INVERSE ASSIGNMENT PROBLEM FOR TIMETABLING IN TUTORING SCHOOL

Takuro Hidaka and Tomomi Matsui

Department of Information and System Engineering,
Faculty of Science and Engineering,
Chuo University,
Kasuga, Bunkyo-ku, Tokyo 112-8551, Japan.

Abstract

We consider a problem for constructing timetable of a tutoring school. For each time slot, we have a set of students and a set of teachers. We need to assign each student to a teacher subject to an upper bound of the number of students assigned to a teacher. The problem finds an assignment which maximizes the sum of fitness of selected student-teacher pairs.

When we use the assignment model, we need to determine a value of fitness for each student-teacher pair. We propose an inverse optimization problem for finding fitness values which accommodate to real schedule data used in a tutoring school. We show that our inverse optimization problem becomes a linear programming problem.

Keywords: timetabling, assignment problem, inverse optimization

1. Introduction

In this paper, we consider a problem for constructing timetable of a tutoring school. For each time slot, we have a set of students and a set of teachers. We need to assign each student to a teacher subject to an upper bound of a number of students assigned to a teacher. The problem finds an assignment which maximizes the sum of fitness of selected student-teacher pairs.

The problem is essentially equivalent to a weighted assignment problem (weighted bipartite matching problem). Thus, a network algorithm solves the problem efficiently. In addition, it is well-known that a linear relaxation problem has an optimal solution which is also optimal to the original weighted assignment problem.

When we use the assignment model described above, we need to determine a value of fitness for each student-teacher pair. We propose an inverse optimization problem for finding fitness values which accommodate to real schedule data used in a tutoring school.

When solving an optimization problem, we usually assume that parameters such as fitnesses are known and that we are interested in finding an optimal solution. However, in practice, it may happen that we know that certain solutions are optimal from experiments. The idea of inverse optimization is to find a cost function (fix parameters) as to make a given feasible solution optimal.

Geophysical scientists started to study inverse (optimization) problems. The book by Tarantola (Tarantola 1987) gives a discussion on inverse (optimization) problems in the geophysical sciences. Burton and Toint (Burton 1992, Burton 1994) studied inverse shortest path problems arising in seismic tomography used in predicting the movement of earthquakes. Since then, many inverse optimization problems, including inverse spanning tree problems (Sokkalingam 1999, Ahuja 2000, Hochbaum 2003), and inverse network flow problems (Xu 1995, Ahuja 2002, Jianga 2010) have been considered by many researchers. For comprehensive lists of related works, see survey papers (Heuberger 2004, Ahuja 2001).

In Section 2, we formulate our timetabling problem as a variation of assignment problem. In Section 3, we discuss a variation of inverse assignment problem. We use complementary slackness theorem and characterizes fitness values. Section 4 gives a linear programming formulation of our inverse optimization problem.

Copyright © 2011 by the Japan Society of Mechanical Engineers
2. Assignment Problem

In this section, we introduce our timetabling problem and formulate the problem to a variation of assignment problem.

Let $S$ be a set of students and $T$ be a set of teachers. A non-negative integer $b$ denotes a capacity of teacher, which represents the maximum number of students assigned to a teacher. In our application setting, $b$ is equal to 3. For each pair $(i, j) \in S \times T$, we introduce a weight $w_{ij}$ which gives fitness (and/or compatibility) of a student-teacher pair $(i, j)$. A problem for finding an assignment of students to teachers maximizing sum of weights is formulated as follows:

\begin{equation}
\text{AP: max. } \sum_{(i,j) \in S \times T} w_{ij}x_{ij}
\end{equation}

\text{s.t.} \begin{align*}
\sum_{i \in S} x_{ij} & \leq b \quad (\forall j \in T), \\
\sum_{j \in T} x_{ij} & = 1 \quad (\forall i \in S), \\
x_{ij} & \in \{0, 1\} \quad (\forall (i, j) \in S \times T),
\end{align*}

where each 0-1 variable $x_{ij}$ is equal to 1 if and only if student $i$ is assigned to teacher $j$.

A successive shortest path method solves problem AP in $O(|S|(|S| + |T|)^2)$ time (see Ahuja 1993 for example).

3. Inverse Assignment Problem

If we use an assignment model proposed in the previous section, we need to determine the weight vector $w$. In this section, we characterize a weight vector which accommodates to a real schedule used in a tutoring school.

Let $\tilde{S} \times T$ be a feasible solution of AP which is an example of real schedule used in a tutoring school. We consider a problem for finding a weight vector $w \in \mathbb{R}^{S \times T}$ satisfying that $\tilde{x}$ is optimal to AP.

