Journal of Global Optimization **24:** 79–88, 2002.

A Concavity Result for Network Design Problems

S. KETABI[1] and F.J.M. SALZBORN[2]
[1]*The University of Isfahan, Isfahan, Iran (e-mail: sketabi@cc.iut.ac.ir);* [2]*The University of Adelaide, Adelaide, Australia*

Abstract. The Network Design Problem has been studied extensively and in many of these models the cost is assumed to be a concave function of the loads on the links. In this paper we investigate under which conditions this is indeed the case for the communication networks. The result is presented as a theorem, the Concavity Theorem, and a list of conditions that can easily be verified. It is also shown how the theorem can be extended to other applications, like in the area of road transportation.

Key words: Multicommodity network, concave costs, transportation, telecommunication.

1. The Network Design Problem

Consider a graph $\Psi = (N, L)$. Traffic between two nodes is called a commodity and for each commodity k the offered traffic t_k is specified. Also for each commodity a set of possible paths P_k is specified that can be used for that commodity; for instance this may simply be all possible paths between the two nodes. The allocation of traffic to the paths is called routing and it is assumed that all offered traffic is routed. To accommodate the resulting traffic in the network, capacities have to be assigned to the links of the paths, which are physical connections between nodes. A link can be shared by several paths and the load on the link is the sum of the traffic allocated to the paths that use it.

If f_e is the load on link $e \in L$ and h_μ is the amount of traffic assigned to path μ then these assumptions can be expressed as

$$\sum_{\mu \in P_k} h_\mu = t_k, \tag{1}$$

$$h_\mu \geqslant 0, \tag{2}$$

$$f_e = \sum_{\mu | e \in \mu} h_\mu. \tag{3}$$

The *load* on the network is the vector of all link loads $f = (f_e | e \in L)$ and Ω is the set of all loads that can occur as the result of any routing of the offered traffic. Note that Ω is a convex set.

The capacity of a link is interpreted differently from model to model but we assume that it can be expressed by a single positive number c_e and that $c_e = 0$ indicates that the link is not available. So this gives rise to another vector, the *capacity* vector $c = (c_e | e \in L)$. The relationship between the load f and the capacity c is assumed to be expressed by a cost function $F(f, c)$ and a set of constraints $(f, c) \in S$. The capacity constraints may simply be $f_e \leqslant c_e$ or may be missing all together, but some models have much more elaborate constraints.

The Network Design Problem in its most general form is then

$$\min F(f, c) \text{ such that } f \in \Omega, (f, c) \in S. \tag{4}$$

This problem can be tackled in many ways and there exists an extensive literature on this subject [6, 10, 17, 20].

An important step in obtaining a practical mathematical formulation for a network design problem is to correctly model the interaction between the loads and capacities, and the performance and design costs. The performance cost is the operational cost related to transporting the flow on the links that may include the travel time, the delay time, the cost of losing customers (because of the lack of capacity) or even the maintenance costs. The design cost is the (monetary) cost of constructing the link capacity or facilities to provide a certain rate of service. The cost function $F(f, c)$ consists of these two components.

The solution methods depend of course heavily on the specification of the cost function F and the constraints S. This problem may be considered as a bilevel optimization [4]. Therefore many of these methods, but not all, start by eliminating the capacity variables c [8, 15] ; the remaining problem is then a routing problem with the objective function

$$H(f) = \min_{\{c|(f,c)\in S\}} F(f, c). \tag{5}$$

This can be regarded as a multicommodity flow problem, which is important not only due to the major relevance of its applications [1], but also because it poses considerable modeling and algorithmic challenges [3, 10, 16, 20].

In many of the models the function $H(f)$ turns out to be concave. This is assumed in a number of applications including production and inventory planning, transportation and communication network design, facilities location and VLSI design [13, 21] . Each application area gives rise to problems with different features for both the costs and the feasible set. In each case, the concavity of the monetary costs originates from start–up costs, discounts or economies of scale. It is the purpose of this paper to investigate under which assumptions the objective function $H(f)$ is concave . The aim is to make the assumptions as general as possible so that many models are covered.

