Discussion Paper No. 660

PROBABILISTIC ANALYSIS OF THE CAPACITATED TRANSPORTATION PROBLEM

by

Refael Hassin*
and
Eitan Zemel**†

August 1985

---

*Department of Statistics, Tel Aviv University, Tel Aviv 69978, Israel.

**Department of Managerial Economics and Decision Sciences, J. L. Kellogg Graduate School of Management, Northwestern University, 2001 Sheridan Road, Evanston, Illinois 60201.

†The research of this author was supported in part by the National Science Foundation under Grant ECS-812141.

Abstract

We consider the capacitated transportation problem defined by sets of supplies a_i, $i \in I$, demands b_j, $j \in J$, and capacities c_{ij}, $i \in I$, $j \in J$. Assuming that the capacities are random variables, we prove asymptotic conditions on the supplies and demands which assure that a feasible solution exists almost surely. The proof is constructive and supplies an algorithm whose running time is $O(|I|\ |J|)$. We then apply the results to the maximum flow problem.

1. Introduction

Probabilistic analysis of combinatorial problems is the subject of many recent investigations (c.f. the survey by Karp, Lenstra, McDiarmid and Rinnooy Kan (1984)). Some of these papers deal with conditions that assure with high probability the existence of feasible solutions to such problems. For example, Shamir and Upfal (1981) prove such results for f-factors, Frieze (1984a) for perfect matchings, Pósa (1976), Komlós amd Szemerédi (1983), Fenner and Frieze (1983), Bollobás (1983), and Frieze (1984b) for Hamiltonian cycles, Loulou (1982) and Karmarkar (1982) for perfect packings and Vercellis (1984) for set covers.

In this paper we consider another fundamental, combinatorial problem, the capacitated transportation problem (CTP). In the version studied below, we seek to find a feasible solution (or to prove that none exists) to the following set of equations:

$$\sum_{i\in I} x_{ij} = b_j, \quad j \in J$$

$$\sum_{j\in J} x_{ij} = a_i, \quad i \in I$$

$$0 \leqslant x_{ij} \leqslant c_{ij}, \; i \in I, \; j \in J$$

It is well-known that

$$\sum_{i\in I} a_i = \sum_{j\in J} b_j$$

is a necessary condition for feasibility which is also sufficient if the supplies, a, and the demands, b, are relatively "small" compared to the capacities C. But how small is "small enough?" This problem was first treated by Doulliez and Jamoulle (1972). They consider a general (i.e., not necessarily bipartite) network with independent discrete random arc capacities and flow supplies and demands at its nodes. They present an elegant algorithm which computes the probability that a feasible solution exists by sequentially generating feasible and nonfeasible subsets of the capacity state space.

In this paper we consider the asymptotic behavior of CTP where the elements c_{ij} are random. Surprisingly, it turns out that under very general conditions on the distribution of C, a and b can be "very large" provided they are "evenly spread." Our proof is constructive and provides an algorithm whose running time is linear in the number of elements of the capacity matrix C and which almost surely produces a feasible solution whenever the conditions stated in the theorem hold. The algorithm employs a novel method for scanning the rows and columns of the transportation matrix.

We then apply our results to analyze the maximum flow problem in a complete or random graph with random capacities. This problem was considered earlier by Frank and Hakimi (1965), (1967), (1968), and by Frank and Frisch (1971). There, both the problems of testing hypotheses on this value and of computing its distribution are considered. It is found that exact computation of the probability distribution of the max flow value is a formidable task. However, Karp (1979), Grimmett and Welsh (1982), and Grimmett and Suen (1982) obtained very strong asymptotic results for complete graphs with i.i.d. capacities. In particular, they have shown that the minimum cut is almost surely either the star emanating from the source or the star entering the sink. This result, for a much more general probabilistic model, is an immediate byproduct of our theorem on CTP. Moreover, our approach for the proof is different since it works directly on the maximum flow rather than on the minimum cut. Thus, it can be used constructively to generate in linear time an exact maximum flow. (As pointed out by Picard and Queryranne (1982), there is an algorithmic assymetry between minimum cuts and maximum flows. Given a maximum flow, a minimal cut can be easily identified but the availability of a minimum cut does not seem to help if one seeks the maximum flow.)

