NETWORK FLOW AND SYSTEMS OF REPRESENTATIVES

L. R. FORD, Jr., AND D. R. FULKERSON

Introduction. The theory developed for the study of flows in networks (**2**; **3**; **4**; **5**; **6**; **7**) sometimes provides a useful tool for dealing with certain kinds of combinatorial problems, as has been previously indicated in (**3**; **4**; **6**; **7**). In particular, Hall-type theorems for the existence of systems of distinct representatives which contain a prescribed set of marginal elements (**10**; **11**), or, more generally, whose intersection with each member of a given partition of the fundamental set has a cardinality between prescribed lower and upper bounds (**9**), can be obtained in this way (**7**). In this note we apply network flow theory to generate necessary and sufficient conditions for (a) the existence of a system of restricted representatives, by which we mean a system of representatives such that each element a_i of the fundamental set occurs at least α_i times in the system, and at most β_i times, and (b) the existence of a common system of restricted representatives for two different collections of subsets of the fundamental set. While problem (b) clearly includes (a), we have chosen to treat the two separately.

Section 1 describes relevant portions of flow theory. In §2 we show how Hall's condition for the existence of a system of distinct representatives and a similar condition for problem (a) may be deduced from maximal network flow problems. Section 3 deals with problem (b) and resolves (a) as a special case.

We emphasize that the present approach may be used not only to yield existence conditions for certain kinds of systems of representatives, but may also be used to provide explicit algorithms for constructing such as well. On the other side of the ledger, it can be shown, although we do not demonstrate it in this paper, that each of the problems we have mentioned can be reduced, by suitably manipulating the network which represents the problem, to an application of Hall's theorem.

1. Network Flow. A basic problem concerning network flows is the following. Suppose given a finite network (linear graph) N with node set $\{s, \ldots, x, y, \ldots, s'\}$ and oriented arcs joining pairs of nodes, the arc from x to y being denoted by (x, y), and suppose each (x, y) has associated with it a *capacity* $c(x, y)$, where $c(x, y)$ is either a non-negative real number or plus infinity. Subject to the conditions (i) the flow in (x, y) is no greater than $c(x, y)$, (ii) the total flow into node x ($x \neq s, s'$) is equal to the flow out, find a maximal flow from s (the *source*) to s' (the sink).

Received March 18, 1957.

Thus, letting $f(x, y)$ be the flow in (x, y), the problem may be described as a linear programme:

$$(1)\quad \begin{cases} \text{(a)} & \sum_y [f(s, y) - f(y, s)] - v = 0 \\ \text{(b)} & \sum_y [f(x, y) - f(y, x)] = 0 \qquad (x \neq s, s') \\ \text{(c)} & \sum_y [f(s', y) - f(y, s')] + v = 0 \\ \text{(d)} & 0 \leqslant f(x, y) \leqslant c(x, y) \\ \text{(e)} & \text{maximize } v. \end{cases}$$

If $(f; v)$ is a solution of the constraints (1a) – (1d), f is a *flow* and v its *value.*

There are algorithms available for solving such problems. The best known of these is probably G. Dantzig's simplex method **(1)** for solving the general linear programming problem of maximizing a linear function subject to linear equations and inequalities. However, problem (1) is a special kind of linear programme for which simple (and computationally more efficient) algorithms have been constructed **(2; 4)**. These algorithms may be used to prove an intuitively plausible theorem which is basic in the study of network flow. To state this theorem, we require some definitions. A *cut* in N with respect to s, s' is a partition of the nodes into two complementary sets L, L' with $s \in L$, $s' \in L'$. The *value* of a cut is

$$\sum_{x \in L, y \in L'} c(x, y).$$

MINIMAL CUT THEOREM **(3; 4; 5)**. *For any network, the maximal flow value is equal to the minimal cut value.*

We remark that it is obvious that flow values are bounded above by cut values. Thus the content of the theorem is the assertion that there is a flow and a cut for which equality of values holds.

In addition to this theorem, we need one other result for the combinatorial applications to be presented in the sequel.

INTEGRITY THEOREM **(3; 4)**. *If the capacity function is integral valued, there exists a maximal flow which is also integral valued.*

The integrity theorem can also be deduced in a variety of ways. For example, the algorithms for constructing maximal flows which were referred to previously can be shown to produce integral flows in case the arc capacities $c(x, y)$ are integers. The theorem also follows from the fact that all the extreme points of the convex polyhedron defined by (1a)-(1d) are integral.