2. Store-and-forward communication networks

In communication systems both the load f_e (which is the expected load in a stationary situation) and the capacity c_e (which is the service rate) are measured in the same units; the bit rate. The service rate of a link depends on the facilities which are installed on the link and the adjacent nodes; the load depends on the offered traffic and the routing.

The cost function $F(f, c)$ consists of two components:

$$F(f, c) = G(f, c) + X(c). \tag{6}$$

The component $X(c)$ is the cost of providing capacity for the links, and therefore does not depend on the load: this is called the *design* cost. The other cost, $G(f, c)$ is the *performance* costs and is a measure of how well the network operates; this is a penalty cost associated with the delays in the network, which of course depends on both the loads and the capacities. A class of problems which has been widely studied is determining both the loads and capacities to minimize the performance and delay costs. These problems are difficult to solve because typically they are nonconvex [7].

The average delay on a link is defined as the mean time that each packet of data will have to wait to be transmitted through that link. The formula for the average delay $D_e(f_e, c_e)$ given below (in the steady state case where $f_e < c_e$) can be obtained by assuming Poisson arrival and service times:

$$D_e(f_e, c_e) = \frac{1}{c_e - f_e}. \tag{7}$$

Then, assuming that the loads on the links are independently distributed, one obtains for the total delay in the network the formula [11]

$$G(f, c) = \sum_e \frac{f_e}{c_e - f_e} = \sum_e \frac{f_e/c_e}{1 - f_e/c_e}. \tag{8}$$

Of course this is only a crude approximation of the actual delay, but the approximation improves when the size of the network increases. This can therefore be used to measure the performance and an appropriate multiple of this function can then be used as the performance component of the objective function $F(f, c)$.

3. The Concavity Result

In this section, we will present the assumptions and proof of the concavity theorem for communications networks. The assumptions are so general that can cover other networks as well. Also an example with a specific design cost is presented that will approve the result for the objective function.

3.1. ASSUMPTIONS

We will start with listing the assumptions for communication networks which are needed to prove concavity of the function $H(f)$. The following extra notation will be used.

For a given vector $x = (x_e | e \in L) \geqslant 0$, the set of all links with a positive value $L(x) = \{e | x_e > 0\}$ is called the *support* of x.

We will use the sign * for element–by–element vector multiplication. That means, if $x = (x_e|e \in L)$ and $y = (y_e|e \in L)$ then

$$x * y = (x_e y_e|e \in L). \tag{9}$$

1. To start with it has to be assumed that the function $H(f)$ exists at all! In other words that there exists for each load $f \in \Omega$ a capacity vector $c(f)$ with $(f, c(f)) \in S$ such that

$$H(f) = F(f, c(f)) \geqslant F(f, c) \text{ for all } c \text{ such that } (f, c) \in S. \tag{10}$$

2. It is assumed that it never pays to provide capacity at a link that is not used at all and that capacity has to be assigned to links that carry traffic, in other words that

$$L(f) = L(c(f)). \tag{11}$$

3. The design cost $X(c)$ is a concave function. This is a condition commonly assumed for these types of costs, reflecting economies of scale. Although in some practical applications, like the facility installing in communication and telecommunication networks, the costs are step functions, but they can be well approximated by concave functions.
4. The following assumption is made concerning the term $G(f, c)$

$$0 \leqslant G(\alpha * f, \alpha * c) \leqslant G(f, c), \text{ for all } \alpha \geqslant 0. \tag{12}$$