The structure of the paper is as follows. The problem and our main theorem are stated in Section 2. Section 3 contains the algorithm and an outline of the proof of the theorem. Section 4 addresses the maximum flow problem. Finally, in Section 5, we give the necessary details to complete the proof of the theorem.

2. The Capacitated Transportation Problem

Let n be a parameter describing the size of problem CTP. Let α, β, γ, δ, ε be constants and consider problem instances of CTP which satisfy, for $n = 1,2,3,\ldots$, the following five conditions:

(a) $$0 \leq a_i \leq n^{\alpha}, \quad i \in I$$

(b) $$0 \leq b_j \leq n^{\beta}, \quad j \in J$$

(c) $$\sum_{i \in I} a_i = \sum_{j \in J} b_j = n^{\gamma}$$

(d) $$|I| \leq n^{\varepsilon}$$

(e) $$|J| \leq n^{\delta}$$

A set $C = \{c_k\colon k \in K\}$ of nonnegative random variables is called proper with constant $c > 0$ if for each $y \geqslant 0$ and each $k \in K$

$$\Pr = [\ c_k \leqslant y: S] \leqslant cy$$

where the conditioning event S is any event concerning the random variables c_ℓ, $\ell \neq k$. A collection of sets $\{C^n\}$, $n = 1,2,\ldots$ is called proper if all its members are proper with the same constant c. A typical example of a proper collection of random matrices is where for each n, all c^n_{ij} are i.i.d. random variables with a common continuous probability distribution F which does not depend on n and which satisfies $F'(0) < \infty$ (cf. Frieze (1985), Frieze and Grimmette (1985), Hassin and Zemel (1984)). However, the concept of properness is much more general. In particular, it allows for variables which are not independent, which are not identically distributed, and where distribution is not independent on n. All which is required is that the random variables involved will not be "too small" with too high a probability. For instance, any collection of positive random variables whose support is uniformly bounded away from zero is proper.

Let $\alpha^+ = \max\{\alpha, 0\}$. Then we have:

Theorem 1. Assume that $\gamma > \alpha^+ + \beta + \max\{2\varepsilon/3,\ (\varepsilon + \beta)/2,\ \beta\}$ and that the matrices (c^n_{ij}) form a proper collection. Then CTP is feasible almost surely.

Remark 1. We point out that the randomness presented in our formulation is limited to the capacities C only. In particular, the supplies and demands (a and b) are treated as given constants. This allows for extra flexibility which is useful in the analysis of Section 3.

Remark 2. The reader may observe that the proof of Theorem 1 holds with trivial modifications to cases where only a part of the martix C is proper.

In particular, we mention the cases where the underlying graph is random where each edge exists independently with probability $p > 0$. The requirement in this case is that the set of capacities associated with existing edges be proper. Another interesting case is where for each pair $\{i,j\}$ only one edge (i,j) or (j,i) exists and the choice is made randomly and independently with probability $0 < p < 1$. This case is used in the analysis of the maximum flow problem on randomly directed graphs.

We devote the next section to an outline of a proof of Theorem 1. In fact, we describe an algorithm which, under the conditions stated, almost surely finds a feasible solution for CTP.

3. The Algorithm

The basic step of the algorithm is a column or row scan. For any two nonnegative vectors $q = (q_k: k \in K)$, $r = (r_k: k \in K)$ and two constants $s \geqslant t \geqslant 0$, we define the following procedure:

```
Scan (q, r, s, t)
     begin Scan
          for k ∈ K
               x = min{q_k, r_k, s - t}
               q_k = q_k - x
               r_k = r_k - x
               s = s - x
          end for
     end Scan.
```

A scan is called successful if it terminates with $s = t$. Obviously, a necessary and sufficient condition for success is that

$$s - t \leqslant \sum_{k \in K} \min\{q_k, r_k\}.$$

A naive approach for solving CTP is the "northwest corner" method, adopted in the obvious way to account for capacities:

```
NW (C,a,b)
    begin NW
        for j ∈ J,
            Scan (a, c_{I,j}, b_j, 0)        :scan the jth column of C
        end for
    end NW
```

This may work well for the first several columns, but eventually, as the row supplies are being depleted, the procedure is likely to fail. The procedure we develop below is a modification of this method which is designed to allocate "sufficient" supplies to enable a successful scan of the "last" columns of the matrix C. It is based on the following trivial observation:

Lemma 1. Assume that $b_j \leq \min\{c_{ij}: i \in I\}$ for every $j \in J$. Then, procedure NW yields a feasible solution for CTP.