A linear relaxation problem (LRP) of AP is a linear programming problem obtained by replacing integrality constraints of AP with non-negativity constraints:

\begin{equation}
\text{LRP: max. } \sum_{(i,j) \in S \times T} w_{ij}x_{ij}
\end{equation}

\text{s.t.} \begin{align*}
\sum_{i \in S} x_{ij} & \leq b \quad (\forall j \in T), \\
\sum_{j \in T} x_{ij} & = 1 \quad (\forall i \in S), \\
x_{ij} & \geq 0 \quad (\forall (i, j) \in S \times T).
\end{align*}

It is well-known that every basic feasible solution of LRP is 0-1 valued, since a coefficient matrix of LRP has total unimodularity. Thus, every optimal basic solution of LRP is also optimal to AP.

The dual of LRP is formulated as follows:

\begin{equation}
D: \text{min. } \sum_{j \in T} y_j + \sum_{i \in S} b_{ij}
\end{equation}

\text{s.t.} \begin{align*}
y_j + y_j & \geq w_{ij} \quad (\forall (i, j) \in S \times T), \\
y_j & \geq 0 \quad (\forall j \in T).
\end{align*}

Let us recall a well-known complementary slackness theorem.

**Theorem 1** (complementary slackness theorem)

Let $(x, y)$ be a pair of solutions feasible to LRP and $D$, respectively. The pair $(x, y)$ is optimal to LRP and $D$, if and only if $(x, y)$ satisfies the conditions that:

- **C1:** $x_{ij} > 0 \rightarrow y_i + y_j = w_{ij}$.
- **C2:** $y_j + y_j > w_{ij} \rightarrow x_{ij} = 0$.
- **C3:** $y_j > 0 \rightarrow \sum_{i \in S} x_{ij} = b$.
- **C4:** $\sum_{i \in S} x_{ij} < b \rightarrow y_j = 0$.

The above theorem directly implies the following.

**Theorem 2**

A feasible solution $\tilde{x}$ of AP is optimal to AP, if and only if there exists a feasible solution $y$ of $D$ satisfying that:

1. $\tilde{x}_{ij} = 1 \rightarrow y_i + y_j = w_{ij}$,
2. $\sum_{i \in S} \tilde{x}_{ij} < b \rightarrow y_j = 0$.

**Proof.** First, we assume that $\tilde{x}$ is optimal to AP. It is well-known that the 0-1 valued solution $\tilde{x}$ is also optimal to LRP (linear relaxation problem). Let $y^*$ be an optimal solution of dual problem $D$. Then, C1 and C4 directly imply properties (1) and (2), respectively.

Next, we consider the case that there exists a feasible solution $y$ of $D$ satisfying above properties (1) and (2). We only need to show that $(\tilde{x}, y)$ satisfies conditions C1, C2, C3 and C4.

C1: Since $\tilde{x}$ is 0-1 valued, property (1) directly implies condition C1.

C2: Property (1) implies that

\begin{align*}
y_i + y_j & \neq w_{ij} \rightarrow \tilde{x}_{ij} \neq 1.
\end{align*}
Clearly, \( y_i + y_j > w_{ij} \rightarrow y_i + y_j \neq w_{ij} \) holds. Since \( x \) is feasible to AP, \( x \) is a 0-1 vector and thus \( x_{ij} \neq 1 \rightarrow x_{ij} = 0 \) holds. From the above, we have Condition C2.

C3: Property (2) implies that
\[
y_j \neq 0 \rightarrow \sum_{i \in S} \tilde{x}_{ij} \geq b.
\]
Clearly, \( y_j > 0 \rightarrow y_j \neq 0 \) holds. Since \( x \) is feasible to AP, \( \sum_{i \in S} \tilde{x}_{ij} \geq b \) implies \( \sum_{i \in S} \tilde{x}_{ij} = b \). Thus, we obtain Condition C3.

C4: Property (2) is equivalent to condition C4.

From the above, a solution \( \tilde{x} \) feasible to AP is optimal to AP, if and only if an inequality system

CP: \[
\begin{align*}
y_i + y_j & \geq w_{ij} \quad (\forall (i, j) \in S \times T), \\
y_j & \geq 0 \quad (\forall j \in T), \\
y_i + y_j & = w_{ij} \quad (\text{if } \tilde{x}_{ij} = 1), \\
y_j & = 0 \quad (\text{if } \sum_{i \in S} \tilde{x}_{ij} < b),
\end{align*}
\]
has a feasible solution \((y, w)\). Consequently, we can find a weight vector \( w \) accommodative to real schedule \( \tilde{x} \) by solving inequality system CP.

