Computers & Operations Research 56 (2015) 134–150

Contents lists available at ScienceDirect

Computers & Operations Research

journal homepage: www.elsevier.com/locate/caor

A unified mixed-integer programming model for simultaneous fluence weight and aperture optimization in VMAT, Tomotherapy, and Cyberknife

Kerem Akartunalı [a,*], Vicky Mak-Hau [b,1], Thu Tran [c,2]

[a] Department of Management Science, University of Strathclyde, 40 George Street, Glasgow G1 1QE, UK
[b] School of Information Technology, Deakin University, Burwood, Vic 3125, Australia
[c] Andrew Love Cancer Centre, Barwon Health, Geelong, Vic 3220, Australia

ARTICLE INFO

Available online 22 November 2014

Keywords:
OR in medicine
Integer programming
Heuristics
Radiotherapy treatment planning
Metaheuristics
Lagrangian relaxation

ABSTRACT

In this paper, we propose and study a unified mixed-integer programming model that simultaneously optimizes fluence weights and multi-leaf collimator (MLC) apertures in the treatment planning optimization of VMAT, Tomotherapy, and CyberKnife. The contribution of our model is threefold: (i) Our model optimizes the fluence and MLC apertures simultaneously for a given set of control points. (ii) Our model can incorporate all volume limits or dose upper bounds for organs at risk (OAR) and dose lower bound limits for planning target volumes (PTV) as hard constraints, but it can also relax either of these constraint sets in a Lagrangian fashion and keep the other set as hard constraints. (iii) For faster solutions, we propose several heuristic methods based on the MIP model, as well as a meta-heuristic approach. The meta-heuristic is very efficient in practice, being able to generate dose- and machinery-feasible solutions for problem instances of clinical scale, e.g., obtaining feasible treatment plans to cases with 180 control points, 6750 sample voxels and 18,000 beamlets in 470 seconds, or cases with 72 control points, 8000 sample voxels and 28,800 beamlets in 352 seconds. With discretization and down-sampling of voxels, our method is capable of tackling a treatment field of 8000–64, 000 cm^3, depending on the ratio of critical structure versus unspecified tissues.

1. Introduction

Radiation therapy or radiotherapy has become one of the most common treatment methods for cancer, besides chemotherapy and surgery, with almost two-thirds of all cancer patients expected to have radiotherapy at some stage in their treatment plan [11]. In this treatment, high-energy radiation is used to shrink tumors and kill cancer cells, where radiation damages the DNA of cancer cells [17]. The radiation can be delivered either by radioactive source(s) placed in or near the tumor, called *brachytherapy*; or by a machine outside the body, called *external radiation therapy*, the most common form of radiotherapy. Since our focus is external radiotherapy, we will refer to that simply as radiotherapy in the remainder of the paper.

In radiotherapy, radiation beams are produced by a Linear Accelerator (LINAC), aimed towards the tumor and its surrounding tissues where cancer may have spread. The LINAC is mounted on a gantry, and the gantry rotates the source of radiation beams around the body of a patient. Volumetric-modulated arc therapy (VMAT), Tomotherapy and CyberKnife are recent major advances in external beam radiotherapy.

In VMAT, the gantry can rotate around a patient's body by 360° in a co-planar manner (see Fig. 1). Co-planar treatments are possible through rotation of the LINAC couch. Radiation is continuously delivered through one or multiple arcs (see, e.g., Elekta Infinity [9] or Varian's RapidArc [32]). An “arc” does not necessarily have to be a full 360° rotation. In Tomotherapy, the source of radiation will continuously rotate around the body of a patient in a helical manner (see [2]), hence it is capable of delivering non-coplanar beams. In CyberKnife, the source of radiation is mounted on a robotic arm, and therefore it can deliver radiation beam from almost any point in space (see [1]). In addition to cancer treatment, VMAT and Tomotherapy have also been used for Total Marrow Irradiation (TMI) in reducing Leukemia relapse ratio (see [10,38]).

The modulation of the radiation beams is carried out using collimators. For VMAT, a multi-leaf collimator (MLC) is mounted in the LINAC head. The MLC is made up of leaves, which will block the

---

[*] Corresponding author. Tel.: +44 141 5484542.
E-mail addresses: kerem.akartunali@strath.ac.uk (K. Akartunalı), vicky.mak@deakin.edu.au (V. Mak-Hau), thut@barwonhealth.org.au (T. Tran).
[1] Tel.: +61 3 92517670.
[2] Tel.: +61 3 42152607.

http://dx.doi.org/10.1016/j.cor.2014.11.009
0305-0548/© 2014 Elsevier Ltd. All rights reserved.

radiation and are arranged parallel to each other in two sets of opposing banks. These leaves can move independent from each other and can create customized beam shape by positioning themselves in a planned position. The radiation field formed by the MLCs are known as *aperture* (see Fig. 2). For example, if an aperture is formed by 20 leaves, each 1 cm thick, and we have 20 leaf positions, then the MLC is said to have 400 bixels (or *beamlets*).

On the other hand, TomoTherapy uses simply binary MLCs, i.e., each leaf has only two positions, open or close. The equipment often uses 32 leaves, seen as rows in a beam's eye view as per Fig. 2. CyberKnife has various collimators, e.g., the M6 FI model uses multiple circular collimators of different sizes (called IRIS collimators), whereas the M6 FM model uses an MLC called InCise.

