SOLVING AN EXTENSION OF A GENERALISED ASSIGNMENT PROBLEM FOUND IN SUGAR CANE HARVEST SCHEDULING: A DYNAMIC TABU SEARCH APPROACH

Andrew J. Higgins
CRC for Sustainable Sugar Production
CSIRO Tropical Agriculture
306 Carmody Road
St. Lucia 4067
Australia
Fax: +617 3214 2206
Ph: +617 3214 2340
E-mail: Andrew.Higgins@tag.csiro.au

AND

Russell Muchow
CSIRO Tropical Agriculture
306 Carmody Road
St. Lucia 4067
Australia
Fax: +617 3214 2206
Ph: +617 3214 2261
E-mail: Russell.Muchow@tag.csiro.au

ABSTRACT

A sugar mill region contains large differences in sugar yield per hectare due to harvest date, harvest age of crop, geographical location and crop class. While it is desirable to harvest all cane when the likely sugar yields are at the season's peak, limited harvesting, mill crushing and cane transport capacity dictate that the harvesting of cane be carried out over a harvest season of several months. In Australia, the harvest season currently extends from early winter to late spring, since this is when the sugar content of cane is at its highest. A typical mill region in Australia, producing about 1.5 million tonnes of cane, contains around 200 - 300 privately owned farms that are contracted to between 20 and 100 harvesting groups. Each farm usually contains between 20 and 100 paddocks, which can differ in terms of cane variety, crop class and age at harvest. A crop class is whether the cane on a farm paddock is plant crop (around 1 year after planting), first ratoon (around 2 years after planting) and so on with a ratoon being the re-growth after a crop is harvested. A crop cycle is a series of crop classes until the paddock is ploughed out. As a result, decisions have to be made throughout the harvest season, on the harvest time interval (fortnightly basis) of each farm paddock and whether the paddock is ploughed out for a new crop cycle, to maximise industry and farmer profitability. The costs associated with farmers and the mill are accounted for in the objective. A planning horizon of many harvest seasons is used since a crop cycle can contain as many as six ratoons. The modelled constraints are transport, mill crushing and harvesting capacity.

The above sugar cane harvest scheduling problem represents an extension of a very large-scale generalised assignment problem (GAP) where farm paddocks are jobs and harvest time intervals are agents. Unlike the GAP, the sugar cane harvest scheduling problem contains two types of decision variables as well as a many year planning horizon, and more complex constraints. Most existing approximation solution methods for solving the GAP are designed for small to medium size

problems (up to 200 jobs) and are not designed for the more complex sugar cane harvest scheduling model. Here, a two stage solution technique is proposed for which the second stage uses the new version of tabu search (TS).

In the first stage of the solution technique, an initial solution is found by firstly solving a linear program approximation model. In this approximation model, the main decision variables are the tonnes of cane harvested at each harvest time interval for a given harvest age by crop class by geographical location. Given the linear program solution, the discrete decision variables of the main model (harvest dates of farm paddocks) are matched as closely as possible to the linear program variables to produce the initial solution. This was a much more effective method for generating an initial solution than using random methods.

A new version of tabu search, designed for very large generalised assignment and other integer programming problems, is proposed for the second stage of the solution process. There are three neighbourhoods for the sugar cane harvest scheduling model. The first two neighbourhoods are insert and swap moves, which are common to the GAP. The third neighbourhood is for the decision variables of whether a paddock is ploughed out in a given year for replanting. Unlike the applications to the GAP's in the literature, a full neighbourhood search in the sugar cane harvest scheduling model is too inefficient due to the size being up to about 80 million possible moves. Because of this, a sample of the neighbourhood is chosen. Depending on the problem at hand, as well as the how close the solution is to optimal during the solution process, the three neighbourhoods will have different likelihoods of improving the solution or producing the best possible move in a random sample. Choosing the neighbourhood sample sizes that appeared to produce the best results required experimenting with different combinations of sample sizes using the mill region case studies. The sample size combination that produced the best all round results in the experimentation was 300 for the insert neighbourhood, 100 for the swap and ploughout neighbourhoods respectively.

Existing versions of TS in the literature vary in terms of how they handle infeasibility search. The best of the existing versions applies oscillation that allows the TS to swap between feasible and infeasible search spaces. However, as the TS proceeds over time, the benefits of search through feasible versus infeasible space will vary. A new approach is to have the fitness of a solution to be a dynamic weighting between the objective function value and a measure of how much the constraints are violated. This dynamic weighting is according to whether the search through feasible versus infeasible space at that point of the solution process produces the greatest gains in objective function value. This is implemented by calculating the solution improvement through feasible and infeasible search over the b most recent moves implemented, where b is selected to give the best results (= 50 for the sugar cane harvesting problem). If the search through feasible space produces much greater gains than the search through infeasible space, the fitness of the solution will be sensitive to infeasibilities and will force the search more towards feasible space where the gains will be greater.

While the best fixed neighbourhood sample sizes are defined above, the best neighbourhood sample sizes do vary throughout the solution process. For example, if at a certain stage of the solution process, the constraints were tight and solution improvements were less frequent, swap moves will more likely improve the solution than insert moves. For the sugar cane harvesting model and given the last b moves implemented, the sample size of a neighbourhood is the ratio of solution improvement per move for that neighbourhood as opposed to all neighbourhoods, multiplied by the total sample size of all neighbourhoods. Having the dynamic neighbourhood sample sizes and oscillation reduces the amount of wasted search throughout the solution process and thus produces a faster increase in solution quality per unit time.

The new version of TS was firstly tested on randomly generated large scale GAP of up to 50000 jobs and 40 agents. The best existing version of TS for these size problems used a fixed weighted oscillation between feasible and infeasible space. When run for 10 minutes of CPU time on a Pentium II PC, the new TS produced superior average results compared to the best of the existing TS versions in the literature on all occasions. For a given solution quality, the new TS required at least 50% less CPU time to find such a solution, and on average over 70% less CPU time.

Through sugar industry participation, the sugar cane harvest scheduling model and the two stage solution technique were applied to five large Australian sugar mill regions to assess the scope for profitability for optimal cane supplies as opposed to current practice decisions. These regions were selected for case studies due to strong sugar industry partnership in model application and delivery of benefits to the growing and milling sectors. Given that the mill regions contain up to 10000 paddocks and the planning horizon is 15 years, this is equivalent to a generalised assignment problem of 150000 jobs and 12 machines (harvest season of about 12 fortnights). Application of the model and solution technique to the five mill regions achieved up to a 7% increase in industry profitability (or A$6M per mill per year) over current practice farming decisions.