

Price Method and Network Congestion Control

Krzysztof Malinowski, Ewa Niewiadomska-Szynkiewicz, and Przemysław Jaskóła

Abstract—Price instruments are useful in achieving market balance conditions in various markets. Those instruments can be also used for control of other composite systems. The formulation and basic properties of the Price Method are reviewed and then the congestion control by price instruments in a computer network is described and tested.

Keywords—computer network, congestion control, coordination strategy, price, price instruments, retail pricing problem.

1. Introduction

In this paper the objective is to present emerging, opportunities to use price mechanisms for management of computer networks. In recent years it was observed that the use of the price instruments could be made both to propose the new techniques for congestion control in data networks, in particular for Internet congestion control (ICC), and to better explain the existing congestion control mechanisms (the current TCP (transmission control protocol) congestion control protocols). It is useful to note at this point that Internet pricing is essential not only for better possible understanding of the network operation but it may also provide means to achieve rational behavior of the network users, who otherwise may “overgraze” the existing resources. Proper pricing of the network services is also necessary for many other purposes, not only for congestion control [1], [2]. Dynamic pricing may be useful, in particular, for balancing demand for access to application servers and for proper valuation of different classes and qualities of service [3].

2. Price Method; General Formulation and the Basic Facts

The Price Method, also known under the name of the Interaction Balance Method [4]–[7] represents a versatile approach to modeling, optimization and management of complex systems. This follows the natural role played by prices on various markets – to achieve the balance between the demand and the supply. The method allows for various problem formulations and price adjustment (price coordination) strategies. Problem formulation may either result from a formal partitioning (decomposition) of a large-scale optimization problem or it may result from more practical, application case-oriented, considerations. Price adjustment strategies can be derived from dual function optimization,

from solving sets of the coordinating conditions or from practical possibilities existing in a particular application.

Consider the following, fairly general, deterministic problem related to complex system optimization:

$$\max_{x,u,y} \sum_i U_i(x_i, u_i, y_i), \quad (1)$$

$$\text{subject to } x_i \in I_i, \quad i = 1, \dots, N, \quad (2)$$

$$y_i = F_i(x_i, u_i) \quad \text{and} \quad u_i = \sum_j H_{ij} y_j, \quad i, j = 1, \dots, N, \quad (3)$$

$$\sum_i r_{li}(x_i) \leq c_l, \quad l = 1, \dots, L. \quad (4)$$

In the above formulation the objective is to maximize the aggregate utility across N involved entities (system elements); x_i is the vector of local decision variables (of dimension n_{x_i}), u_i is the local (vector) interaction input (dim $u_i = n_{u_i}$) and y_i (dim $y_i = n_{y_i}$) is the local interaction output of the i th system element. In Eq. (3) it is assumed that the interaction output y_i is a unique deterministic function of x_i and u_i , and that the interaction coupling constraints $u_i = \sum_j H_{ij} y_j$ are linear. Obviously, these relations can be made more general. Similarly, local decision constraint sets I_i can be imposed also on u_i . Often these sets are in form of box constraints $I_i = [x_i^{\min}, x_i^{\max}]$ (with x_i^{\min} being the vector of lower bounds and x_i^{\max} the vector of upper bounds on the components of x_i). Constraint (4) is the global resource constraint; there are assumed to be L resources which may be required by the local entities; $r_{li}(x_i)$ represents the consumption of the l th resource by system element i . It is assumed that the optimization problem (1)–(4) has a solution.

The above optimization problem can be decomposed into N parallel local problems by introducing prices (in mathematical terms Lagrange multipliers) λ_i (dim $\lambda_i = n_{u_i}$), $i = 1, \dots, N$, associated with the coupling constraints $u_i = \sum_j H_{ij} y_j$, and μ_l , $l = 1, \dots, L$, representing prices of the global resources. The i th local problem is then defined – for given prices – as follows:

$$\max_{x_i, u_i} \left[U_i(x_i, u_i, y_i) - \lambda_i^T u_i + \sum_{j=1}^N \lambda_j^T H_{ji} y_j - \sum_{l=1}^L \mu_l r_{li}(x_i) \right], \quad (5)$$

