

Exact and Approximate Algorithms for Topological Design of Wide Area Networks with Non-simultaneous Single Commodity Flows

Andrzej Kasprzak

Wroclaw University of Technology, Chair of Systems and Computer Networks
50-370 Wroclaw, Poland
kasprzak@zssk.pwr.wroc.pl

Abstract. The problem is concerned with designing of wide area computer networks. This problem is very important from practical point of view because the network, when designed efficiently, can and does become the lifeblood of a business organization. The problem consists in selection of the non-simultaneous single commodity flow routes, channel capacities and the wide area network topology in order to maximize the value of the min-max flow subject to the budget constraint. The problem is NP-complete. Then, the branch and bound method is used to construct the exact algorithm. Several properties of the problem are formulated and proved. Computational results are reported.

1 Introduction

Data transmission in a wide area network (WAN) environment involves the design of a communication network. Several different formulations of the design problem can be found in the literature; generally, they correspond to different choices of criterion, of design variables and of constraints [1], [2], [3]. Usually delay, cost, throughput and network reliability are applied as criteria. For the purpose of this paper the throughput of the WAN represented by non-simultaneous single commodity flow requirements is selected as criterion. The design variables are following: topology, channel capacities and routing (i.e. flows). The techniques used for the solution of topological design problem depend on the nature of cost-capacity functions. The problem is considered for linear, concave and discrete cost-capacity functions [3]. Since the channel capacities can be chosen from discrete sequence defined by ITU-T (International Telecommunications Union – Telecommunications Sector) recommendations, then only discrete cost-capacity function is important from practical point of view.

Consider the network and maximal flows between each pair of its nodes. Of all the maximal flows in the network the one the value of which is minimal is called the min-max flow. The non-simultaneous single commodity flow requirements are understood here as the value of the min-max flow in the network.

In the paper, we consider the new problem which is concerned with selection of the capacities of channels and their geographical layout (i.e. network topology) so as to satisfy the investment budget constraint and maximize the min-max flow value. The

budget constraint limits the number of channels that may be included in the optimal network topology as well as the values of capacities of these channels chosen from the set of discrete values of capacities defined by ITU-T recommendations. The discrete cost-capacity function is considered. Also we assume that there are only half-duplex channels in the WAN. The considered problem can be formulated as follows.

Given: node location, possible channels locations, channel capacity options and costs, budget of the WAN
Maximize: value of min-max flow
over: topology, channel capacities, routing (i.e. min-max flow)
subject to: flow constraints, channel capacity constraints, budget constraint.

First, several cases of the network synthesis problem with non-simultaneous single commodity flow requirements were studied by Gomory and Hu [2], [4]. They study minimum cost network design problems with the linear cost-capacity function and without budget constraint. The formulation of the considered problem differs from the formulation of Gomory and Hu problem because: we assume the discrete costs-capacity function (Gomory and Hu assume only linear cost-capacity function), we take into account the budget constraint (Gomory and Hu problem is solved without the budget constraint) and we use other optimization criterion (value of min-max flow) than criteria used in [2], [4]. The main difference between the problem [2], [4] and the problem considered here is such that the problem [2], [4] belongs to the class of continuous problems and our problem belongs to the class of discrete problems. The formulation of the problem considered here is similar to formulations of topological design problems of WAN presented for example in the papers [1], [3].

2 Problem Formulation

Let p be the number of channels which must be installed in the WAN and let m be the number of possible channels which may be used to build the WAN. The capacity of the i -th channel which must be installed is denoted by c_0^i . For each potential i -th channel there is the set $ZC^i = \{c_1^i, c_2^i, \dots, c_{r(i)-1}^i\}$ of alternative capacities from among which exactly one must be chosen if the i -th channel was chosen to build the network. Let d_k^i be the cost of leasing capacity c_k^i . Let the network $S = \langle G; c, d \rangle$ be the model of the WAN. The network S is defined by a undirected multigraph $G = \langle N, L \rangle$, by a capacity function c and by a discrete cost function d ; L denotes the set of arcs.

$N = \{1, \dots, n\}$ - the set of nodes; N corresponds to the set of nodes of the WAN,

$A = \{w^{m+1}, \dots, w^p\}$ - the set of fixed undirected arcs; each arc from this set corresponds to the channel which must be installed in the WAN; $p = p' + m$.