Matrices which satisfy the stipulations of Theorem 1 are unlikely to satisfy the stipulations of Lemma 1. To overcome this difficulty, we partition the rows of C into two subsets I_1 and I_2, denoting the two resulting submatrices by C_1 and C_2 respectively. We first scan the columns of C_1 "setting aside" a certain portion $\bar{b}_j \leq b_j$ of each column demand so that Lemma 1 holds for C_2. Specifically, Lemma 1 is applicable to C_2 if we let $\bar{b}_j = \min\{b_j, \{\min c_{ij}: i \in I_2\}\}$. Since the supply and demand of C_1 are generally not balanced, we scan this matrix again, this time going over it row by row. Finally, we apply NW to C_2. The trick is to choose I_1 and I_2 so that the row and column scans of C_1 are almost surely successful. Before showing how this can be done, we summarize our algorithm as procedure Solve below:

```
     Solve (C, a, b, I_1, I_2)
          begin Solve
               for j ∈ J
(1)                 b̄_j = min{b_j, min{c_ij: i ∈ I_2}}      :Set aside demands for C_2
               end for
               for j ∈ J
(2)                 Scan (a, c_{I_1,j}, b_j, b̄_j),         :Scan column j of C_1
               end for
               for i ∈ I_1
(3)                 Scan (b, c_{i,J}, a_i, 0)              :Scan row i of C_1
               end for
(4)                 NW (C_2, a, b)                         :Scan the colums of C_2
          end Solve
```

We now examine the partition of I into (I_1, I_2) so that Steps (2) and (3) of Solve are likely to succeed. As mentioned previously, our choice of $\bar{b}_j$ in (1) is such that Step (4) is guaranteed to succeed by Lemma 1. The conditions of Theorem 1 imply the existence of a constant ξ satisfying the stipulations of the following lemma and therefore this lemma constitutes a proof of Theorem 1.

<u>Lemma 2.</u> Let I_2 be the set of indices of the n^{ξ} largest elements of a. Assume that $\xi > 0$ satisfies conditions (i), (ii), and (iii) below. Then, procedure Solve succeeds almost surely.

(i) $\quad \xi < (\gamma - \alpha - \beta)/2$

(ii) $\quad \xi < \gamma - \alpha - 2\beta$

(iii) $\quad \xi > \alpha^+ + \beta + \varepsilon - \gamma$

Outline of Proof. As the algorithm progresses, shipments x_{ij} are allocated to routes. Denote the residual supplies, demands, and capacities by a', b' and C', respectively. We use the standard order notation to compare the asymptotic growth of functions, e.g., $f(n) = \Omega(g(n))$ if there exists a constant $c > 0$ such that $f(n) \geq cg(n)$ for every $n \geq 1$.

We have to show that steps (2) and (3) of Solve are successful. Below we consider each of these steps.

Step 2: Scan of column j of C_1

This scan is successful if

$$(5) \qquad \sum_{i \in I_1} \min \{a_j', c_{ij}\} \geq b_j - \bar{b}_j$$

First note that $a_i' = a_i$ for $i \in I_2$ and thus

$$\sum_{i \in I_1} a_i' + \sum_{i \in I_2} a_i = \sum_{i \in I} a_i' = \sum_{j \in J} b_j' \geq \sum_{j \in J} \bar{b}_j$$

so that

$$(6) \qquad \sum_{i \in I_1} a_i' \geq \sum_{j \in J} \bar{b}_j - \sum_{i \in I_2} a_i$$

We estimate the magnitude of each term in the right side of (6) separately. First consider

$$\sum_{j \in J} \bar{b}_j = \sum_{j \in J} \min \{b_j, \min \{c_{ij}: i \in I_2\}\}$$

$\mathrm{Min}\{c_{ij}: i \in I_2\}$ is of order $n^{-\xi}$. The worst case for $\sum_{j\in J} \bar{b}_j$ is when $b_j = n^{\beta}$ for $n^{\gamma-\beta}$ columns and zero elsewhere. This yields

