## ARTICLE IN PRESS

European Journal of Operational Research xxx (2004) xxx-xxx
www.elsevier.com/locate/dsw

## Discrete Optimization

# A factor $\frac{1}{2}$ approximation algorithm for two-stage stochastic matching problems 

Nan Kong, Andrew J. Schaefer *<br>Department of Industrial Engineering, University of Pittsburgh, 1048 Benedum Hall, Pittsburgh, PA 15261, USA

Received 26 November 2003; accepted 22 October 2004


#### Abstract

We introduce the two-stage stochastic maximum-weight matching problem and demonstrate that this problem is $\mathscr{N} \mathscr{P}$-complete. We give a factor $\frac{1}{2}$ approximation algorithm and prove its correctness. We also provide a tight example to show the bound given by the algorithm is exactly $\frac{1}{2}$. Computational results on some two-stage stochastic bipartite matching instances indicate that the performance of the approximation algorithm appears to be substantially better than its worst-case performance.


© 2004 Elsevier B.V. All rights reserved.
Keywords: Stochastic programming; Approximation algorithms; Matching; Combinatorial optimization

## 1. Introduction

Let $G=(V, E)$ be a graph, and let each edge $e \in E$ have an edge weight $c_{e}$. The maximumweight matching problem (Cook et al., 1998) is $\max \left\{\sum_{e \in E} c_{e} x_{e} \mid \sum_{e \in \delta(v)} x_{e} \leqslant 1, \forall v \in V ; x_{e} \in\{0,1\}, \forall_{e} \in E\right\}$.

[^0]It is well known that the maximum-weight matching problem is polynomially solvable (Edmonds, 1965). Consider a stochastic programming extension of this problem as follows. Each edge has two weights, a first-stage weight $c_{\mathcal{e}}$, and a discretely distributed second-stage weight $\tilde{d}_{e}$. The first-stage decision $x$ is to choose a matching in $G$. After the decision, a scenario of the second-stage edge weights is realized. That is, each edge weight is assigned to one of the $r$ possible values $d_{e}^{1}, \ldots, d_{e}^{r}$ with corresponding probabilities $p_{1}, \ldots, p_{r}$. For each scenario $s=1, \ldots, r$, the second-stage decision $y^{s}$ is to choose a matching over those vertices unmatched by the first-stage matching. Without loss of generality, the edge weights $c_{e}$ and $d_{e}^{s}$ for

## ARTICLE IN PRESS

each scenario $s=1, \ldots, r$ are nonnegative, since any edge with negative $c_{e}$ or $d_{e}^{s}$ won't be chosen in any optimal solution. The goal is to maximize the total expected edge weight in these matchings. The stochastic programming extension of (1) can then be written as:
$\max \sum_{e \in E} c_{e} x_{e}+\sum_{s=1}^{r} p_{s} \sum_{e \in E} d_{e}^{s} y_{e}^{s}$
subject to
$\sum_{e \in \delta(v)} x_{e}+\sum_{e \in \delta(v)} y_{e}^{s} \leqslant 1, \quad \forall v \in V, s=1, \ldots, r$,
$x_{e} \in\{0,1\}, y_{e}^{s} \in\{0,1\}, \quad \forall e \in E, s=1, \ldots, r$.

For an introduction to stochastic programming, we refer to Kall and Wallace (1994) and Birge and Louveaux (1997). Interestingly, unlike the polynomially solvable deterministic maximum-weight matching problem, this stochastic programming extension is $\mathscr{N} \mathscr{P}$-complete, as will be shown in Section 2. Therefore, it is natural to develop approximation algorithms that finds solutions with a performance guarantee in a polynomial number of steps for the stochastic programming extension. Hochbaum (1997) and Vazirani (2001) provided surveys of approximation algorithms. There have been very few studies of the computational complexity of stochastic programs and the applications of approximation algorithms to such problems. Dye et al. (2003) studied the computational complexity of the stochastic single-node service provision problem arises from an application of distributed processing in telecommunication networks. They showed the strong $\mathcal{N} \mathscr{P}$-completeness of the problem and presented several approximation algorithms.