2. Hall's Theorem; Systems of Restricted Representatives. Let $\mathscr{S} = \{S_1, \ldots, S_n\}$ be a family of subsets of a given set $A = \{a_1, \ldots, a_m\}$. A list R of (not necessarily distinct) elements

https://doi.org/10.4153/CJM-1958-009-1 Published online by Cambridge University Press

$$a_{i_1}, \ldots, a_{i_n}$$

is a *system of representatives* for $\mathscr{S}$ if

$$a_{i_j} \in S_j, \qquad j = 1, \ldots, n.$$

If we further stipulate that each element $a_i \in A$ occurs in R at least α_i times and at most β_i times, where $0 \leqslant \alpha_i \leqslant \beta_i$, we call R a *system of restricted representatives* (abbreviated SRR). In case $\alpha_i = 0$, $\beta_i = 1$ for all i, then R is a *system of distinct representatives* (abbreviated SDR). A well-known theorem of P. Hall **(8)** states that a necessary and sufficient condition for the existence of an SDR is that, for each $k = 1, \ldots, n$, every union of k sets of $\mathscr{S}$ contains at least k elements. The necessity of the condition is of course obvious.

As an exercise, let us construct a network maximal flow problem which represents the problem of finding an SDR and deduce Hall's condition from it. To this end, let

$$s, \bar{S}_1, \ldots, \bar{S}_n, \bar{a}_1, \ldots, \bar{a}_m, s'$$

be the nodes of N, and define arcs and capacities as follows:

$$\begin{array}{lll} (s, \bar{S}_j) & \text{with capacity } 1, & j = 1, \ldots, n, \\ (\bar{S}_j, \bar{a}_i) & \text{with capacity } \infty, & i, j \ni a_i \in S_j, \\ (\bar{a}_i, s') & \text{with capacity } 1, & i = 1, \ldots, m. \end{array}$$

We assert that an SDR exists for $\mathscr{S}$ if and only if the maximal flow value in N is n. For, given an SDR, we can construct an integral flow of value n as follows. Let

$$f(s, \bar{S}_j) = 1$$

$$f(\bar{S}_j, \bar{a}_i) = \begin{cases} 1 \text{ if } a_i \text{ occurs in the SDR,} \\ 0 \text{ otherwise} \end{cases}$$

$$f(\bar{a}_i, s') = \begin{cases} 1 \text{ if } a_i \text{ occurs in the SDR,} \\ 0 \text{ otherwise.} \end{cases}$$

This is clearly a flow in N of value n; it is certainly maximal since the cut value

$$\sum_{j=1}^{n} c(s, \bar{S}_j)$$

is also n. Conversely, if the maximal flow value is n, we may select (by the integrity theorem) an integral flow of value n, and let a_i represent S_j if and only if $f(\bar{S}_j, \bar{a}_i) = 1$. Then all sets S_j are represented (since the flow has value n and $c(s, \bar{S}_j) = 1$) and no a_i occurs more than once in the representation (since $c(\bar{a}_i, s') = 1$). Thus an SDR exists for $\mathscr{S}$ if and only if the maximal flow value (that is, minimal cut value) for the associated network is n.

To discover Hall's condition, we simply examine all candidates for minimal cuts, and insist that their values exceed n. First let us introduce some notation. Given two disjoint subsets X, Y of the nodes of a network N, let (X, Y) denote the set of arcs from any node of X to any node of Y, and let

$$c(X, Y) = \sum_{(x,y \in X,Y)} c(x, y).$$

Also, for any set X of nodes, let $J(X)$ be the set of nodes of N which are joined to some node of X by an arc. Finally, let $|X|$ denote the cardinality of set X.

Suppose now that (L, L') is a cut in a representing network N for the SDR problem. Let $\bar{S} = \{\bar{S}_1, \ldots, \bar{S}_n\}$, $\bar{A} = \{\bar{a}_1, \ldots, \bar{a}_m\}$, and define subsets of the nodes of N as follows:

$$X = L \cap \bar{S};\; X' = L' \cap \bar{S};\; Y = L \cap \bar{A};\; Y' = L' \cap \bar{A}.$$

Then the condition which is equivalent to the existence of a flow of value n is

$$c(L, L') = c(s, X') + c(X, Y') + c(Y, s') \geqslant n \tag{2}$$

for all cuts (L, L'), or equivalently, for all $X \subset \bar{S}$, $Y \subset \bar{A}$.

Now (2) holds automatically unless (X, Y') is vacuous, and (X, Y') empty implies $J(X) \cap \bar{A} \subset Y$. Thus the set of inequalities (2) is equivalent to the set

$$|X'| + |Y| \geqslant n \tag{3}$$

for all $X \subset \bar{S}$, all $Y \supset J(X) \cap \bar{A}$, and hence to

$$|X'| + |J(X) \cap \bar{A}| \geqslant n \tag{4}$$

for all $X \subset \bar{S}$. Replacing $|X'|$ by $n - |X|$ in (4) yields

$$|X| \leqslant |J(X) \cap \bar{A}|, \qquad \text{all } X \subset \bar{S}. \tag{5}$$

All that remains is to restate (5) in the language of sets: for any subset X of the indices $\{1, \ldots, n\}$,

$$|X| \leqslant |I(X)|, \tag{6}$$

where $I(X) \subset \{1, \ldots, m\}$ is the index set of $\bigcup S_j (j \in X)$.