(7) $$\sum_{j\in J} \bar{b}_j = \Omega(n^{\gamma-\beta-\xi})$$

with high probability (detailed proof is given in Section 5). For the second term, observe that $|a_i| \leqslant n^{\alpha}$ and that $|I_2| = n^{\xi}$. Thus,

$$\sum_{i\in I_2} a_i \leqslant n^{\xi+\alpha}.$$

Since $\gamma - p - \xi > \xi + \alpha$ (condition (i)), (6) yields

$$\sum_{i\in I_1} a_i' = \Omega(n^{\gamma-\beta-\xi}).$$

Consider now the left-hand side of (5). The worst case is where $a_i' = n^{\alpha}$ for $\Omega(n^{\delta-\beta-\xi-\alpha})$ rows and zero elsewhere. This yields with high probability

(8) $$\sum_{i\in I_1} \min\{a_i', c_{ij}\} = \Omega(n^{\gamma-\beta-\xi-\alpha}).$$

Since $\gamma - \beta - \xi - \alpha > \beta$ (condition (ii)), the right hand side of this expression is more than n^{β} which in turn is more than $b_j - \bar{b}_j$. Therefore, (5) holds with high probability.

Step 3: Scan of row i of C_1. This scan is successful if:

(9) $$\sum_{j\in J} \min\{b_j', c_{ij}'\} \geqslant a_i'.$$

We want to use similar techniques to those of step 2. However, c'_{ij}, having been depleted, are no longer proper. To overcome this, let x_{ij} be the shipment in route (i,j). Add x_{ij} to each of the terms in (9) to get an equivalent expression:

$$\sum_{j\in J} \min \{b'_j + x_{ij}, c'_{ij} + x_{ij}\} \geq a'_i + \sum_{j\in J} x_{ij}.$$

But $c'_{ij} + x_{ij} = c_{ij}$, and $a'_i + \sum_{j\in J} x_{ij} = a_i$. Therefore, (9) is implied by the stronger requirement:

$$(10) \qquad \sum_{j\in J} \min \{b'_j, c_{ij}\} \geq a_i$$

Expression (10) involves the original matrix C which is proper. As for the b'_j, $j \in J$, these satisfy

$$\sum_{j\in J} b'_j \geq \sum_{i\in I_2} a_i \geq n^{\xi+\gamma-\varepsilon}$$

where the second inequality follows since I_2 contains the n^{ξ} largest values a_i. The worst case for the left side of (10) is when $b'_j = n^{\beta}$ for $n^{\xi+\gamma-\varepsilon-\beta}$ columns and 0 elsewhere. Thus

$$(11) \qquad \sum_{j\in J} \min\{b'_j, c_{ij}\} = \Omega(n^{\xi+\gamma-\varepsilon-\beta})$$

By condition (iii) this is more than n^{α} so that (10) holds with high probability. This completes the outline of the proof. A detailed proof is given in Section 5.

3. Maximum Flows

We now consider the application of Theorem 1 to the maximum s - t flow problem in a directed or undirected, complete or random network G_n on a set of $n + 2$ vertices $V_n = (s,1,\dots,n,t)$ with random edge capacities. For the undirected complete problem, Grimmett and Welsh (1982) have considered the case where each arc capacity c_{ij}, $i \in V$, $j \in V$ is drawn independently from a probability distribution which does not vary with n. Their main result is that $\lim_{n\to\infty} X_n/(n + 1) \to \mu$ almost surely where X_n is the maximum s - t flow on G_n, and μ is the expected value of c_{ij}. Their proof uses the fact that the minimum s - t cut in G_n is one of the two stars rooted at s or t. There, and in Grimmett and Suen (1982), a similar result is proved for a directed complete version in which each arc is oriented with probability p from the lower to the higher indexed node, and with probability (1 - p) in the opposite way.

Theorem 1 enables us to analyze the problem even if the capacities are not i.i.d., and even if their joint distribution depends on n, as long as the matrix C is proper. Furthermore, the theorem can be used to produce an $O(n^2)$ algorithm which constructs the optimum s - t flow almost surely. As it turns out, the optimum flow produced by the algorithm uses paths of at most three edges. To demonstrate the generality of the conditions which could be handled, we mention here the following cases:

<u>Case 1.</u> Let G_n be an undirected complete graph on V_n. Define $c_{0j} = |c_{sj} - c_{jt}|$ and assume that $\{c_{ij}: 0 \leq i < j \leq n\}$ is proper, and that c_{jt}, c_{sj} are uniformly bounded from above for $j = 1,\dots,n$.