The remainder of the paper is organized as follows. In Section 2, we show the $\mathscr{N} \mathscr{P}$-completeness of the stochastic matching problem. In Section 3, we present a factor $\frac{1}{2}$ approximation algorithm and provide a class of instances for which the bound is tight. Section 4 provides computational results that show the performance of the approximation algorithm on a set of randomly generated two-stage stochastic bipartite matching instances.

## 2. The complexity of two-stage stochastic matching

We state the two-stage stochastic matching problem formally.

Instance: Graph $G=(V, E)$, for each $e \in E$, first-stage edge weights $c_{e}$ and second-stage edge weights $d_{e}^{s}$ for $s=1, \ldots, r$, and probability $p_{s}$ for scenario $s$, a positive integer number $r$, and a positive real number $k$.

Question: Are there matchings $M^{0}, M^{1}, \ldots, M^{r}$ in the graph $G$ such that for $s=1, \ldots, r$, $M^{0} \cap M^{s}=\phi, M^{0} \cup M^{s}$ is a matching, and the total expected edge weight given by

$$
\begin{equation*}
\sum_{e \in M^{0}} c_{e}+\sum_{s=1}^{r} p_{s} \sum_{e \in M^{s}} d_{e}^{s} \tag{3}
\end{equation*}
$$

is at least $k$ ?
Theorem 1. Two-stage stochastic matching is $\mathscr{N} \mathscr{P}$ complete.

Aboudi (1986) studied a similar problem, constrained matching, and demonstrated that it is also $\mathscr{N} \mathscr{P}$-complete with a somewhat similar proof.

Proof. Two-stage stochastic matching is clearly in $\mathscr{N} \mathscr{P}$. We assume that for all $s, p_{s}>0$, since any scenario with $p_{s}=0$ may be eliminated. We will use a reduction from CNF-satisfiability to establish the theorem. Let $C$ be an expression in conjunctive normal form with $\rho$ clauses: $C=C_{1} \wedge$ $C_{2} \wedge \cdots \wedge C_{p}$ and $q$ literals $x_{1}, x_{2}, \ldots, x_{q}$. We assume that $x_{i}$ and $\bar{x}_{i}$ do not appear in the same clause, since each clause is a disjunction and thus any clause containing both $x_{i}$ and $\bar{x}_{i}$ is always satisfied. We construct the graph $G$ as follows:

For each $x_{i}$, create vertices $v_{i}, w_{i}$, and $\bar{w}_{i}$. For each $v_{i}$, construct edges ( $v_{i}, w_{i}$ ) and ( $v_{i}, \bar{w}_{i}$ ). For each such edge $e$, let $c_{e}=1$, and let $d_{e}^{s}=0$ for $s=$ $1, \ldots, \rho$. For each $C_{s}$, create a vertex $u_{s}$. For $i=1, \ldots, q$, construct edges ( $w_{i}, u_{s}$ ), and ( $\bar{w}_{i}, u_{s}$ ). For each edge $e=\left(w_{i}, u_{s}\right)$, let $c_{e}=0$, and if $x_{i}$ is in $C_{s}$, let $d_{e}^{s}=\rho$, otherwise, $d_{e}^{s}=0$. For each edge $e=\left(\bar{w}_{i}, u_{s}\right)$, let $c_{e}=0$, and if $x_{i}$ is in $C_{s}$, let $d_{e}^{s}=\rho$, otherwise, $d_{e}^{s}=0$. Define $r \equiv \rho$ and $k \equiv \rho+q$. For $s=1, \ldots, r$, let $p_{s}=\frac{1}{r}$. Note that $G$ is a bipartite graph with bipartition $(V+U, W)$, where
$V, U, W$ are the sets containing all vertices $v_{i}, u_{s}$, and $w_{i}$ and $\bar{w}_{i}$, respectively.