$\bar{W}^i = \{w_1^i, \dots, w_{r(i)-1}^i\}$ - the set of alternative undirected arcs; at most one arc from this set may be used to construct the network S . This set corresponds to the i -th channel which may be used to build the WAN. The number of all \bar{W}^i is equal to m .

The functions c and d are defined as follows:

$$c(w_j^i) = c_j^i \text{ and } d(w_j^i) = d_j^i \text{ for every } w_j^i \in \bar{W}^i, i=1, \dots, m,$$

$$c(w^i) = c_0^i \text{ and } d(w^i) = 0 \text{ for every } w^i \in A,$$

where $c(w_j^i)$ and $d(w_j^i)$ are the capacity and the cost of the arc w_j^i , respectively.

All arcs from the set \bar{W}^i connect the same pair of nodes. Each arc w_j^i corresponds to capacity c_j^i from the set ZC^i . If the arc w_j^i is chosen to construct the network S , then it means that the i -th channel with capacity c_j^i is used to build the WAN.

Let's call the arcs $w_{r(i)}^i$ for $i=1, \dots, m$ the empty arcs; then $c(w_{r(i)}^i) = 0$ and $d(w_{r(i)}^i) = 0$. Then the set $W^i = \bar{W}^i \cup \{w_{r(i)}^i\}$ is the set of alternative arcs from among which exactly one arc must be used to construct the network S . If the arc $w_{r(i)}^i$ is chosen then the i -th channel is not used to build the WAN. Let the set W be the sum of all sets W^i . Thus, $L = W \cup A$. Let $X_r \subset W$ be the set of arcs such that X_r includes exactly one arc from each set W^i . X_r is called a selection. Let R be the set of all such selections. Each selection X_r generates the undirected graph $G_r = \langle N, A \cup X_r \rangle$ and the associated network $S_r = \langle G_r; c, d \rangle$. Therefore, the network S_r defines unique WAN because the selection X_r determines simultaneously potential channels which are used to build the WAN and the values of capacities for these channels.

Let $v_r(i, j)$ be the maximal flow value from source i to sink j in the network S_r . $v_r(i, j)$ is obtained solving the capacited maximal flow problem [2]. To find the maximal flows between all pairs of nodes of the network the Gomory and Hu algorithm [2] may be applied. It is easy to see that the flow between each pair of nodes in the network S_r corresponds to the packets flow in the WAN which topology and channel capacities are given by X_r .

Let $V(X_r)$ be the value of min-max flow in the network S_r , then

$$V(X_r) = \min_{i, j \in N} v_r(i, j). \tag{1}$$

Let DB be the budget of the WAN and let $d(X_r)$ be the building cost of S_r . Then, the considered topological design problem can be formulated as follows.

$$\max_{X_r} V(X_r) \tag{2}$$

subject to

$$X_r \in R \tag{3}$$

$$d(X_r) = \sum_{w \in X_r} d(w) \leq DB. \tag{4}$$

3 Calculation Scheme

Assuming that $r(i)=2$ and $c(w_1^i)=1$ for $i=1, \dots, m$, $A=\emptyset$, the problem (2-4) is resolved itself into the "two-connectivity" problem. Since the two-connectivity problem is NP-complete [5] then the problem (2-4) is also NP-complete as more general. Then, the branch and bound method can be used to construct the exact algorithm for solving the considered problem. The detailed description of the calculation scheme of the branch and bound method may be found in the literature.

Starting with the initial selection $X_1 \in R$ we generate a sequence of selections. The arc $w \in W$ is called normal if $w \in X_1$. The arc which does not belong to X_1 is called reverse. A replacement of any arc by other arc from W^i is called complementing.

The generation of a new selection involves the choice of a certain normal arc from X_r for complementing; it is called the branching rules. For each selection X_r we perform a testing operation (i.e. upper bound) to check the possibility of the generation of X_s with the greater value of min-max flow than already found.