This assumption needs some clarification. Firstly, it is easily seen that it implies that $G(\alpha * f, \alpha * c) = G(f, c)$ if $L(f) = L(c)$ and if, for all $e \in L(f)$, $\alpha_e > 0$. This is so because for

$$\bar{\alpha}_e = \begin{cases} \dfrac{1}{\alpha_e} & \text{if } e \in L(f) \\ 0 & \text{if } e \notin L(f) \end{cases},$$

applying 12 twice, we get

$$G(f, c) \geqslant G(\alpha * f, \alpha * c) \geqslant G(\bar{\alpha} * \alpha * f, \bar{\alpha} * \alpha * c) = G(f, c). \tag{13}$$

This means that if all links with capacity have a load and all links with a load have capacity (for instance if $c = c(f)$), then $G(f, c)$ is a function of the ratios $\frac{f_e}{c_e}$ on the used links i.e. the links in $L(f) = L(c)$, as long as the topology of the network does not change.

Secondly, we note that the function G does not increase if the loads and capacities of some links are set equal to zero. In other words, for every given vector (f, c), if $(\bar{f}, \bar{c})$, corresponding to an arbitrary set $L^0 \subset L$, is defined by

$$(\bar{f}_e, \bar{c}_e) = \begin{cases} (0, 0) & \text{if } e \in L^0 \\ (f_e, c_e) & \text{if } e \notin L^0 \end{cases},$$

then as a result of 12, choosing $\alpha_e = 1$ for $e \notin L^0$ and $\alpha_e = 0$ for $e \in L^0$, we have

$$G(\bar{f}, \bar{c}) \leqslant G(f, c). \tag{14}$$

Although this assumption seems restrictive, the performance costs naturally do not increase if both the flow and the capacity are multiplied by the same factor. Indeed in many practical cases, theses costs are varied as functions of the ratio f/c.

5. The term $Y(f)$ as a multiplier in the performance cost function is introduced to cover certain models of road transportation as well. For all other models (and more naturally) this term vanishes as $Y(f) = 1$. However to make our model as general as possible we will assume that $Y(f)$ is a linear function and that $Y(f) \geqslant 0$ if $f \in \Omega$.
6. The set S is assumed to have the property that

$$\text{if } (f, c) \in S \text{ then } (\alpha * f, \alpha * c) \in S \text{ for all } \alpha \geqslant 0. \tag{15}$$

Note that this is for instance true if the only constraints are $f \leqslant c$.

Now we can state the main result of this paper.

3.2. THE CONCAVITY THEOREM

THEOREM 1. *If all conditions listed above are satisfied then*

$$H(f) = \min_{\{c|(f,c)\in S\}} F(f, c)$$

is a concave function of $f \in \Omega$.

Proof of the theorem. Choose any two feasible loads f^1 and f^2, and $0 < \lambda < 1$ and let

$$f^0 = \lambda f^1 + (1 - \lambda) f^2. \tag{16}$$

Since Ω is a convex set, $f^0 \in \Omega$. It will be shown that

$$\lambda H(f^1) + (1 - \lambda) H(f^2) \leqslant H(f^0). \tag{17}$$

Let c^i be the optimal capacity corresponding to f^i, so that

$$H(f^i) = Y(f^i) G(f^i, c^i) + X(c^i), \text{ for } i = 0, 1, 2. \tag{18}$$

Define the components α_e^i of the vectors α^i for $i = 1, 2$ by

$$\alpha_e^i = \begin{cases} \dfrac{f_e^i}{f_e^0} & \text{if } f_e^0 \neq 0 \\ 0 & \text{if } f_e^0 = 0 \end{cases}. \tag{19}$$

Notice that $f_e^0 = 0$ if and only if f_e^1 and f_e^2 are both zero. Hence for $i = 1, 2$,

$$f^i = \alpha^i * f^0 \tag{20}$$

and

$$\lambda\alpha_e^1 + (1-\lambda)\alpha_e^2 = \begin{cases} 1 & \text{if } f_e^0 \neq 0 \\ 0 & \text{if } f_e^0 = 0 \end{cases} . \tag{21}$$