We now claim that $G$ contains matchings $M^{0} \cup M^{s}$ for $s=1, \ldots, r$, and the total expected edge weight given in (3) is at least $k$ if and only if the expression $C$ is satisfiable. To see this we demonstrate the correspondence between matchings with which the value of (3) is at least $k$ and a literal assignment which satisfies $C$.

Suppose that there exists a literal assignment that satisfies $C$. Construct the matchings $M^{0}, M^{1}, \ldots, M^{r}$ as follows.

1. For all $x_{i}$, if $x_{i}$ is True, add $\left(v_{i}, w_{i}\right)$ to $M^{0}$.
2. For all $x_{i}$, if $x_{i}$ is False, add $\left(v_{i}, \bar{w}_{i}\right)$ to $M^{0}$.
3. For all clauses $C_{s}$, pick any literal that satisfies $C_{s}$. If $x_{i}$ is chosen, add $\left(u_{s}, \bar{w}_{i}\right)$ to $M^{s}$. If $\bar{x}_{i}$, is chosen, add $\left(u_{s}, w_{i}\right)$ to $M^{s}$.

It is easy to check that for $s=1, \ldots, r$, $M^{0} \cap M^{s}=\emptyset$ and $M^{0} \cup M^{s}$ is a matching, and the total expected edge weight is $k$.

Now let us suppose that there exist matchings $M^{0}, M^{1}, \ldots, M^{r}$ such that for $s=1, \ldots, r$, $M^{0} \cap M^{s}=\emptyset$ and $M^{0} \cup M^{s}$ is a matching, and the total expected edge weight is at least $k$. Note that no more than $q$ edges with $c_{e}>0$ can be in $M^{0}$, and $c_{e}=1$ for all such edges. Also, note that for each $s$, no more than one edge with $d_{e}^{s}>0$ can be in $M^{s}$, and $d_{e}^{s}=\rho$ for this edge. Hence, $\sum_{e \in M^{0}} c_{e} \leqslant q$ and $\sum_{e \in M^{s}} p_{s} d_{e}^{s} \leqslant 1$. The latter inequality implies that $\sum_{s=1}^{r} \sum_{e \in M^{s}} p_{s} d_{e}^{s} \leqslant \rho$ and thus the value of (3) is at most $q+\rho=k$. Since the total expected edge weight is at least $k$, it follows that $M^{0}$ matches every vertex in $V$ with a weight 1 edge and each $M^{s}$ matches vertex $u_{s}$ with a positively weighted edge.

Consider any literal in $C$, we construct the literal assignment as follows.

1. If $\left(v_{i}, w_{i}\right)$ is in $M^{0}, x_{i}$ is True.
2. If ( $v_{i}, \bar{w}_{i}$ ) is in $M^{0}, x_{i}$ is False.

It is easy to check that this literal assignment satisfies $C$.

The above transformation is clearly polynomial, so we conclude that Two-stage stochastic matching is $\mathscr{N} \mathscr{P}$-complete.

Table 1
Edge weights of the graph constructed from $C$

| $e$ | $c_{e}$ | $d_{e}^{1}$ | $d_{e}^{2}$ |
| :--- | :--- | :--- | :--- |
| $\left(v_{1}, w_{1}\right)$ | 1 | 0 | 0 |
| $\left(v_{1}, \bar{w}_{1}\right)$ | 1 | 0 | 0 |
| $\left(v_{2}, w_{2}\right)$ | 1 | 0 | 0 |
| $\left(v_{2}, \bar{w}_{2}\right)$ | 1 | 0 | 0 |
| $\left(w_{1}, u_{1}\right)$ | 0 | 2 | 0 |
| $\left(\bar{w}_{1}, u_{1}\right)$ | 0 | 0 | 0 |
| $\left(w_{2}, u_{1}\right)$ | 0 | 0 | 0 |
| $\left(\bar{w}_{2}, u_{1}\right)$ | 0 | 2 | 0 |
| $\left(w_{1}, u_{2}\right)$ | 0 | 0 | 2 |
| $\left(\bar{w}_{1}, u_{2}\right)$ | 0 | 0 | 0 |
| $\left(w_{2}, u_{2}\right)$ | 0 | 0 | 2 |
| $\left(\bar{w}_{2}, u_{2}\right)$ | 0 | 0 | 0 |