If the new selection X_s is generated from X_r by complementing the normal arc by the reverse arc, then we constantly fix this reverse arc. It means that this reverse arc cannot be complemented by any other in every possible successor of X_s . So, for each selection X_r we constantly fix a set $F_r \subset X_r$. The arcs which do not belong to F_r are called free in X_r . If we want backtrack from X_1 then the algorithm terminates.

4 Testing Operation

The basic task of this operation is the computation of upper bound UB_r of value of min-max flow for every possible successor generated from X_r . Let \bar{C}_r be the upper bound of the sum of capacities of all arcs in S_r and in all its successors – it may be calculated using the algorithm similar to one presented in [3].

For each network S_s we have $v_s(x, y) \leq c_s(x, N)$ for every $x, y \in N, x \neq y$, where $c_s(x, N)$ is the sum of capacities of all arcs connected to node x in S_s . Summing over all nodes $x, y \in N, x \neq y$, we obtain

$$v_s(N, N) = \sum_{x=1}^n \sum_{\substack{y=1 \\ y \neq x}}^n v_s(x, y) \leq (n-1) \sum_{x=1}^n c_s(x, N) = (n-1)c_s(N, N) \tag{5}$$

where $c_s(N, N)$ is double sum of capacities of all arcs in S_s . Since $V(X_s) \leq v_s(x, y)$ for every $x, y \in N, x \neq y$, then after summing over all nodes x, y we obtain

$$n(n-1)V(X_s) \leq v_s(N, N). \tag{6}$$

It follows from (5) and (6) that $V(X_s) \leq c_s(N, N)/n$. Since \bar{C}_r is the upper bound of sum of all arc capacities in S_r and its successor S_s then $c_s(N, N) \leq 2\bar{C}_r$. Then the value of min-max flow $V(X_s)$ in every successor X_s of the X_r is bounded and

$$V(X_s) \leq 2\bar{C}_r/n. \tag{7}$$

Then the right hand side of the inequality (7) is the upper bound of min-max flow value in every successor of X_r . The following corollary follows from (6) and (7).

Corollary 1. Let X_s be the successor of $X_r \in R$. The value $V(X_s)$ is maximal iff the maximal flows between every pair of nodes in S_s are equal to $2\bar{C}_r/n$.

Let $v_r^u(x, y)$ be the maximal flow in the network $S_r^u = \langle G_r^u; c, d \rangle$, where $G_r^u = \langle N, A \cup F_r \rangle$. The network S_r^u contains only arcs from the sets A and F_r which are not changed in every successor of S_r . Then, for the maximal flows $v_s(x, y)$ in every successor S_s of the network S_r the following condition is satisfied

$$v_s(x, y) \geq v_r^u(x, y) \text{ for every } x, y \in N, x \neq y. \tag{8}$$

Let us assume that there is $v_r^u(k, l)$ such that $v_r^u(k, l) > 2\bar{C}_r/n$ in S_r . It follows from the above assumption and from (8) that for each successor S_s of the network S_r the condition $v_s(k, l) = v_s(l, k) > 2\bar{C}_r/n$ is satisfied.

Then, it follows from the corollary 1 that there also exist maximal flows less than $2\bar{C}_r/n$ in the network S_s . In such case, the value $V(X_s)$ may be bounded using condition (6) only for $x, y \in N$ such that $x \neq y$ and $(x, y) \neq (k, l)$. Then, taking into account the conditions (5), (8) and that $c_s(N, N) \leq 2\bar{C}_r$ we obtain

$$V(X_s) \leq \frac{2(n-1)\bar{C}_r - (v_r^u(k, l) + v_r^u(l, k))}{n(n-1) - 2}. \tag{9}$$

Next, if there exists $v_r^u(a, b)$, $(a, b) \neq (k, l)$ greater than the right hand side of (9) then the above procedure should be done again. Based on (9) the following algorithm for calculation the upper bound UB_r may be formulated.

Step 1. Compute \bar{C}_r . Let $B = \{(x, y) : x, y \in N, x \neq y\}$, $e = 2\bar{C}_r/n$, $a=0$ and $b=0$.

Calculate all $v_r^u(x, y)$ in S_r^u .

Step 2. If there are the pairs (x, y) such that $v_r^u(x, y) > e$ in the set B then perform $a = a + v_r^u(x, y)$, $b = b + 1$, $B = B - \{(x, y)\}$ for each pair (x, y) and go to step 3. Otherwise go to step 4.