With this as background, let us next turn to the question of the existence of an SRR. For this problem, let

$$s, \bar{S}_1, \ldots, \bar{S}_n, \bar{a}_1, \ldots, \bar{a}_m, t, s'$$

be the nodes of N; the arcs and capacities are

$(s, \bar{S}_j)$	with capacity	$1,$	$j = 1, \ldots, n,$
$(\bar{S}_j, \bar{a}_j)$	with capacity	$\infty,$	$i, j \ni a_i \in S_j,$
$(\bar{a}_i, t)$	with capacity	$\beta_i - \alpha_i,$	$i = 1, \ldots, m,$
$(\bar{a}_i, s')$	with capacity	$\alpha_i,$	$i = 1, \ldots, m,$
(t, s')	with capacity	$n - \sum_{i=1}^{m} \alpha_i.$	

(Notice that we are tacitly assuming $n \geqslant \sum \alpha_i$, obviously a necessary condition for the existence of an SRR.)

It is not difficult to see that an SRR exists if and only if the maximal flow through N has value n. Define X, Y, X', Y' as before, and suppose first that $t \in L$. Then the relevant condition is

$$c(s, X') + c(X, Y') + c(Y, s') + c(t, s') \geqslant n \tag{7}$$

for all $X \subset \bar{S}$, $Y \subset \bar{A}$. Proceeding as before, (7) leads to

$$|X'| + \sum_{Y} \alpha_i + n - \sum_{i=1}^{m} \alpha_i \geqslant n \tag{8}$$

for all $X \subset \bar{S}$, all $Y \supset J(X) \cap \bar{A}$, and thus to

$$|X| \leqslant n - \sum_{i=1}^{m} \alpha_i + \sum_{J(X) \cap \bar{A}} \alpha_i \tag{9}$$

for all $X \subset \bar{S}$.

In a similar manner, if $t \in L'$, one obtains the condition

$$|X| \leqslant \sum_{J(X) \cap \bar{A}} \beta_i \tag{10}$$

for all $X \subset \bar{S}$. Thus we may state (since (9) includes the condition

$$n - \sum_{i=1}^{m} \alpha_i \geqslant 0):$$

THEOREM 1. *An* SRR *exists for* $\mathscr{S} = \{S_1, \ldots, S_n\}$ *if and only if, for every subset* X *of the indices* $\{1, \ldots, n\}$

$$|X| \leqslant \min \left(n - \sum_{i=1}^{m} \alpha_i + \sum_{I(X)} \alpha_i, \sum_{I(X)} \beta_i\right) \tag{11}$$

where $I(X) \subset \{1, \ldots, m\}$ *is the index set of* $\bigcup S_j (j \in X)$.

Observe that (11) reduces to Hall's condition in case $\alpha_i = 0$, $\beta_i = 1$ for all i. Also, if $\alpha_i = 1$, for $i = 1, \ldots, q$, and $\alpha_i = 0$ for $i = q + 1, \ldots, m$, all $\beta_i = 1$, (11) yields the Hoffman-Kuhn condition **(10)** for the existence of an SDR containing a prescribed set of marginal elements $a_1, \ldots, a_q$.

3. Existence of Common SRR. Since the only ingenuity required in solving problems of the kind we have discussed lies in finding a representing network (if one exists), we shall merely give a description of such a network for the common SRR problem and a statement of the conditions, leaving the proof to the reader.

Let $\mathscr{S} = \{S_1, \ldots, S_n\}$, $\mathscr{T} = \{T_1, \ldots, T_n\}$ be the two families of subsets of $A = \{a_1, \ldots, a_m\}$. Define a network N consisting of nodes

$$s, \bar{S}_1, \ldots, \bar{S}_n, \bar{a}_1, \ldots, \bar{a}_m, \tilde{a}_1, \ldots, \tilde{a}_m, \bar{T}_1, \ldots, \bar{T}_n, s'$$

and arcs

$(s, \bar{S}_j)$	with capacity 1,	$j = 1, \ldots, n,$
$(s, \tilde{a}_i)$	with capacity α_i,	$i = 1, \ldots, m,$
$(\bar{S}_j, \bar{a}_i)$	with capacity ∞,	$i, j \ni a_i \in S_j,$
$(\bar{a}_i, \tilde{a}_i)$	with capacity $\beta_i - \alpha_i$,	$i = 1, \ldots, m,$
$(\bar{a}_i, s')$	with capacity α_i,	$i = 1, \ldots, m,$
$(\tilde{a}_i, \bar{T}_j)$	with capacity ∞,	$i, j \ni a_i \in T_j,$
$(\bar{T}_j, s')$	with capacity 1,	$j = 1, \ldots, n.$

As a common SRR exists for $\mathscr{S}, \mathscr{T}$ if and only if there is a flow from s to s' of value

$$n + \sum_{i=1}^{m} \alpha_i,$$

similar procedures lead to the following theorem.

