-purpose code, which incorporates pre-processing, reduced-cost fixing, cut generation, and fast heuristics, has been quite effective in solving the instances reported in MIPLIB3 [2]. For the prostate cancer instances reported in this paper, the matrix reduction and approximation scheme, the penalty-based adaptive primal heuristic procedure, and penalty branching strategy described in section 3 were implemented to assist in the solution process.

None of the instances for model 1 were solved to proven-optimality, whereas for

model 2, all except one were solved to optimality. In tables 4a and 4b the solution statistics for both models are given. We set the running time limit to be 10,000 CPU seconds for model 1. In each table, the column labeled *Pt* denotes the patient case; the column labeled *Initial LP obj.* lists the optimal objective value of the initial LP relaxation; and the columns *First heuristic (secs, obj.)* list the elapsed time when the heuristic procedure is first called and the objective value corresponding to the feasible integer solution returned by the heuristic. For table 4a, the columns *Best LP obj.* and *Best IP obj.* report, respectively, the LP objective bound corresponding to the best node in the remaining branch-and-bound tree and the incumbent objective value corresponding to the best integer feasible solution upon termination of the solution process (10,000 CPU seconds). In table 4b, the columns *Optimal IP obj.*, *bb nodes*, and *Elapsed time* report, respectively, the optimal IP objective value, the total number of branch-and-bound tree nodes solved, and the total elapsed time for the solution process.

Table 5 compares the running time and number of branch-and-bound nodes solved for model 2 when matrix reduction was used. The column *CPU factor* shows the improvement in running time for each instance, and the column *% Nonzero reduced* records the percentage of nonzero entries in the constraint matrix reduced by the procedure. Using the reduction approach, the running time for model 2, with the exception of one instance, decreased by 1.5–92 times for the 19 instances, with an average decrease of 14.6 times. The readers are referred to [9] for more details regarding this approach.

Table 4a
Solution statistics for model 1 (maximization).

Pt	Initial LP obj.	First heuristic		Best LP obj.	Best IP obj.
		secs	obj.		
1	1888013.3	376.1	1752286	1873433.93	1766609
2	1809964.8	397.5	1736946	1796642.43	1736946
3	687448.6	33.4	587712	633843.00	593228
4	803564.9	250.9	753672	802134.58	765115
5	2855667.4	1345.9	2638679	2835825.38	2649950
6	2925181.3	1349.2	2805284	2907792.52	2805284
7	651682.5	62.7	582314	639160.50	598630
8	1132430.4	232.3	1062561	1112930.1	1075670
9	677253.3	111.4	639527	669073.94	641643
10	286986.4	25.5	252368	274188.69	257492
11	2585974.0	695.0	2453886	2529795.54	2462279
12	983328.6	97.7	804213	945400.35	817875
13	862373.4	67.2	744450	27676.91	795149
14	1611020.9	507.3	1509329	1590484.06	1531009
15	438667.7	39.9	376087	428376.60	396064
16	1273297.8	274.2	1170743	1248805.82	1204870
17	892239.9	136.4	747929	817014.30	757446
18	683918.1	94.6	581684	666083.02	592861
19	425871.6	16.0	341328	403235.91	376179
20	360474.3	26.4	288973	343623.43	309499

Table 4b
Solution statistics for model 2 (minimization).