<u>Case 2.</u> Let G_n be a directed complete graph on V_n, i.e., both (i,j) and (j,i) exist. Assume that $\{C_{ij}: i \neq j\}$ is proper and that c_{jt}, c_{ij} are uniformly bounded from above for $j = 1,\dots,n$.

Case 3. Let G_n be the directed graph obtained from the complete undirected graph on V_n as follows. For each pair $\{i,j\}$ $i < j$ either (i,j) or (j,i) are in G_n but not both, and the probability that (i,j) is in G_n, denoted by $0 < p < 1$, is independent of n (for this case we may assume that $s \equiv 0$ and $t \equiv \infty$). Assume that the sets of capacities associated with the arcs of G_n form a proper collection and that c_{jt}, c_{st} are uniformly bounded from above for $j = 1,\ldots n$.

Case 4. In any of the above cases the underlying graph may be random, where the probability of selecting an arc is independent of n (in Case 3 this means randomly directing the arcs of a random graph).

In all cases let

$$F_s = \sum_{j=1}^{n} c_{si},$$

$$F_t = \sum_{j=1}^{n} c_{jt}.$$

Theorem 2.

(a) $F_{max} = \min\{F_s, F_t\}$, almost surely.

(b) A max flow in G can be almost surely found in $O(|E|)$ time, where E is the set of edges of G.

Proof. Without loss of generality, assume $F_s \leq F_t$. Arbitrarily reduce the capacities c_{it} $i = 1,\ldots,n$, until $F_s = F_t$. We show that a flow of value F_s almost surely exists in the new network. The cases of directed graphs follow easily from Theorem 1 as follows. Let $x_{si} = c_{si}$, $x_{it} = c_{it}$, and let x_{ij} $i,j \in \{1,\ldots,n\}$ be the solution to CTP with $a_i = c_{si}$, $b_j = c_{jt}$, and capacities

c_{ij}. Note that this problem almost surely satisfies the stipulation of Theorem 1 with $\alpha = \beta = 0$, $\gamma = \delta = \varepsilon = 1$. For the undirected case we cannot use this argument since with the condition $c_{ij} = c_{ji}$ the collection $\{c_{ij}: i \neq j\})$ is not necessarily proper. To overcome this difficulty let

$$I = \{i: c_{si} > c_{it}\},$$

$$J = \{i: c_{si} < c_{it}\}.$$

Let $x_{si} = c_{si}$, $x_{it} = c_{it}$ and x_{ij} be the solution to the CTP with $a_i = c_{si} - c_{it}$ $i \in T$, $b_j = c_{jt} - c_{sj}$ $j \in J$, and capacities c_{ij}, $i \in I$, $j \in J$. Note that the condition of properness with respect to c_{0j} in this case implies that $\gamma = 1$ almost surely so that Theorem 1 can be used again. $\square$

We note that Case 3 stipulates $0 < p < 1$. The case of $p = 0$ was considered by Grimmett and Welsh (1982) and Grimmett and Suen (1982). They proved that, for this case too, $F_{max} \sim (n + 1)p\mu$. This case is not covered by our theorem. Indeed, it is not true that $F_{max} = \min\{F_s, F_t\}$ almost surely. In fact, since no edge (n,j) exists for $j \neq t$ then the event $c_{sn} > c_{nt}$ implies $F_{max} < F_s$. Similarly, since no edge $(i,1)$ exists for $i \neq s$ then the event $c_{s1} < c_{1t}$ implies $F_{max} < F_t$. Thus, if both events occur then $F_{max} < \min\{F_s, F_t\}$.