THEOREM 2. *A common* SRR *exists for* $\mathscr{S} = \{S_1, \ldots, S_n\}, \mathscr{T} = \{T_1, \ldots, T_n\}$ *if and only if, for every* $X, Y \subset \{1, \ldots, n\}$,

$$|X| + |Y| \leqslant n - \sum_{i=1}^{m} \alpha_i + \sum_{I(X) \cup I(Y)} \alpha_i + \sum_{I(X) \cap I(Y)} \beta_i \tag{12}$$

where $I(X) \subset \{1, \ldots, m\}$ *is the index set of* $\cup S_j$ $(j \in X)$, *and* $I(Y) \subset \{1, \ldots, m\}$ *is the index set of* $\cup T_j (j \in Y)$.

Notice that, for any given X, taking Y empty yields

$$|X| \leqslant n - \sum_{i=1}^{m} \alpha_i + \sum_{I(X)} \alpha_i$$

and taking Y the full set yields

$$|X| \leqslant - \sum_{i=1}^{m} \alpha_i + \sum_{I(X) \cup I(Y)} \alpha_i + \sum_{I(X) \cap I(Y)} \beta_i \leqslant \sum_{I(X)} \beta_i$$

which combine to give (11). Conversely, if $S_i = T_i$ for all i, and if (11) holds for all X, then (12) holds for all X, Y. To see this,[1] suppose given any two sets $X, Y \subset \{1, \ldots, n\}$ and apply (11) to the sets $X \cup Y$, $X \cap Y$, obtaining in particular

$$|X \cup Y| \leqslant n - \sum_{i=1}^{m} \alpha_i + \sum_{I(X \cup Y)} \alpha_i = n - \sum_{i=1}^{m} \alpha_i + \sum_{I(X) \cup I(Y)} \alpha_i$$

$$|X \cap Y| \leqslant \sum_{I(X \cap Y)} \beta_i \leqslant \sum_{I(X) \cap I(Y)} \beta_i.$$

Adding these two inequalities gives (12).

By taking $\alpha_i = 0$, $\beta_i = 1$ in (12), one obtains conditions for the existence of a common SDR.

COROLLARY. *A common* SDR *exists for* $\mathscr{S}$ *and* $\mathscr{T}$ *if and only if*

$$|X| + |Y| \leqslant n + |I(X) \cap I(Y)|, \tag{13}$$

where $I(X)$, $I(Y)$ *are as defined in Theorem* 2.

---

[1]This short proof is due to O. Gross.

References

1. G. B. Dantzig, A. Orden, and P. Wolfe, *The generalized simplex method for minimizing a linear form under linear inequality constraints*, Pacific J. Math., *5* (1955), 183–195.
2. G. B. Dantzig and D. R. Fulkerson, *Computation of maximal flows in networks*, Naval Research Logistics Quarterly, *2* (1955), 277–283.
3. ———, *On the min cut max flow theorem of networks*, Annals of Mathematics Study No. 38, Linear Inequalities and Related Systems, ed. H. W. Kuhn and A. W. Tucker (Princeton, 1956), 215–221.
4. L. R. Ford, Jr., and D. R. Fulkerson, *A simple algorithm for finding maximal network flows and an application to the Hitchcock Problem*, Can. J. Math., *9* (1957), 210–218.
5. ———, *Maximal flow through a network*, Can. J. Math., *8* (1956), 399–404.
6. D. Gale, *A Theorem on flows in networks*, RAND Corporation, Research Memorandum RM-1737, 1956 (to appear in Pacific J. Math.).
7. D. Gale and A. Hoffman, *Circulation in networks* (unpublished notes).
8. P. Hall, *On representatives of subsets*, J. Lond. Math. Soc., *10* (1935), 26–30.
9. A. J. Hoffman and H. W. Kuhn, *On systems of distinct representatives*, Annals of Mathematics Study, No. 38, Linear Inequalities and Related Systems, ed. H. W. Kuhn and A. W. Tucker (Princeton, 1956), 199–206.
10. ———, *Systems of distinct representatives and linear programming*, Amer. Math. Monthly, *63* (1956), 455–460.
11. H. B. Mann and H. J. Ryser, *Systems of distinct representatives*, Amer. Math. Monthly, *60* (1953), 397–401.

Santa Monica, California