### 2.1. Example of the reduction

Consider the expression
$C=\left\{\bar{x}_{1} \vee x_{2}\right\} \wedge\left\{\bar{x}_{1} \vee \bar{x}_{2}\right\}$.
There are two literals and two clauses, so $r=\rho=2, q=2$ and $k=4$. Then $G$ is as in Fig. 1 and the edge weights are as in Table 1. All edge weights are also labeled in Fig. 1. The two scenarios are assigned with equal probability.


Fig. 1. Example of the graph $G$ constructed from the expression $C$ given in (4). The edge weights are represented by $\left(c_{e}, d_{e}^{1}, \ldots, d_{e}^{r}\right)$.

## ARTICLE IN PRESS

From Theorem 1, there exist matchings $M^{0}, M^{1}$ and $M^{2}$ such that $M^{0} \cap M^{1}=\emptyset, M^{0} \cap M^{2}=\emptyset$, and $M^{0} \cup M^{1}$ and $M^{0} \cup M^{2}$ are matchings and
$\sum_{e \in M^{0}} c_{e}+\frac{1}{2} \sum_{e \in M^{1}} d_{e}^{1}+\frac{1}{2} \sum_{e \in M^{2}} d_{e}^{2} \geqslant 4$
if and only if there exists a literal assignment satisfying $C$.

Matchings $\quad M^{0}=\left\{\left(v_{1}, \bar{w}_{1}\right),\left(v_{2}, \bar{w}_{2}\right)\right\}, \quad M^{1}=$ $\left\{\left(w_{1}, u_{1}\right)\right\}$ and $M^{2}=\left\{\left(w_{2}, u_{2}\right)\right\}$ have a total expected weight of $\left(1+1+2 \times \frac{1}{2}+2 \times \frac{1}{2}=4\right)$, and these matchings correspond to the assignment of literal $x_{1}$ to False and literal $x_{2}$ to False which satisfies $C$. Note that $M^{0} \cup M^{1}$ and $M^{0} \cup M^{2}$ are matchings. An alternative is matchings $M^{0}=$ $\left\{\left(v_{1}, \bar{w}_{1}\right),\left(v_{2}, w_{2}\right)\right\}, M^{1}=\left\{\left(w_{1}, u_{1}\right)\right\}, \quad$ and $\quad M^{2}=$ $\left\{\left(\bar{w}_{2}, u_{2}\right)\right\}$, which also have a total expected weight of 4 , and correspond to the assignment of literal $x_{1}$ to False and literal $x_{2}$ to True which satisfies $C$ as well.

## 3. A factor $\frac{1}{2}$ approximation algorithm

Definition 1. A first-stage myopic solution is an optimal solution to:

$$
\begin{gathered}
\text { (MYOPIC1): } \max \left\{\sum_{e \in E} c_{e} x_{e} \mid \sum_{e \in \delta(v)} x_{e} \leqslant 1, \forall v \in V ;\right. \\
\left.x_{e} \in\{0,1\}, \forall e \in E\right\} .
\end{gathered}
$$

Definition 2. A second-stage myopic solution for scenario $s$ is an optimal solution to:
(MYOPIC2) : $\max \left\{\sum_{e \in E} d_{e}^{s} y_{e} \mid \sum_{e \in \delta(v)} y_{e} \leqslant 1, \forall v \in V ;\right.$

$$
\left.y_{e} \in\{0,1\}, \forall e \in E\right\} .
$$

A first- (second-) stage myopic solution is the solution to a deterministic maximum-weight matching problem with the appropriate choice of objective.