Note that because $(f^0, c^0) \in S$, by using 20, we will have

$$(f^i, \alpha^i * c^0) = (\alpha^i * f^0, \alpha^i * c^0) \in S, \tag{22}$$

and, because c^i is the optimal capacity for f^i $(i = 1, 2)$,

$$F(f^i, c^i) \leqslant F(f^i, \alpha^i * c^0). \tag{23}$$

Therefore using 23, 22 and 12, we have

$$\lambda H(f^1) + (1-\lambda)H(f^2) = \lambda F(f^1, c^1) + (1-\lambda)F(f^2, c^2) \tag{24}$$

$$\leqslant \lambda F(f^1, \alpha^1 * c^0) + (1-\lambda)F(f^2, \alpha^2 * c^0) \tag{25}$$

$$= \lambda Y(f^1)G(\alpha^1 * f^0, \alpha^1 * c^0) + \lambda X(\alpha^1 * c^0) \tag{26}$$

$$+ (1-\lambda)Y(f^2)G(\alpha^2 * f^0, \alpha^2 * c^0) + (1-\lambda)X(\alpha^2 * c^0) \tag{27}$$

$$\leqslant (\lambda Y(f^1) + (1-\lambda)Y(f^2))G(f^0, c^0) \tag{28}$$

$$+ \lambda X(\alpha^1 * c^0) + (1-\lambda)X(\alpha^2 * c^0) \tag{29}$$

$$\leqslant Y(f^0)G(f^0, c^0) + X((\lambda\alpha^1 + (1-\lambda)\alpha^2) * c^0), \tag{30}$$

where, in the last inequality, the concavity of X, linearity of Y and non–negativity of G have been used. Now note that assumption 2 implies that $c_e^0 = 0$ if $f_e^0 = 0$. Hence from 21 it can be concluded that

$$(\lambda\alpha^1 + (1-\lambda)\alpha^2) * c^0 = c^0. \tag{31}$$

This completes the proof.

3.3. THE CONCAVITY RESULT FOR COMMUNICATION NETWORKS

In this section we will approve the conditions of the theorem for the communication networks. The important observation is that $G(f, c)$ in equation (8) depends only on the ratios f_e/c_e and therefore the assumptions of the Concavity Theorem are satisfied, either if this function is used as the performance cost (i.e., if $F(f, c) = G(f, c) + X(c)$) or if it is used to restrict the delay to a specified limit (i.e., $F(f, c) = X(c)$ and constraint $G(f, c) \leqslant q$).

Even if a more elaborate expression is used for the total delay, it is still likely to depend on the ratios f_e/c_e only and the argument put forward here is still valid: the Concavity Theorem still applies, of course assuming that all the other conditions which are not related to the performance cost also hold, which effectively means that $X(c)$ must be concave as for this model $Y(f) = 1$.

Note that the function H need not be separable in the links even if the delay and the design costs are. For instance, for the function

$$H(f) = \min_c \sum_{e \in L} \gamma_e c_e \mid G(f, c) \leqslant q,\ f \in \Omega \tag{32}$$

where $G(f, c)$ is defined by Kleinrock's formula, one can derive an explicit expression

$$H(f) = \sum_e \gamma_e f_e + \frac{1}{q}(\sum_e \sqrt{\gamma_e f_e})^2. \tag{33}$$

which is not separable into link costs, but as we know from the Concavity Theorem must be concave.