We conclude this section with another example demonstrating the necessity of the condition that C be proper. Let the capacities of G_n be i.i.d. with $Pr[c_{ij} = 1] = 1/n$, $Pr[c_{ij} = 0] = 1 - 1/n$. Then, with probability $1 - (1 - 1/n)^n \to 1 - 1/e$ there exists an edge (s,i), $i \neq t$, such that $c_{si} = 1$. Therefore, with probability at least $(1 - 1/e)/e = (e - 1)/e^2$ the star rooted at s is not minimal. A similar argument applies to the other star. Thus, the probability that $F_{max} < \min\{F_s, F_t\}$ is at least $(e - 1)^2/e^4$.

4. Proof of Lemma 2. We now complete the necessary details for the proof of Lemma 2. In preparation for the proof, we need the following lemmas:

Lemma 3 (Renyi, 1970): Let $0 < p < 1$, $q = 1 - p$, $x < \frac{1}{2}\sqrt{\frac{n}{pq}}$. Then

$$\sum_{|r-np|>x\sqrt{npq}} \binom{n}{r} p^r q^{n-r} \leqslant 2\exp(-x^2/4). \qquad \square$$

Lemma 4: Let x_i, $i = 1,\ldots,n$ be independent Bernoulli random variables with $\Pr[x_i = 1] \equiv p$. Let r_i be nonnegative constants with $r_i \leqslant U$, $i = 1,\ldots,n$. Define $d = 8(1 - p)/p + 1$, $A = \sum_{i=1}^{n} r_i - dU$ and

$$S = \sum_{i=1}^{n} x_i r_i$$

Then

$$\Pr[S < Ap\frac{d-1}{8d}] \leqslant 2\exp(-\frac{A}{16dU})$$

Proof: The assertion is trivial if $A \leqslant 0$. For the other case, aggregate the weights r_i, $i = 1,\ldots,n$ into bins with capacity dU each. The number of bins used must be larger than one. It is possible to pack the bins such that all but at most one of them contains at least $(d - 1)U$ total weight. We refer to such bins as full. The number of full bins, ℓ, is obviously at least 1 and furthermore satisfies

$$\ell \geqslant \frac{\Sigma r_i}{dU} - 1 = \frac{A}{dU}$$

Consider the set of elements which fall into the j^{th} full bin, N_j. Then

$$\sum_{i \in N_j} r_i \geq (d - 1)U.$$

Choose arbitrary constants $0 \leq r_i' \leq r_i$, $i \in N_j$, so that

$$\sum_{i \in N_j} r_i' = (d - 1)U.$$

Consider the sum

$$S_j' = \sum_{i \in N_j} x_i r_i'$$

then

$$E(S_j') = p(d - 1)U$$

$$\mathrm{Var}(S_j') = p(1 - p) \sum_{i \in N_j} (r_i')^2 \leq p(1 - p)(d - 1)U^2$$

Thus, by Chebyshev's inequality

$$\Pr[S_i' \leq \frac{p(d - 1)U}{2}] \leq \frac{4(1 - p)}{(d - 1)p} = \frac{1}{2}$$

where the last equaltiy follows our choice of d.

Call a full bin j a <u>success</u> if its contribution to S, namely

$$S_j = \sum_{i \in N_j} r_i x_i$$

satisfies $S_j \geq \frac{p(d-1)U}{2}$. Since $r_i \geq r_i'$, $i \in N_j$, the probability of success is at least 1/2. Since the number of full bins is at least A/dU, the number of successes is at least A/4dU with probability at least $1 - 2\exp(-\frac{A}{16dU})$ (Lemma 3). But then S must satisfy

$$S \geq (\frac{A}{4dU})(\frac{p(d-1)U}{2}) = Ap\,\frac{d-1}{8d}$$

with the prescribed probability. $\square$

Lemma 5: Let r be a vector of nonnegative constants with $M = \sum_{k \in K} r_k$. Let $Q = q_{\ell k}$, $\ell = 1,\ldots,L$, $k \in K$ be a proper matrix with constant c. Define

$$S = \sum_{k \in K} \min\{r_k, \min\{q_{\ell k}: \ell = 1,\ldots,L\}\}$$

Assume that M is large enough, $Lt \geq 5/c$, and $t \geq \max\{r_k: k \in K\}$. Then