The intuition behind our approximation algorithm is straightforward. We consider $r+1$ solutions: one first-stage myopic solution, and $r$ second-stage myopic solutions for all scenarios. We compare two objective values: the objective value of the first-stage myopic solution, and the expected objective value of the second-stage myopic solutions over all scenarios. Of these two values, the larger one gives the output of the approximation algorithm.

We state the algorithm formally:

## Algorithm 1

INPUT: A two-stage stochastic maximum-weight matching problem.
Let $x_{1}$ be a first-stage myopic solution, and let $z_{1}=c x_{1}$.
For scenario $s$, Let $y_{2}^{s}$ be a second-stage myopic solution, and let $z_{2}^{s}=d^{s} y_{2}^{s}$.
Let $\hat{z}=\max \left\{z_{1}, \sum_{s=1}^{r} p_{s} z_{2}^{s}\right\}$.
OUTPUT: If $\hat{z}=z_{1}$, then return $\left(x_{1}, \mathbf{0}, \ldots, \mathbf{0}\right)$ and $z_{1}$; otherwise, return $\left(\mathbf{0}, y_{2}^{1}, \ldots, y_{2}^{r}\right)$ and $\sum_{s=1}^{r} p_{s} z_{2}^{s}$.

Theorem 2. Algorithm 1 is an approximation algorithm with performance guarantee $\frac{1}{2}$ for the twostage stochastic maximum-weight matching problem given in (2).

Proof. Solutions $\left(x_{1}, \mathbf{0}, \ldots, \mathbf{0}\right)$ and $\left(\mathbf{0}, y_{2}^{1}, \ldots, y_{2}^{\prime}\right)$ are clearly feasible to (2). Let $x^{*}=\left(x_{0}, y_{0}^{1}, \ldots, y_{0}^{r}\right)$ and $z^{*}$ be an optimal solution and the optimal objective value to (2), respectively. Since solution $x_{0}$ is feasible to (MYOPIC1),

$$
\begin{equation*}
\hat{z} \geqslant z_{1} \geqslant c x_{0} . \tag{5}
\end{equation*}
$$

Since solution $y_{0}^{s}$ is feasible to (MYOPIC2) for scenario $s, s=1, \ldots, r$,
$\hat{z} \geqslant \sum_{s=1}^{r} p_{s} z_{2}^{s} \geqslant \sum_{s=1}^{r} p_{s} d^{s} y_{0}^{s}$.
Summing up inequalities (5) and (6) yields $2 \hat{z} \geqslant c x_{0}+\sum_{s=1}^{r} p_{s} d^{s} y_{0}^{s}=z^{*}$, and thus the result follows.

Since both (MYOPIC1) and (MYOPIC2) are polynomially solvable, Algorithm 1 runs in polynomial time.

## ARTICLE IN PRESS

### 3.1. A tight example for Algorithm 1

We give a tight example of two-stage stochastic bipartite matching. The problem is defined on the graph $G=(V, E)$ as in Fig. 2 and its objective function is given as in (2).

Let $G$ be a bipartite graph with bipartition $V=(S, T)$ where $|S|=|T|$ and furthermore let $S=\left(S_{1}, S_{2}\right)$ with $\left|S_{1}\right|=\left|S_{2}\right|$ and let $T=\left(T_{1}, T_{2}\right)$ with $\left|T_{1}\right|=\left|T_{2}\right|$. Let $l$ be any positive integer. For all edges $e=(u, v)$ with $u \in S_{1}$ and $v \in T_{1}$, let $c_{e}=l$. For all other edges connecting $S$ and $T$, let $c_{e}=0$. For any scenario $s, s=1, \ldots, r$, for all edges $e=(u, v)$ with $u \in S_{2}$ and $v \in T_{2}$ let $d_{e}^{s}=l$; for all other edges connecting $S$ and $T$, let $d_{e}^{s}=0$. Fig. 2 illustrates such an instance with $|S|=|T|=4$.