4. Application of the result to transportation networks

In a transportation system the average delay $D(f_e, c_e)$ on a link is often assumed to be a function of the ratio f_e/c_e [18], and the total delay for the link is then $f_e D(f_e/c_e)$, unlike the model discussed above for a communication network, where the total delay is a function of the ratio $\frac{f_e}{c_e}$. If the total delay in the network is selected again as the performance measure

$$G(f, c) = \sum_{e \in L} f_e D(\frac{f_e}{c_e}) \tag{34}$$

then it is clear that this is not of the form required for the Concavity Theorem. However the theorem still applies if we assume that the design cost is also separable in the links, i.e. if the objective function to be minimized is

$$F(f, c) = \sum_{e \in L} F_e(f_e, c_e) = \sum_{e \in L} [f_e D(\frac{f_e}{c_e}) + X_e(c_e)] \tag{35}$$

and if there are no constraints involving more than one link, because in that case the function $H(f)$ is also separable in the links,

$$H(f) = \sum_{e \in L} H_e(f_e) = \sum_{e \in L} \min_{c_e} [f_e D(\frac{f_e}{c_e}) + X_e(c_e)], \tag{36}$$

and the capacity of each link is optimized separately.

However for a single link the Concavity Theorem does apply, assuming that X_e is concave. This can be easily verified by substituting of f_e as $Y(f)$ in the theorem, *i.e.* $F_e(f_e, c_e) = f_e D(f_e/c_e) + X_e(c_e)$ is then of the required form. So the theorem states that then $H_e(f_e)$ is concave for every link e and therefore $H(f)$ is concave.

4.1. AN EXAMPLE

Consider a network containing one link ($|L| = 1$). Suppose that G is a differentiable, increasing and strictly convex function where $G(0) = 0$, $Y(f) = f$, and X is a piecewise linear concave function as defined below:

$$X(c) = \begin{cases} a_1 c & \text{if } 0 \leqslant c \leqslant c^* \\ a_2 c + (a_1 - a_2)c^* & \text{if } c > c^* \end{cases},$$

where $c^* > 0$ and $a_1 > a_2 > 0$. It is easily verified [2] that H is a piecewise linear and concave function as follows:

$$H(f) = \begin{cases} (G(y_1) + \dfrac{a_1}{y_1})f & \text{if } f \leqslant f^* \\ (G(y_2) + \dfrac{a_2}{y_2})f + (a_1 - a_2)c^* & \text{if } f > f^* \end{cases},$$

where y_i is the unique solution of $y^2 G'(y) = a_i$ ($i = 1, 2$), $y_1 > y_2 > 0$ and

$$f^* = \frac{(a_1 - a_2)c^*}{G(y_1) - G(y_2) + \frac{a_1}{y_1} - \frac{a_2}{y_2}}.$$

5. Discussion

Although we have only discussed two examples, there are no doubt many network design models where the Concavity Theorem applies. Spelling out the conditions as we have done in this paper should make it easier to recognize such situations. In general these are the models where the performance function is a measure of the average delay in the network and the design cost is concave in terms of the capacity. In these models the capacity c_e on a link is interpreted as the service rate that is provided and the stochastic model underlying the derivation of the delay is based on Poisson distributions for the loads, which means that the delays depend on the ratios f_e/c_e only.

There are of course also many models where the theorem is not valid. This is in particular the case when the capacity is interpreted as a physical quantity (for instance number of circuits). Then there is no reason to assume that the delay depends on the ratios f_e/c_e only and the conditions of the theorem will usually not apply. Also if a different performance measure is selected (for instance lost calls, as in telephone networks) then the conditions may not hold.

When the theorem is valid, the design problem may be converted to a routing problem with a concave cost objective function. Then this has several consequences. Firstly, because of the concavity of the objective function, local optimality does not imply global optimality. This indicates that finding the overall minimum corresponds to searching all the extreme points of the feasible set. This will make the problem of finding a minimum cost network very difficult, computationally, as there usually will be a large number of local optima.

Then, secondly, one probably has to settle for an approximate solution found by using a random search procedure or a construction method that cannot guarantee optimality. At best, one can hope that a proper lower bounding procedure is available. Many such methods are described in the literature [1, 5, 9, 12, 14, 19].