Pt	Initial LP obj.	First heuristic		Optimal IP obj.	bb nodes	Elapsed time
		secs	obj.			
1	29973430.5	21.7	440437196.1	93550763.6	377	9706.0
2	19921521.4	34.7	179171112.9	49156651.9	9184	378857.0
3	−11333869.7	5.2	97625273.7	50517325.3	4051	27724.0
4	2597572.3	18.7	189610043.6	21005621.8	1377	27485.0
5	73684327.8	112.4	467410325.8	93828192.8	1293	748292.3[a]
6	36902037.2	105.3	524058129.4	64216816.0	5293	1136221.7
7	45848681.6	6.5	302836935.1	118325071.3	712	4655.5
8	17614469.1	32.3	250057575.6	73399636.5	62373	1863362.0
9	14691002.3	17.3	344540093.9	57209440.5	1643	41212.1
10	28197622.0	2.1	90862556.4	55251869.2	883	2619.1
11	172211617.5	20.5	616562230.8	293530404.4	627	13904.1
12	292898229.2	11.5	785823995.0	518235776.6	1985	35718.5
13	−163007095.9	4.3	−21671699.9	−77173221.5	481	2817.6
14	40303495.4	27.1	378940132.7	119586431.2	1408	58654.2
15	89432119.5	5.5	236921860.0	191780731.4	10838	55913.8
16	78434032.7	14.1	244541089.6	148828362.1	1282	25969.0
17	−830974566.8	2.7	−717574515.4	−799657523.1	25	178.2
18	155505947.5	9.6	700452425.7	351076662.5	82118	554737.2
19	73628152.3	2.1	204208781.0	149604823.5	377	1207.8
20	−45968824.5	1.8	57904156.7	15635930.3	415	1222.5

[a] Not optimal.

Despite the computational difficulty, high-quality clinically desirable treatment plans were obtained using both models. In table 6 we report some clinically relevant statistics. For the results reported in the table, the treatment plans are those associated with the first feasible solution obtained from our specialized solver applied to model 1 and model 2. The cases are categorized according to the target prescription dose (100 Gy, 120 Gy, or 160 Gy). *Prostate vol (cc)* records the volume of the prostate, *Activity (mCi)* is the activity rate of the implanted seeds, and *conformity* and *coverage* are measures of the quality of the generated plans. Conformity is defined as the ratio of the volume of the prescription isodose surface determined by the plan to the portion of the target volume within this surface. Coverage measures the ratio of the target volume within the prescription isodose surface to the entire target volume. For an ideal plan both the conformity and coverage indices should be 1. A conformity index greater than 1 provides a measure of the amount of healthy normal tissue receiving dose at least as high as the prescription dose. In particular, a smaller conformity index implies that nearby healthy tissue is exposed to less radiation, thus reducing the probability of complications. Compared to currently used computer-aided planning methods, plans derived via the MIP approach use fewer seeds (20–30 fewer) and needles, and provide better coverage and conformity indices [4,8,15]. There are only marginal differences in the clinical properties between the first plans returned by the two MIP models.

Table 5
Contrasting solution time from matrix reduction technique for model 2 (minimization).

Pt	Original		Reduced system			
	No. bb nodes	CPU secs elapsed	No. bb nodes	CPU secs elapsed	CPU factor	% Nonzero Reduced
1	377	9706.0	468	6406.4	1.5	34.2
2	9184	378857.0	2190	22285	17	44.3
3	4051	27724.0	2554	4766.5	5.8	39.4
4	1377	27485.0	980	7852.8	3.5	35.9
5	1293	748292.3[a]	104	10539.3	71	44.6
6	5293	1136221.7	504	12301.6	92.4	52.8
7	712	4655.5	242	802.9	5.8	30.3
8	62373	1863362.0	5460	64638.8	28.8	38.1
9	1643	41212.1	1442	8862.7	4.7	23.5
10	883	2619.1	620	741.7	3.5	12.5
11	627	13904.1	720	6525.8	2.1	48.1
12	1985	35718.5	1607	5011.4	7.2	29.1
13	481	2817.6	242	702.9	4.2	37.2
14	1408	58654.2	611	17585.9	3.3	42.6
15	10838	55913.8	5699	12292.6	4.5	15.0
16	1282	25969.0	3874	13779.3	4.5	32.3
17	25	178.2	56	324.9	0.55	42.6
18	82118	554737.2	31172	102346.2	5.4	25.8
19	377	1207.8	209	162.5	7.5	22.0
20	415	1222.5	390	289.6	4.2	25.0

[a] Not optimal.