The first-stage myopic solution is given by choosing any complete matching from $S_{1}$ to $T_{1}$, together with any matching from $S_{2}$ to $T_{2}$. Hence, the output of Algorithm 1 is to use the first-stage myopic solution in the first stage and to choose no edges in the second stage. This output gives the total expected edge weight $l \cdot \frac{|V|}{4}$. For any scenario, the second-stage myopic solution is given by choosing any complete matching from $S_{2}$ to


Fig. 2. Example on the complete bipartite graph $G$ when the bound is tight. Edges with presence have edge weights. Solid edges are weighted $l$ in the first stage and dashed edges are weighted $l$ in the second stage. All other edges have zero weight in both stages. $S_{1}=\left\{u_{1}, u_{2}\right\}, S_{1}=\left\{u_{3}, u_{4}\right\}, T_{1}=\left\{v_{1}, v_{2}\right\}$, $T_{1}=\left\{v_{\mathrm{e}}, v_{4}\right\}$.
$T_{2}$, together with any matching from $S_{1}$ to $T_{1}$. Hence, the output of Algorithm 1 is to choose no edges in the first stage and to use the $s$ th secondstage myopic solution in the second stage if scenarios $s$ is realized. This output gives the total expected edge weight $l \cdot \frac{|V|}{4}$. The maximum of these two solutions is $l \cdot \frac{|V|}{4}$, so the approximation algorithm gives a solution of $l \cdot \frac{|V|}{4}$.

The optimal solution to the two-stage stochastic bipartite matching problem is to choose any complete matching from $S_{1}$ to $T_{1}$ as the first-stage decision, and for each scenario, choose any complete matching from $S_{2}$ to $T_{2}$ as the second-stage decision. The first-stage matching gives edge weight $l \cdot \frac{|V|}{4}$, and the expected second-stage edge weight is $l \cdot \frac{|V|}{4}$, so the total expected edge weight is $l \cdot \frac{|V|}{2}$. Thus the approximation algorithm returns a solution whose total expected edge weight is exactly $\frac{1}{2}$ of the optimal objective value.

## 4. Computational results

We tested our approximation algorithm on a set of randomly generated two-stage stochastic bipartite matching instances with 10 vertices in each side of the bipartition and 100 scenarios. In our computational experiments, we used CPLEX 7.0 to find the solutions to (MYOPIC1) and (MYOPIC2). To check the performance of our approximation algorithm, we also solved the stochastic programming formulation directly using the Lshaped method (Van Slyke and Wets, 1969), a variant of Benders' decomposition (Benders, 1962) and a standard technique for exactly solving twostage stochastic linear programs with continuous or integer first stage. For some large instances, the L-shaped method tended to be very time-consuming, so we imposed a 1 -hour CPU time limit on it and obtained the solution of the restricted master problem, which is an upper bound on the exact solution.

In all test instances, the first-stage and secondstage edge weights were normally distributed. All scenarios were realized with equal probability. We tested four groups of instance classes, in each of which only one of the four distribution parameters (mean and standard deviation of the first-

## ARTICLE IN PRESS

stage and second-stage edge weights) was varied and other three were fixed. For example, in group 1, we varied the mean of the first-stage edge weights from 5 to 25 . We generated 100 instances for each instance class and reported the average CPU time. Table 2 presents the characteristics and computational results of these instance classes. Our computational experiments indicate that the CPU time of the approximation algorithm is insensitive to the distribution parameter settings. On the other hand, when the first-stage and sec-
ond-stage edge weights are generated from the same or similar distributions, the L-shaped method is relatively less efficient due to the symmetry between the edge weights in the two stages. We also report the average ratio of the approximation solution to the exact solution in the table. When the first-stage and second-stage edge weights are generated from the same or similar distributions, this ratio tends to be the lowest for the same reason. The last column in Table 2 shows the integrality gap of the LP-relaxation for these instances.