4. Finding a Weight Vector

In this section, we discuss a problem for finding an appropriate weight vector \( w \in \mathbb{R}^{S \times T} \) from a set of real schedule data used in a tutoring school. Consider the case that we have \( L \) sets of real schedule data
\[
\{(S^\ell, T^\ell, x^\ell) \mid \ell = 1, 2, \ldots, L\},
\]
where \( S^\ell \) is a set of students, \( T^\ell \) is a set of teachers, and \( x^\ell \) is a solution feasible to AP (defined on pair of sets \( S^\ell \) and \( T^\ell \)).

In our application setting, student-teacher pairs are classified into five categories. We introduce five (pairwise disjoint) subsets
\[
\{E_0^\ell, E_1^\ell, \ldots, E_4^\ell\} \subseteq S^\ell \times T^\ell \quad (\ell = 1, 2, \ldots, L)
\]
declared as follows:

\( E_2^\ell \): set of possible pairs \((i, j) \in S^\ell \times T^\ell \) satisfying that teacher \( j \) is not a supervisor of student \( i \), and \( j \) is a skilled teacher.

\( E_1^\ell \): set of possible pairs \((i, j) \in S^\ell \times T^\ell \) satisfying that teacher \( j \) is a supervisor of student \( i \), and a class at irregular day.

\( E_4^\ell \): set of possible pairs \((i, j) \in S^\ell \times T^\ell \) satisfying that teacher \( j \) is a supervisor of student \( i \), and a class at regular day.

We assume that \( \{E_0^\ell, E_1^\ell, \ldots, E_4^\ell\} \) is a partition of all the pairs in \( S^\ell \times T^\ell \).

In the rest of this paper, we find a weight vector \( w^\ell \in \mathbb{R}^{S^\ell \times T^\ell} \) satisfying conditions that
\[
\begin{align*}
&\exists v = (v_0, v_1, \ldots, v_4) \in \mathbb{R}^5, \\
&0 = v_0 < v_1 < \cdots < v_4 \text{ and } w^\ell_{ij} = v_k \text{ if } (i, j) \in E_k^\ell.
\end{align*}
\]
Given feasible solutions
\[
\tilde{x}^\ell \in \{0, 1\}^{S^\ell \times T^\ell} \quad (\ell = 1, 2, \ldots, L)
\]
of AP defined on a pair of sets of students \( S^\ell \) and teachers \( T^\ell \), we solve the following problem:

IOP: \[
\begin{align*}
\text{min.} & \quad v_4 \\
\text{s.t.} & \quad y_i^\ell + y_j^\ell \geq v_k \quad \forall \ell = 1, 2, \ldots, L, \\
& \quad v_k = 0, 1, \ldots, 4, \\
& \quad \forall (i, j) \in E_0^\ell, \\
& \quad y_j^\ell \geq 0 \quad \forall j \in T^\ell, \\
& \quad \forall \ell = 1, 2, \ldots, L, \\
& \quad \forall (i, j) \in E_{E_1^\ell}^\ell, \\
& \quad y_i^\ell + y_j^\ell = v_k \quad \text{if } x_{ij}^\ell = 1 \quad \text{and } (i, j) \in E_k^\ell, \\
& \quad y_j^\ell = 0 \quad \left(\text{if } \sum_{i \in S} \tilde{x}_{ij}^\ell < b\right), \\
& \quad v_0 \geq 0, \\
& \quad v_0 + 1 \leq v_1, \\
& \quad v_1 + 1 \leq v_2, \\
& \quad v_2 + 1 \leq v_3, \\
& \quad v_3 + 1 \leq v_4.
\end{align*}
\]
Obviously, the last five inequalities are essentially equivalent to the conditions that \( 0 \leq v_0 < v_1 < \cdots < v_4 \).

We have solved problem IOP by using five sets of real schedule data. Each data consists of 10~30 students and 5~10 teachers. An obtained optimal solution is \((v_0, v_1, v_2, v_3, v_4) = (0, 1, 2, 3, 4)\), which indicates that we can
assign a simple weight vector \((0, 1, 2, 3, 4)\) to the above set of five categories.

For finding an assignment, we solve problem AP by setting \(w_{ij} = k\) if \((i, j) \in E_k\) and adding equality constraints
\[
x_{ij} = 0 \ (\forall (i, j) \in E_0).
\]

5. Conclusion

In this paper, we discussed a timetabling problem arising in a tutoring school. We formulate the problem as a variation of assignment problem.

When we apply the assignment model, we need to determine the fitness of each student-teacher pair. We proposed an inverse optimization model for finding fitness values accommodating to real schedule data.

By using a well-known complementary slackness theorem, we can formulate the inverse optimization problem as a linear programming problem.

References