where $y_i = F_i(x_i, u_i)$ and subject to $x_i \in I_i$. Assuming unique solutions to the above N problems, $x_i(\lambda, \mu)$, $u_i(\lambda, \mu)$ and $y_i(\lambda, \mu)$, (where $y_i(\lambda, \mu) = F_i(x_i(\lambda, \mu), u_i(\lambda, \mu))$) and $\lambda^T = (\lambda_1^T, \dots, \lambda_N^T)$, $\mu^T = (\mu_1^T, \dots, \mu_L^T)$ one can seek such

coordinating values λ^C and μ^C of λ and μ for which the coupling constraints (3) and the resource constraints (4) are satisfied, that is for $i = 1, \dots, N$

$$y_i(\lambda^C, \mu^C) = F_i(x_i(\lambda^C, \mu^C), u_i(\lambda^C, \mu^C)),$$

$$u_i(\lambda^C, \mu^C) = \sum_j H_{ij} y_j(\lambda^C, \mu^C), \quad (6)$$

$$\sum_i r_{li}(x_i(\lambda^C, \mu^C)) \leq c_l, \quad l = 1, \dots, L, \quad (7)$$

$$\mu_l^C \geq 0, \quad \mu_l^C \left[\sum_i r_{li}(x_i(\lambda^C, \mu^C)) - c_l \right] = 0, \quad l = 1, \dots, L. \quad (8)$$

The above conditions (7) and (8) result from the requirement to satisfy overall optimality conditions. In simple terms they state that the optimal prices of the resources must be nonnegative and that a positive price can be charged for the commonly available resource only when this resource is fully utilized, i.e., when the respective resource constraint is active. It should be observed that some or even all values of the components of λ_i^C may be negative. The assumption about the uniqueness property of the local problem solutions – for given price vectors μ and λ – is an essential one [4] for the conditions guaranteeing the existence of the coordinating prices λ^C, μ^C . If these solutions do not have this property, then the coordinating prices satisfying eqns. (6)–(8) may easily not exist. One could expect that the lack of uniqueness of $x_i(\lambda, \mu), u_i(\lambda, \mu)$ for an isolated point $(\lambda^0, \mu^0), \mu^0 \geq 0$ should not matter too much. Alas, it is easy to demonstrate with many important examples that if the local solutions are not unique for some pair (λ^0, μ^0) , they are also not unique for any pair (λ, μ) that could satisfy the coordinating conditions. Now, the simple sufficient conditions guaranteeing the uniqueness of $x_i(\lambda, \mu), u_i(\lambda, \mu)$ (and hence of $y_i(\lambda, \mu)$) for any feasible pair (λ, μ) are: the strict concavity of the function $U_i(x_i, u_i, y_i)$ with respect to its arguments, linearity of F_i and the convexity of r_{li} and of the set I_i . More general conditions can also be given [5], [8] but in a general nonlinear case the desirable uniqueness property of the local problem solutions is, unfortunately, not easily achievable. Assuming the existence of λ^C, μ^C , one can propose a number of algorithms for iterating the values of λ, μ . The basic strategy is to use the following gradient method:

$$\lambda_i^{(k+1)} = \lambda_i^{(k)} + \gamma \left[u_i(\lambda^{(k)}, \mu^{(k)}) - \sum_j H_{ij} y_j(\lambda^{(k)}, \mu^{(k)}) \right], \quad (9)$$

$$\mu_l^{(k+1)} = \left\{ \mu_l^{(k)} + \gamma \left[\sum_i r_{li}(x_i(\lambda^{(k)}, \mu^{(k)})) - c_l \right] \right\}_+ \quad (10)$$

for $i = 1, \dots, N, l = 1, \dots, L$, where γ is a positive step size, and k is the iteration index. The algorithm defined by (9) and (10) is a gradient strategy – with price projection on the feasible range – since, assuming the uniqueness of $x_i(\lambda, \mu), u_i(\lambda, \mu)$ for every feasible pair (λ, μ) and each i , the expressions in square brackets in (9), (10) are the slopes (with minus sign) of the dual function defined as the sum

of the maximum local performance values Eq. (5) minus the term $\sum_l \mu_l^{(k)} c_l$. The dual function attains its minimum at λ^C, μ^C .

The above price adjustment strategy will converge for sufficiently small value of γ if all local utility functions are strongly concave, functions F_i are linear, functions r_{li} are convex, sets I_i are convex and there exists a feasible point satisfying all the constraints, such that all the inequality constraints are inactive. Algorithm (9) and (10) has a distributed character. In particular, the adjustments of the resource prices μ_l can be performed for each price independently from the other price adjustments. This property appears to be most useful.