Step 3. Calculate $e = (2(n-1)\bar{C}_r - a)/(n(n-1) - b)$ and go to step 2.

Step 4. Perform $UB_r = e$ and stop.

5 Branching Rules

The purpose of branching rules is to find the normal arc from X_r for complementing and generating a successor of X_r with the possibly greatest value of the min-max flow.

The arc $w_j^i = \langle x, y \rangle$ is adjacent to the cut (X, \bar{X}) in the network S_r if there exists $x \in X$ and $y \in \bar{X}$. The minimum capacity cut which is connected with the min-max flow is called the min-max cut. It follows from the *max-flow – min-cut theorem* [2] that if we want increase the maximal flow in the network then the minimum capacity cut must increase. Then, it is easy to observe that the following corollary is true.

Corollary 2. Let $X_r \in R$. If there exists a network S_s (successor of S_r) such that $V(X_s) > V(X_r)$ then the selection X_s contains at least one reverse arc $w_j^i \in W - X_r$ of the normal arc $w_k^i \in X_r$ adjacent to the min-max cut in S_r .

Let $E_r = X_r - F_r$ be the set of all normal arcs which may be complemented by reverse arcs in X_r , and let M_r be the set of all reverse arcs of normal arcs from the set E_r which are adjacent to min-max cut in S_r . Also let B be the subset of the set N , and let $z(x, B)$ be the number of node a such that $a \in B$ and $v_r(x, a) = \min_{b \in B} v_r(x, b)$.

Theorem 1. Let $X_r \in R$ and let (Z, \bar{Z}) be the min-max cut in S_r . If S_s is obtained from S_r by complementing the normal arc w_k^i by the reverse arc $w_j^i = \langle x, y \rangle \in M_r$ then $V(X_s) \geq V(X_r) + \Delta_{kj}^{ir}$, where $\Delta_{kj}^{ir} = \min(\Delta_{kj}^{ir,1}, c(w_j^i) - c(w_k^i), \Delta_{kj}^{ir,2})$, $\Delta_{kj}^{ir,1} = v_r(z(x, Z), x) - V(X_r)$ and $\Delta_{kj}^{ir,2} = v_r(z(y, \bar{Z}), y) - V(X_r)$.

Proof. Let $e \in Z, t \in \bar{Z}$ and let (Q, \bar{Q}) be the minimal capacity cut separating nodes x and e in S_r (Fig. 1). Then $v_r(e, t) = V(X_r)$. After complementing w_k^i by w_j^i we can located "additional flow" $\Delta v_r(e, t)$ from e to t on paths from e to t which contain the arc w_j^i because it is the one way to increase the maximal flow $v_s(e, t)$ in S_s . Then, after complementing the value $v_s(e, t) - v_r(e, t)$ will be greater or equal to minimum of the three values: the value of additional flow from e to x , the difference between capacities of arcs w_j^i and w_k^i and the value of additional flow from y to t . Thus,

$$v_s(e, t) - v_r(e, t) \geq \min(\Delta v_r(e, x), c(w_j^i) - c(w_k^i), \Delta v_r(y, t)). \tag{10}$$

Assume that $\bar{Z} \subset \bar{Q}$. Otherwise we obtain the same results. Let $U = Z \cap Q$ and $D = Z \cap \bar{Q}$ (Fig. 1). It is easy to observe that $c(Z, \bar{Z}) = c(U, \bar{Z}) + c(D, \bar{Z})$. The additional flow from the node e to the node x is $\Delta v_r(e, x) = c(U, D) - c(U, \bar{Z})$. Because $(U, \bar{Z}) \subset (Q, \bar{Q})$ and $(D, \bar{Z}) \not\subset (Q, \bar{Q})$ then $c(U, \bar{Z}) \leq c(D, \bar{Z})$. Hence,

$$\begin{aligned} \Delta v_r(e, x) &= (c(U, D) + c(U, \bar{Z})) - (c(U, \bar{Z}) + c(D, \bar{Z})) + c(D, \bar{Z}) - c(U, \bar{Z}) = \\ &= v_r(e, x) - v_r(e, t) + c(D, \bar{Z}) - c(U, \bar{Z}) \geq v_r(e, x) - v_r(e, t). \end{aligned} \tag{11}$$