However concavity of the objective function also has another consequence by exploring the characterization of the extreme points. This is that one only has to look for single path routing, which means that all traffic of one commodity is allocated to one single route [13]. It is a well-known fact that there is such an optimal routing if the objective function (as a function of the load) is concave. This observation may simplify the computations considerably and existing methods invariably make use of this fact.

References

1. Ahuja, R. K., Magnanti, T. L. and Orlin, J. O. (1993), *Network Flows*, Prentice Hall, Englewood Cliffs.
2. Ashtiani, H. Z. (1989), On The Objective Function for Network Design Problems, Technical Report, Isfahan University of Technology, Unpublished.
3. Assad, A. A. (1978), Multicommodity Network Flows – A Survey, *Networks*, 8, 37–91.
4. Bard, J. F. (1998), *Practical Bilevel Optimization*, Kluwer Academic Publishers, Dordrecht.
5. Balakrishnan, A. and Graves, S. C. (1989), A Composite Algorithm for a Concave-Cost Network Flow Problem, *Networks*, 19, 175–202.
6. Bertsekas, D. and Gallager, R. (1992), *Data Networks*, Prentice Hall, Englewood Cliffs, NJ.
7. Bretthauer, K. M. (2000), Optimal Service And Arrival Rates in Jackson Queueing Networks, *Naval Res. Logis.*, 47, 1–17.
8. Dantzig, G. D., Harvey, R., Lansdowne, Z., Robinson, D. and Maier, S. Formulation and Solving the Network Design Problem by Decomposition, *Transportation Research B*, 13 5–17.
9. Florian, M. (1986), Nonlinear Cost Network Models in Transportation Analysis, *Mathematical Programming Study*, 26, 176–196.
10. Gendron, B. and Crainic, T. and Frangioni, A. (1998), Multicommodity Capacitated Network Design, in B. Sans‘o and P. Soriano (eds), *Telecommunications Network Planning*, pp. 1–19, Kluwer, Norwell.
11. Gerla, M. and Kleinrock, L. (1977), On the Topological Design of Distributed Computer Networks, *IEEE Transaction on Communications* 25, 48–60.
12. Guisewite, G. M. (1995), Network Problems, in R. Horst and P. M. Pardalos (eds), *Handbook of Global Optimization*, pp. 609–648. Kluwer, Dordrecht.
13. Guisewite, G. M. and Pardalos, P. M. (1990), Minimum Concave–cost Network Flow Problems: Applications, Complexity, and Algorithms, *Annals of Operations Research* 25, 75–100.
14. Homqvist, K. and Migdalas, A. and Pardalos, P. M. (1998), A Grasp Algorithm for the Single Source Uncapacitated Minimum Concave-Cost Network Flow Problems, in P. M. Pardalos and

D. Z. Du (eds.), *Network Design: Connectivity and Facilities Location*, *DIMACS Series on Discrete Mathematics and Theoretical Computer Science*, 40, 131–142.

15. Humes, C. J. (1992), A Projection-Feasible Direction Method for Continuous Capacity and Flow Assignment, *Mat. Apl. Comput.*, 11, 243–254.
16. Kennington, J. L. (1978), A Survey of Linear Cost Multicommodity Network Flows, *Operations Research* 26, 209–236.
17. T. L. Magnanti and R. T. Wong, Network design and transportaion planning: models and algorithms, Transportation Science, 18:1–55, 1984.
18. Manheim, M. L. (1979), *Fundamentals of Transportaion System Analysis*, MIT Press, Cambridge, MA.
19. Minoux, M. (1976), Multiflots de cout concaves, *Annales Des Telecommunications*, 31, 77–92.
20. Minoux, M. (1989), Network Synthesis and Optimum Network Design Problems: Models, Solution Methods and Application, *Networks*, 19, 313–360.
21. Runggeratigul, S. and Tantaratana, S. (1999), Link Capacity Assignment in Packet-switched Networks: The Case of Piecewise Linear Concave Cost Function, *IEICE Transactions on Communications*, E82–B(10), 1566–1576.