(12) $$\Pr[S < M/16cLt] \leq 2\exp(-M/32t)$$

Proof: Let $K_1 = \{k \in K: r_k \geq 1/2cL\}$, $K_2 = K\backslash K_2$. Also, call index $k \in K$ a success if

$$\min_{\ell=1,\ldots L} \{q_{\ell k}\} \geq 1/2cL$$

Clearly the probability of success is at least $(1 - 1/2L)^L \geq 1/2$. Let

$$M_1 = \sum_{k \in K_1} r_k, \quad M_2 = M - M_1.$$

Case 1: $M_1 \geq M/2$. Let

$$S_1 = \sum_{k \in K_1} \min\{r_k, \min\{q_{\ell k}: \ell = 1,\ldots,L\}\}$$

Let N_1 be the number of successes in K_1. Note that $|K_1| \geq M/2t$ so that we get from Lemma 3 (with $n = M/2t$, $p = 1/2$)

$$\Pr[N_1 < M/8t] \leq 2 \exp\{-M/32t\}$$

since $S \geq S_1$ and since each success contributes at least $1/2cL$ to S_1, we obtain (12).

Case 2: $M_2 \geq M/2$. Let

$$S_2 = \sum_{k \in K_2} \min\{r_k, \min\{q_{\ell k}: \ell = 1,\ldots,L\}\}.$$

Then

$$S \geq S_2 \geq T \equiv \sum_{\substack{k \in K_2 \\ k \text{ is a success.}}} r_k$$

Note that T stochastically dominates $\sum_{k \in K_2} x_k r_k$ where the x_k are independent Bernoulli random variables with $\Pr[x_k = 1] = 1/2$. Apply Lemma 4 with $p = 1/2$ (which yields $d = 9$), $\Sigma r_i = M/2$, $U = 1/2cL$ to obtain

(13) $$\Pr[S < \frac{A}{18}] \leq 2 \exp(-\frac{AcL}{72})$$

where A is as in Lemma 4. The condition $M \gg 1$, and $Lt > 5/c$, imply that

$$\frac{M}{16cLt} < \frac{A}{18} = \frac{M}{36} - \frac{1}{4cL}$$

and

$$\frac{M}{32t} < \frac{AcL}{72} = \frac{McL - 9}{144}$$

so that (13) implies (12) □

Proof of Lemma 2: We have to show that the inequalities (7), (8) and (11) of the outline hold simultaneously almost surely for all $i \in I$, $j \in J$. To prove (7) use Lemma 5 with $r = b$, $Q = C_2$, $L = n^{\xi}$, $t = n^{\beta}$, $M = \sum_{j \in J} b_j = n^{\gamma}$ to obtain

$$\Pr[\sum_{j \in J} \bar{b}_j \geq \frac{n^{\gamma-\beta-\xi}}{16c}] \geq 1 - 2\exp(-\frac{n^{\gamma-\beta}}{32})$$

Since $\gamma > \beta$, (7) occurs almost surely. To prove (8) for column j use Lemma 5 with $r = (a_i^{'}: i \in I_1)$, $L = 1$, $Q = (c_{ij}: i \in I_1)$, $M \geq \theta n^{\gamma-\beta-\xi}$ for some constant θ ($\theta = \frac{1}{16} c$ will almost surely do as per (7)), and $t = n^{\alpha}$ to obtain

$$\Pr[\sum_{i \in I_1} \min\{a_i^{'}, c_{ij}\} \geq \theta^{'} n^{\gamma-\alpha-\beta-\xi}] \geq 1 - 2\exp(-\theta^{''} n^{\gamma-\alpha-\beta-\xi})$$

where $\theta^{'} = \theta/16c$ and $\theta^{''} = \theta/32$. Since the number of columns satisfies $|J| \leq n^{\delta}$, (8) holds simultaneously for all $j \in J$ almost surely. Finally, to prove (11) for row $i \in I_2$ apply Lemma 5 with $t = n^{\beta}$, $L = 1$, $Q = (c_{ij}: j \in J)$, $r = b^{'}$ and $M = \sum_{j \in J} b_j \geq \sum_{i \in I_2} a_i \geq n^{\xi+\gamma-\varepsilon}$ to obtain

$$\Pr[\sum_{j \in J} \min\{b_j^{'}, c_{ij}\} \geq \theta n^{\xi+\gamma-\varepsilon-\beta}] \geq 1 - 2\exp(-\frac{n^{\xi+\gamma-\varepsilon-\beta}}{32})$$

By condition (iii), $\xi + \gamma - \varepsilon - \beta > 0$. Thus, since $|I_2| < |I| \leqslant n^{\varepsilon}$, (11) holds simultaneously for all $i \in I_1$ almost surely. This concludes the proof of the lemma. $\square$

References

Bollabás, B. (1983), "Almost All Regular Graphs are Hamiltonian," European J. Combin. 4, 97-106.