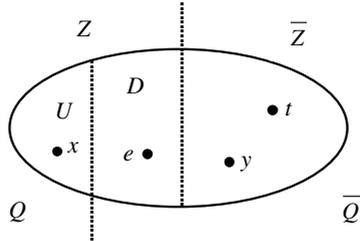


Fig. 1. The location of nodes and sets in the set of all nodes \$N\$

Taking into account that \$v_r(e, t) = V(X_r)\$ we obtain \$\Delta v_r(e, x) \geq v_r(e, x) - V(X_r)\$. The minimum value of the right hand side of the last inequality is for \$e=z(x, Z)\$, then

$$\Delta v_r(e, x) \geq v_r(z(x, Z), x) - V(X_r) = \Delta_{kj}^{ir,1}. \tag{12}$$

Analogically we obtain that

$$\Delta v_r(y, t) \geq v_r(z(y, \bar{Z}), y) - V(X_r) = \Delta_{kj}^{ir,2}. \tag{13}$$

Substituting (12) and (13) into (10) and taking into account that \$v_r(e, t) = V(X_r)\$ we obtain that \$v_s(e, t) - V(X_r) \geq \min(\Delta_{kj}^{ir,1}, c(w_j^i) - c(w_k^i), \Delta_{kj}^{ir,2})\$. Then

$$v_s(e, t) \geq V(X_r) + \Delta_{kj}^{ir}. \tag{14}$$

To find the upper bound of \$V(X_s)\$ we should bounded the value of maximal flow between any pair of nodes \$a\$ and \$b\$ in \$S_s\$. There are three possible cases:

- A. \$a \in Z, b \in \bar{Z}\$. We obtain, using (14) that \$v_s(a, b) \geq V(X_r) + \Delta_{kj}^{ir}\$,
- B. \$a, b \in Z\$. Taking into account the triangle inequality [2] we obtain \$v_s(a, b) \geq \min(v_s(a, t), v_s(b, t))\$. Since values \$v_s(a, t)\$ and \$v_s(b, t)\$ are greater than right hand side of (14) then \$v_s(a, b) \geq V(X_r) + \Delta_{kj}^{ir}\$,
- C. \$a, b \in \bar{Z}\$. Analogically to case B: \$v_s(a, b) \geq \min(v_s(e, a), v_s(e, b)) \geq V(X_r) + \Delta_{kj}^{ir}\$.

By definition the value of min-max flow is not greater than the value of the maximal flow between each pair of nodes, then \$V(X_s) \geq V(X_r) + \Delta_{kj}^{ir}\$. ■

The choice criterion for complementing is as follows: \$\delta_{kj}^{ir} = |V(X_r) + \Delta_{kj}^{ir} - UB_r|\$.

6 Approximate Algorithm

The exact algorithm involves to find the initial selection X_1 such that $X_1 \in R$ and for which the condition (4) is satisfied. Moreover, the initial selection X_1 should be the near-optimal solution of the problem (2-4). To find the feasible initial selection X_1 as above the following approximate algorithm is proposed.

Step 1. Assign maximum available capacities arcs to X_1 i.e. perform $i=1$ and $X_1^{(i)} = \{w_1^1, w_1^2, \dots, w_1^m\}$.

Step 2. Compute $V(X_1^{(i)})$. If $d(X_1^{(i)}) \leq DB$ then the feasible initial selection is found. Perform $X_1 = X_1^{(i)}$ and the algorithm terminates. Otherwise go to step 3.

Step 3. Choose the arcs $w_k^l = \langle x, y \rangle \in X_1^{(i)}$ and $w_{k+1}^l \in W^l - X_1^{(i)}$ for which the value $v_i(x, y) - c(w_k^l) + c(w_{k+1}^l) - V(X_1^{(i)})$ is maximal; $v_i(x, y)$ denotes the maximal flow from the node x to the node y in $S_1^{(i)} = \langle G_1^{(i)}; c, d \rangle$ where $G_1^{(i)} = \langle N, A \cup X_1^{(i)} \rangle$. Perform $i=i+1$ and $X_1^{(i)} = (X_1^{(i-1)} - \{w_k^l\}) \cup \{w_{k+1}^l\}$. Go to step 2.