Table 2
Characteristics and computational results of small instances

| Group | Instance |  | Average approx. CPU | Average <br> L-shaped CPU | Average \# of L-shaped iterations | Average performance ratio | Average integrality gap (\%) |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | First stage | Second stage |  |  |  |  |  |
| 1 | $\sim N\left(5,15^{2}\right)$ | $\sim N\left(10,15^{2}\right)$ | 0.08 | 0.71 | 13.39 | 0.976 | 0.1 |
|  | $\sim N\left(10,15^{2}\right)$ | $\sim N\left(10,15^{2}\right)$ | 0.08 | 1.88 | 25.35 | 0.958 | 0.3 |
|  | $\sim N\left(15,15^{2}\right)$ | $\sim N\left(10,15^{2}\right)$ | 0.08 | 2.25 | 23.70 | 0.984 | 0.3 |
|  | $\sim N\left(20,15^{2}\right)$ | $\sim N\left(10,15^{2}\right)$ | 0.08 | 0.41 | 9.05 | 0.998 | 0.0 |
|  | $\sim N\left(25,15^{2}\right)$ | $\sim N\left(10,15^{2}\right)$ | 0.08 | 0.14 | 4.79 | 1.000 | 0.0 |
| 2 | $\sim N\left(10,5^{2}\right)$ | $\sim N\left(10,15^{2}\right)$ | 0.08 | 0.10 | 3.01 | 1.000 | 0.0 |
|  | $\sim N\left(10,10^{2}\right)$ | $\sim N\left(10,15^{2}\right)$ | 0.08 | 0.34 | 7.20 | 0.997 | 0.0 |
|  | $\sim N\left(10,15^{2}\right)$ | $\sim N\left(10,15^{2}\right)$ | 0.08 | 1.88 | 25.35 | 0.958 | 0.3 |
|  | $\sim N\left(10,20^{2}\right)$ | $\sim N\left(10,15^{2}\right)$ | 0.08 | 1.21 | 18.93 | 0.966 | 0.2 |
|  | $\sim N\left(10,25^{2}\right)$ | $\sim N\left(10,15^{2}\right)$ | 0.08 | 0.53 | 11.54 | 0.978 | 0.0 |
| 3 | $\sim N\left(10,15^{2}\right)$ | $\sim N\left(5,15^{2}\right)$ | 0.08 | 1.81 | 21.97 | 0.980 | 0.1 |
|  | $\sim N\left(10,15^{2}\right)$ | $\sim N\left(10,15^{2}\right)$ | 0.08 | 1.88 | 25.35 | 0.958 | 0.3 |
|  | $\sim N\left(10,15^{2}\right)$ | $\sim N\left(15,15^{2}\right)$ | 0.08 | 0.79 | 14.02 | 0.980 | 0.1 |
|  | $\sim N\left(10,15^{2}\right)$ | $\sim N\left(20,15^{2}\right)$ | 0.08 | 0.32 | 7.20 | 0.994 | 0.0 |
|  | $\sim N\left(10,15^{2}\right)$ | $\sim N\left(25,15^{2}\right)$ | 0.08 | 0.15 | 4.11 | 0.998 | 0.0 |
| 4 | $\sim N\left(10,15^{2}\right)$ | $\sim N\left(10,5^{2}\right)$ | 0.08 | 0.16 | 5.74 | 0.989 | 0.0 |
|  | $\sim N\left(10,15^{2}\right)$ | $\sim N\left(10,10^{2}\right)$ | 0.08 | 0.82 | 15.20 | 0.976 | 0.1 |
|  | $\sim N\left(10,15^{2}\right)$ | $\sim N\left(10,15^{2}\right)$ | 0.08 | 1.88 | 25.35 | 0.958 | 0.3 |
|  | $\sim N\left(10,15^{2}\right)$ | $\sim N\left(10,20^{2}\right)$ | 0.09 | 0.97 | 12.66 | 0.991 | 0.0 |
|  | $\sim N\left(10,15^{2}\right)$ | $\sim N\left(10,25^{2}\right)$ | 0.09 | 0.19 | 4.86 | 0.999 | 0.0 |