To contrast the effect of weights on the resulting coverage and conformity scores of plans, we illustrate graphically in figures 1 and 2 the changes in toxicity (for this presentation, toxicity is defined as conformity -1) to external normal tissue and coverage when various objective function weights are selected for the contour points. In particular, we highlight three weight combinations for contour points: objective function weights equal to 1 (C-small); objective function weights eqaul to the ratio of the number of uniformity points to the number of contour points (C-medium); and objective function weights equal to the number of unformity points (C-large). Detailed analysis using different weights, together with the clinical significance, is given in [10]. From the graphs, both models illustrate similar trends: larger weights on contour points provide better coverage while at the same time producing slightly higher toxicity to normal tissue adjacent to the prostate, which is in agreement with the conflicting properties of coverage and conformity. Comparing plans obtained via model 1 and model 2, model 2 tends to yield plans with slightly higher coverage and conformity, with increases in toxicity more significant when large weights are placed on the contour points.

Table 6
Clinical significance of the MIP generated plans.

Pt	Prostate vol. (cc)	Activity (mCi)	Model 1			Model 2		
			Conformity	Coverage	No. seeds	Conformity	Coverage	No. seeds
				100 Gy				
1	49.1	0.592	1.20	0.973	40	1.21	0.973	40
2	53.6	0.450	1.16	0.994	51	1.21	0.994	53
3	34.2	0.334	1.18	0.945	51	1.23	0.945	52
4	31.0	0.400	1.17	0.985	42	1.20	0.978	43
5	68.7	0.590	1.21	0.985	50	1.22	0.988	51
6	68.1	0.450	1.20	0.986	64	1.22	0.982	64
7	26.7	0.400	1.25	0.970	39	1.29	0.973	40
8	40.8	0.450	1.21	0.983	44	1.21	0.985	44
9	28.9	0.500	1.28	0.988	32	1.35	0.992	34
				120 Gy				
10	16.6	0.468	1.29	0.939	28	1.29	0.973	
				160 Gy				
11	66.1	0.520	1.12	0.964	85	1.15	0.967	86
12	38.3	0.544	1.23	0.951	58	1.25	0.951	58
13	39.9	0.450	1.22	0.986	70	1.27	0.990	72
14	48.2	0.450	1.17	0.989	76	1.23	0.993	71
15	24.3	0.550	1.18	0.980	42	1.23	0.973	42
16	45.3	0.592	1.15	0.975	57	1.15	0.980	57
17	50.7	0.463	1.11	0.874	72	1.10	0.873	72
18	26.4	0.500	1.29	0.970	51	1.34	0.986	52
19	25.4	0.450	1.15	0.964	48	1.20	0.977	49
20	25.6	0.400	1.16	0.977	57	1.13	0.965	56

5. Discussion

We presented a mixed integer programming approach to brachytherapy treatment planning and provided numerical and clinical results when applied to permanent prostate implants. The computational work presented herein demonstrates that our models can produce high-quality and clinically acceptable treatment plans for prostate cancer cases. The MIP models provide the flexibility to enforce clinically critical dosimetric conditions, and to prioritize dose level achievement for vital organs and tissues near the diseased structure.

Although the mixed integer programming problem instances are difficult to solve to optimality, with our specialized heuristic procedure and branching routines, good treatment plans are returned within 15 CPU minutes. Compared to currently used computer-aided planning methods, plans derived via the MIP approach use fewer seeds (20–30 fewer) and needles, and provide better coverage and conformity – measures commonly used to assess the quality of treatment plans. This suggests that incorporation of this MIP-based optimization module into a comprehensive treatment planning system for

Normal tissue toxicity index (model 1)