Doulliez, P. and Jamoulle, E. (1972), "Transportation Networks with Random Arc Capacities," Rev. Francaise d'Automatique Informatique et Recherche Operationelle 3, 45-60.

Fenner, T. I. and A. M. Frieze (1983), "On the Existence of Hamiltonian Cycles in a Class of Random Graphs," Discrete Math. 45, 301-305.

Frank, H. and S. L. Hakimi (1965), "Probabilistic Flows Through a Communication Network," I.E.E.E. Trans. Circuit. Theory, CT-12, 413-414.

Frank, H. and S. L. Hakimi (1967), "On the Optimum Synthesis of Statistical Communication Networks--Psuedo Parametric Techniques," J. Franklin Inst., 284, 407-416.

Frank, H. and S. L. Hakimi (1968), "Parametric Analysis of Statistical Communication Networks," Quant. Appl. Math., 26, 249-263.

Frank, H. and I. T. Frisch (1971), Communication, Transmission, and Transportation Networks, Ch. 4 ("Maximal Flow in Probabilistic Graphs"), Addison-Wesley Publishing Co., Inc.

Frieze, A. M. (1984a), "Maximum Matching in a Class of Random Graphs," Graduate School of Industrial Administration, Carnegie-Mellon University, Pittsburgh, Pennsylvania.

Frieze, A. (1984b), "Limit Distribution for the Existence of Hamiltonian Cycles in a Random Graph," MSRR 508, GSIA, Carnegie-Mellon University.

Frieze, A. M. (1985), "On the Value of a Random Minimizing Spanning Tree Problem," Discrete Appl. Math., 10, 47-56.

Frieze, H. and G. R. Grimmett (1985), "The Shortest Path Problem for Graphs with Random Arc Lengths," Discrete Applied Math. 10, 57-78.

Grimmett, G. R. and D. R. A. Welsh (1982), "Flow in Networks with Random Capacities," Stochatics 7, 205-229.

Grimmett, G. R. and H.-C. S. Suen (1982), "The Maximal Flow Through a Directed Graph with Random Capacities," Stochastics 8, 153-159.

Hassin, R. and E. Zemel (1984), "On Shortest Paths in Graphs with Random Weights," Math. of Operations Research (to appear).

Karmarkar, N. (1982), "Probabilistic Analysis of Some Bin-Packing Algorithms," Proc. 23rd Annual IEEE Symp. Foundation of Computer Science, 107-111.

Karp, R. M. (1979), "The Probabilistic Analysis of Combinatorial Optimizations Algorithms," paper presented at the Tenth International Symposium on Mathematical Programming, Montreal, August 1979.

Karp, R. M., J. K. Lenstra, C. J. H. McDiarmid, and A. H. G. Rinnooy Kan (1984), "Probabilistic Analysis of Combinatorial Algorithms: An Annotated Bibliography," Centrum voor Wiskunde en Enformatica OS-R8411.

Komlós, J. and E. Szemerédi (1983), "Limit Distribution for the Exitence of Hamiltonian Cycles in Random Graphs," Discrete Math. 43, 55-63.

Loulou, R. (1982), "Probabilistic Behavior of Optimal Bin Packing Solutions," Faculty of Mangaement, McGill University, Montreal.

Piccard, J. C. and Queyranne, M. (1982), "Selected Applications of Minimum Cuts in Networks," INFOR 20, 394-422.

Pósa, L. (1976), "Hamiltonian Circuits in Random Graphs," Discrete Math. 14, 359-364.

Rényi, A. (1970), Foundation of Probability, Holden Day, Inc., San Francisco.

Shamir, E. and E. Upfal (1981), "On Factors in Random Graphs," Israel J. Math. 39, 296-302.

Vercellis, C. (1984), "A Probabilistic Analysis of the Set Covering Problem," Ann. Oper. Res. 1.