Table 3
Characteristics and computational results of larger instances

| Instance class | Weights |  | Average <br> approx. CPU | Average <br> L-shaped CPU | Average \# of <br> L-shaped iterations | Average <br> performance ratio |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
|  | First stage | Second stage |  | $\geqslant 3600$ | $\geqslant 3$ | 0.954 |
| 1 | $\sim N\left(10,15^{2}\right)$ | $\sim N\left(10,15^{2}\right)$ | 160.4 | $\geqslant 2600$ | $\geqslant 3$ | 0.998 |
| 2 | $\sim N\left(10,15^{2}\right)$ | $\sim N\left(10,20^{2}\right)$ | 144.9 | $\geqslant N\left(10,30^{2}\right)$ | 131.8 | 197.0 |
| 3 | $\sim N\left(10,15^{2}\right)$ | $\sim N\left(15,15^{2}\right)$ | $\sim N\left(10,15^{2}\right)$ | 161.3 | $\geqslant 3600$ | $\geqslant 3$ |
| 4 | $\sim N\left(20,15^{2}\right)$ | $\sim N\left(10,15^{2}\right)$ | 159.7 | $\geqslant 3600$ | $\geqslant 3$ | 0.986 |
| 5 | $\sim N\left(20,15^{2}\right)$ | $\sim N\left(10,30^{2}\right)$ | 132.3 | 191.4 | 2 | 0.997 |
| 6 |  |  |  |  | 1.000 |  |

We then considered some large stochastic bipartite instances with 500 vertices in each side of the bipartition and 10 scenarios. Table 3 presents the characteristics and computational results of these instances. Each instance class consists of 10 instances. As above, each random parameter was generated according to a normal distribution and each scenario was assigned with equal probability. In the table, we also report the average ratio of the approximate solution to the exact solution or its upper bound if the L-shaped method did not terminate within 1 hour.

## 5. Conclusions

As we have shown in this paper, the stochastic programming extension of a polynomially solvable combinatorial optimization problem may become $\mathscr{N} \mathscr{P}$-complete. However, the line between easy and hard stochastic combinatorial optimization problems has yet to be fully explored. Meanwhile, given difficulty of solving stochastic programs, particularly stochastic integer programs, developing approximation algorithms for such problems is a promising direction for future research.

## Acknowledgment

This research was partially supported by National Science Foundation grant DMI-0217190
and a grant from the University of Pittsburgh Central Research Development Fund.

## References

Aboudi, R., 1986. A constrained matching program: A polyhedral approach. Ph.D. Thesis, Cornell University, Ithaca, NY.
Benders, J.F., 1962. Partitioning procedures for solving mixed variables programming problems. Numerische Mathematik 4, 238-252.
Birge, J.R., Louveaux, F.V., 1997. Introduction to Stochastic Programming. Springer, New York.
Cook, W.J., Cunningham, W.H., Pulleyblank, W.R., Schrijver, A., 1998. Combinatorial Optimization. John Wiley and Sons, New York.
Dye, S., Stougie, L., Tomasgard, A., 2003. Approximation algorithm and relaxations for a service provision problem on a telecommunication network. Discrete Applied Mathematics 129 (1), 63-81.
Edmonds, J., 1965. Maximum matching and a polyhedron with $0-1$ vertices. Journal of Research of the National Bureau of Standards 69B, 125-130.
Hochbaum, D.S., 1997. Approximation Algorithms for $\mathscr{N} \mathscr{P}$-Hard Problems. PWS Publishing Company, Boston, MA.
Kall, P., Wallace, S.W., 1994. Stochastic Programming. John Wiley and Sons, Chichester, UK.
Van Slyke, R., Wets, R.J.-B., 1969. L-shaped linear programs with applications to optimal control and stochastic programming. SIAM Journal on Applied Mathematics 17, 638663.

Vazirani, V., 2001. Approximation Algorithms. Springer-Verlag, New York.


[^0]:    * Corresponding author. Tel.: +1412624 5045; fax: +1412 6249831.

    E-mail address: schaefer@ie.pitt.edu (A.J. Schaefer).