Tumor coverage index (model 1)
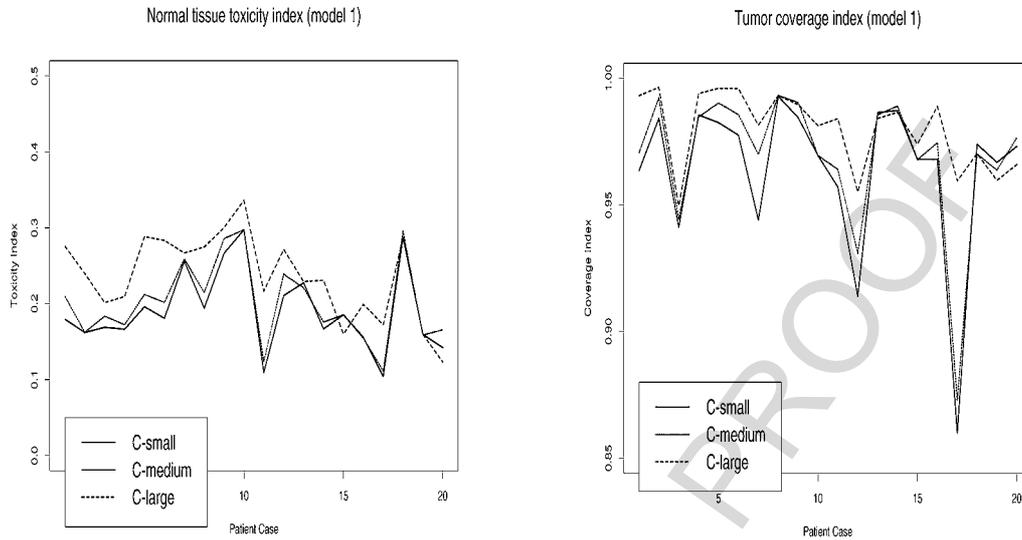
Figure 1. Normal tissue toxicity and tumor coverage for model 1 with varying weights on prostate boundaries.

Normal tissue toxicity index (model 2)

Tumor coverage index (model 2)
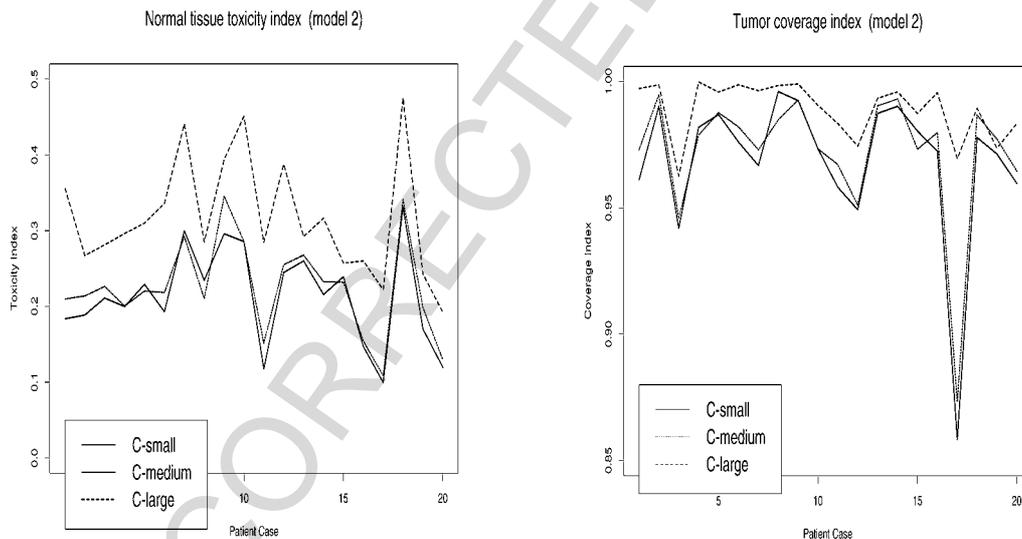
Figure 2. Normal tissue toxicity and tumor coverage for model 2 with varying weights on prostate boundaries.

use in the operating room is feasible. This work has the potential to have a direct positive impact on treatment success, as well as in eliminating the time-consuming task of generating treatment plans via iterative approaches. Interested readers can refer to the publications [4,8,15,17] for additional medical and clinical insight regarding this research.