Using the presented approximate algorithm we obtain the approximate (near optimal) solutions in a short computation time and using the presented exact algorithm we obtain the optimal solutions in a long computation time.

7 Computational Results

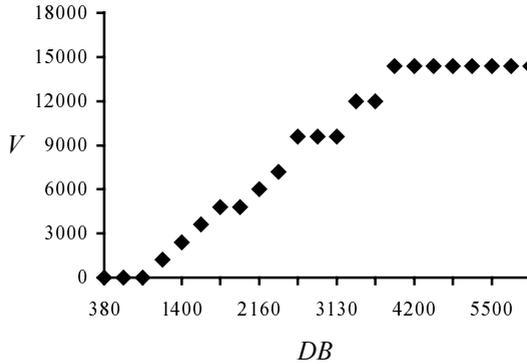
Extensive numerical experiments were performed with presented algorithms for many different network topologies. The experiments were conducted with two main purposes in mind: first, to examine the impact of various parameters on solutions (i.e. min-max flow) and second, to test the computational efficiency of the algorithms.

The typical dependence of optimal min-max flow value on budget is presented in the Fig. 2. It follows from results obtained for many considered networks that the optimal min-max flow value in terms of budget is the function belonging to following class:

$$V = \begin{cases} 0 & \text{for } 0 \leq DB < D_1 \\ a \cdot DB + b & \text{for } D_1 \leq DB \leq D_2 \\ V_{\max} & \text{for } D_2 < DB \leq DB_{\max} \end{cases} \quad (15)$$

where a, b, D_1 and D_2 are constant coefficients, V_{\max} and DB_{\max} are maximal feasible min-max flow value and maximal budget in the network respectively.

It is easy to see that $D_1 = -b/a$ and $D_2 = (V_{\max} - b)/a$. Then, to find the function (15) for some network it is enough to know only the values of a and b . To compute a and b the identification problem must be solved e.g. using the Last Squares Method.



normalized budget $NDB = (DB / DB_{\max}) \cdot 100\%$ which was introduced to compare the results obtained for different network topologies. It follows from numerical experiments (Fig. 4) that the exact algorithm is especially effective from computation point of view for $NDB \in [0,20] \cup [40,100]$.

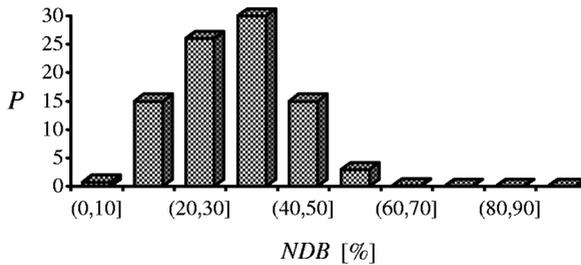


Fig. 4. The dependence of number of iteration P on normalized budget NDB

8 Conclusion

In the paper the exact and approximate algorithms are presented. It follows from computational experiments that about 60% approximate solutions are equal to optimal solutions (Fig. 3). We are of the opinion that the very good performance of the approximate algorithm obviates, in most practical cases, the application of the exact algorithm, that is for more time-consuming. Moreover, the proposed algorithms may be also applied for solving the network design problem maximizing the number of different routes between each pair of nodes subject to budget constraint. In this case, each set W^i contains only the empty arc and the arc the capacity of which is equal to one.

References

1. Amiri A.: A System for the Design of Packet-switched Communication Networks with Economic Tradeoffs. *Computer Communications* 21 (1998) 1670–1680
2. Ford L. R., Fulkerson D. R.: *Flows in Networks*. Princeton University Press, Princeton, New Jersey (1962)
3. Kasprzak A.: *Topological Design of the Wide Area Networks*. Wroclaw University of Technology Press, Wroclaw (2001)
4. Minoux M.: *Synthesis and Optimum Network Design Problems: Models, Solution Methods and Applications*. *Networks* 19 (1989) 313–360
5. Plesnik J., *The Complexity of Design a Network with Minimum Diameter*. *Networks* 11 (1981) 77–85