Since MLCs are common in all three machines, we focus our model on optimizing the radiation intensity or fluence weight as well as the MLC aperture simultaneously. Some MLCs do not allow *interdigitation*, that is, the left leaf of a row cannot collide with the right leaf of the neighboring rows, and vice versa. This is also known as the *interleaf constraints* in the literature. Some MLCs allow full interdigitation, while some allow no interdigitation or no interdigitation with a gap. There are also other machinery constraints for VMAT, such as a limit to the speed a leaf can move in consecutive control points. Such a restriction can be translated to the number of positions, or columns, a leaf is allowed to move between control points before the MLC carriage shifts to enable further proximal/distal positions.

Fig. 1. Gantry that can rotate by 360°.

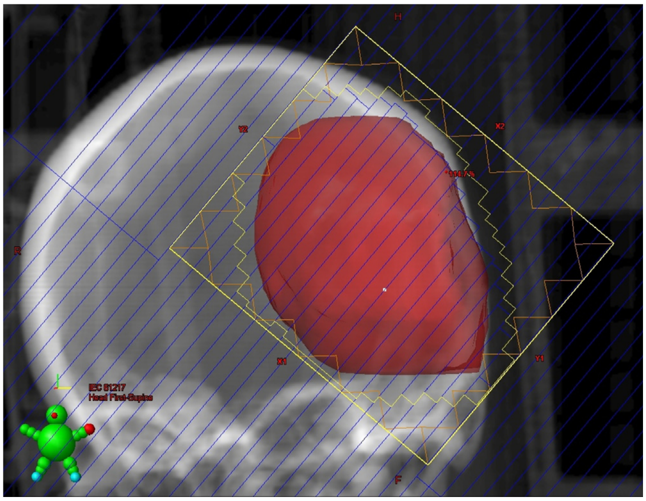

Fig. 2. An MLC can form a shape that resembles the shape of the tumor.

1.1. *Recent developments in treatment planning optimization*

To plan a radiotherapy treatment that involves the use of an MLC, one must decide from *where* the radiation beams should be delivered, the *intensity* of radiation that should be delivered from each location, and *what apertures* the MLC should form at those locations. The general goals of a treatment is (i) for tumors to receive enough radiation so that they can be eliminated, (ii) for organs at risk (OARs) to be spared from radiation as much as possible for minimal damage to healthy tissue, and (iii) for the overall treatment time to be as short as possible for patient's comfort. The tumor area is usually given a margin that will cover some surrounding tissues where cancer may have spread to and to compensate for motion and setup. This area is called the Planning Target Volume (PTV). Before the treatment planning optimization takes place, a radiation oncologist prescribes the dose limits to the different structures: a lower bound on the dose to the PTVs (or a volume constraint such as "at least 95% of the PTV must receive a dose of 73.7 Gray (Gy) or above"); and an upper bound on the OARs (or volume limits such as "no more than 35% of the bladder can receive more than 40 Gy of radiation").

The question of *where* will be answered by a set of locations, commonly referred to as the *control points* (CPs). The question of *intensity* at CP k is determined by the dose rate (r_k, in Monitor Unit (MU) per second), and the gantry speed (s_k, also in MUsec). In recent literature, e.g., Peng et al. [25] and Sun et al. [29,30], a value called *fluence weight* that equals to r_k/s_k is used for simplicity, and we also use this concept in this paper for consistency. Finally the answer to *what aperture* the MLC should form is where combinatorial optimization comes in. Further, there is a limit in the maximum fluence weight that can be delivered from each control point. For VMAT, the change in fluence weight is also limited in consecutive control points.

Notice that for all three treatment modalities, only one MLC aperture is considered at each control point. On the other hand, in intensity-modulated radiotherapy (IMRT), a widely used form of radiotherapy with machinery similar to VMAT, there are very few control points (usually only 5–7) in each treatment, but multiple MLC apertures are used at each control point.

Traditionally, fluence weights and MLC apertures are optimized separately. In IMRT, three separate optimization problems are solved (see, e.g., [13,8,7]). First, a *beam-angle optimization problem* is decided. These angles are either predetermined by an experienced treatment planner or calculated by solving some optimization problems (see, e.g., [22,35,39]). The output is a number of angles (usually 5–7, at most around 10 for the most complex cancer cases, i.e., head and neck) from where radiation is delivered. After that, a *fluence map optimization* problem is solved [26], which will provide us with one intensity matrix for each beam-angle (control point). Finally, a *realization problem* is solved to find the right MLC apertures that will decompose the intensity matrices, and to find the most time-efficient way to do so [18]. The field of minimizing treatment times for the step-and-shoot MLC radiotherapy practically began with the work of Xia and Verhey [36]. The most recent advances in the minimizing of total treatment time can be found in Cambazard et al. [5], Mason et al. [18], Taskin et al. [31].

Since the underlying mathematical problem for the VMAT is different, methods developed for the IMRT cannot be directly applied to the VMAT. The difference mainly stems from the interconnectivity between different control points and the computational complexity due to higher degree of freedom in VMAT, as noted by Yu and Tang [37]. To our knowledge, the treatment planning optimization of the