The preliminary matrix reduction techniques appear to help in reducing the computational effort in solving the MIP instances to optimality. Studies detailing the integer programming aspects, including detailed computational strategies and numerical comparisons for applying matrix reduction techniques in solving the dense MIP instances will be reported in [9].

Acknowledgments

References

[1] G.K. Bahr, J.G. Kereiakes, H. Horwitz, R. Finney, J. Galvin and K. Goode, The method of linear programming applied to radiation treatment planning, Radiology 91 (1968) 686–693.

[2] R.E. Bixby, S. Ceria, C.M. McZeal and M.W.P. Savelsbergh, *An Updated Mixed Integer Programming Library: MIPLIB 3* (1998).

[3] R.E. Bixby and E.K. Lee, Solving a truck dispatching scheduling problem sing branch-and-cut, Ph.D. Thesis, Department of Computational and Applied Mathematics, Rice Univerity, Houston, TX (1993). Companion paper appeared in Operations Research 46(3) (1998) 355–367.

[4] R.J. Gallagher and E.K. Lee, Mixed integer programming optimization models for brachytherapy treatment planning, Journal of the American Medical Informatics Association, Symposium Supplement, ed. D.R. Masys (1997) 278–282.

[5] Interstitial Collaborative Working Group, *Interstitial Brachytherapy: Physical, Biological, and Clinical Considerations* (Raven Press, New York, 1990).

[6] M. Langer, R. Brown, M. Urie, J. Leong, M. Stracher and J. Shapiro, Large scale optimization of beam weights under dose-volume restrictions, International Journal of Radiation Oncology and Biological Physics 18 (1990) 887–893.

[7] E.K. Lee, Computational experience of a general purpose mixed 0/1 integer programming solver (MIPSOL), Technical Report, Georgia Institute of Technology (1997).

[8] E.K. Lee, R.J. Gallagher, D. Silvern, C.S. Wuu and M. Zaider, Treatment planning for brachytherapy: an integer programming model, two computational approaches and experiments with permanent prostate implant planning, Physics in Medicine and Biology 44(1) (1999) 145–165.

[9] E.K. Lee, E.L. Johnson and G.L. Nemhauser, Matrix reduction techniques for dense MIPs, Working paper.

[10] E.K. Lee and M. Zaider, Analysis of the use of different clinical objectives in the design of treatment plans for brachytherapy, Working paper.

[11] J. Legras, B. Legras and J.-P. Lambert, Software for linear and nonlinear optimization in external radiotherapy, Comput. Prog. Biomed. 15 (1982) 233–242.

[12] I.I. Rosen, R.G. Lane, S. Morrill and J.A. Belli, Treatment plan optimization using linear programming, Phys. Med. Biol. 18 (1991) 141–152.

[13] M.W.P. Savelsbergh, Preprocessing and probing for mixed integer programming problems, ORSA Journal on Computing 6 (1994) 445–454.

[14] D.A. Silvern, Automated OR prostate brachytherapy treatment planning using genetic optimization, Ph.D. Thesis, Department of Applied Physics, Columbia University, New York (1998).

[15] D.A. Silvern, E.K. Lee, R.J. Gallagher, L.G. Stabile, R.D. Ennis, C.R. Moorthy and M. Zaider, Treatment planning for permanent prostate implants: genetic algorithm versus integer programming, Medical & Biological Engineering & Computing, Supplement 35, Part 2 (1997) 850.

[16] R.S. Sloboda, Optimization of brachytherapy dose distributions by simulated annealing, Medical Physics 19 (1992) 955–964.

[17] C.S. Wuu, R.D. Ennis, P.B. Schiff, E.K. Lee and M. Zaider, Dosimetric criteria for selecting a source activity and a source type (^{125}I or ^{103}Pd) in the presence of irregular pellet placement in permanent prostate implants, International Journal of Radiation Oncology, Biology and Physics 47 (2000) 815–820.

[18] Y. Yu and M.C. Schell, A genetic algorithm for optimization of prostate implants, Medical Physics 23 (1996) 2085–1091.