3. Network congestion control and price-based schemes

With Internet, the Price Method has found a large-scale system for which this method seems to be quite well suited. In several congestion control mechanisms, as recently proposed [9]–[14], the network is represented by S traffic sources, representing particular source-destination pairs, and a grid of a set of L links. The links, together with associated routers, are the network resources, of limited traffic carrying capacity c_l . Each source i is supposed to use a set $L(i) \subseteq L$ of links, and has at time k an associated transmission rate x_i ; the set of transmission rates determines the aggregate flow $y_l(k)$ through each link, by the equation:

$$y_l(k) = \sum_{i \in S(l)} x_i(k), \quad (11)$$

where $S(l)$ is the set of all sources transmitting through the link l . Then, the feedback mechanism communicates to sources the congestion information about the network. This congestion measure – the price $p_l(k)$ – is a positive valued quantity associated with link l . The fundamental assumption is made that sources have access to the aggregate price $(q_i(k), i = 1, \dots, S)$ of all links in their route:

$$q_i(k) = \sum_{l \in L(i)} p_l(k). \quad (12)$$

4. Network Control by Price Instruments, Application of Price Method

Let us consider the network with S sources and L links. Assume now that the traffic sources (or source-destination pairs) are indeed utility oriented, i.e., they have utilities $U_i(x_i)$ expressed in monetary terms and are willing to maximize profits equal to utility minus payment charged by the network. Then, the equilibrium rate x_i^{se} will solve the following local source problem:

$$\max_{x_i} [U_i(x_i) - q_i^{se} x_i], \quad (13)$$

where $q_i^{se} = \sum_{l \in L(i)} p_l^{se}$ and p^{se} denotes vector of equilibrium prices. The role of prices p is to coordinate the ac-

tions of the individual sources; in fact to ensure that solutions of (13) together solve the network flow optimization problem

$$\max_{x_i \geq 0} \sum_i U_i(x_i) \quad \text{subject to} \quad \sum_{i \in S(l)} x_i \leq c_l, \quad l = 1, \dots, L, \quad (14)$$

where c_l is the l th link capacity. This is particular, simple instant of problem (1)–(4). Hence, it would seem possible – at least in theory – to propose the following congestion control scheme: for given link prices $p_l(k)$, $l = 1, \dots, L$, at time k , the sources solve local problems

$$\max_{x_i \in I_i} [U_i(x_i) - q_i(k)x_i], \quad (15)$$

where $I_i = [x_i^{\min}, x_i^{\max}]$. The solutions $x_i^s(k) = x_i(q_i(k))$ are signaled to all concerned links, which then adjust their link prices – for the next iteration $k + 1$ – according to the following rule

$$p_l^{(k+1)} = \left\{ p_l^{(k)} + \gamma \left[\sum_{i \in S(l)} x_i^s(k) - c_l \right] \right\}_+ \quad (16)$$

and is a positive step – chosen to allow the scheme to converge. Then the new link prices are signaled to links, etc., until the convergence is obtained.

5. Dynamic Routing

Although it would require a serious modification of Internet protocols to implement the described scheme of coordination by price instruments in a distributed manner, this algorithm can be instead used by the centralized controller, so called bandwidth broker. One can observe, that further improvement of the profit can be achieved by conscious choice of routing, in the systems with redundant paths existing between nodes. Here is an additional profit of using price instruments for network coordination. Equilibrium prices p_l , calculated in the process of network balancing can be interpreted as link congestion indicators. As such they can be used by external shortest-path routing algorithm.

There is a problem connected with this approach. The calculation of routes based on dynamic metrics leads to the change in the load pattern, and necessity of re-calculating allocations, and so on. In other words it may lead to oscillations, what was observed in experiments. To avoid this problem, time scales of routing and bandwidth allocation have to be separated. For example, rerouting can be performed every few minutes.

The following rerouting scheme is proposed:

Step 0: Sort the list of active sources in descending order w.r.t. x_i^{\max} .

Step 1: Choose the first element from the sorted list and find a shortest length route for it using Dijkstra algorithm, basing on dynamic metrics defined in Eq. (17); remove this source from the list.

Step 2: Calculate the new equilibrium prices pattern of the network solving the problem (15)–(16), update the metrics.

Step 3: Proceed until the route is found for every pair source-destination.

The critical issue is to propose an additive metric for the shortest length search. It must be a scalar value, bound to the link, which incorporates information about the path load and capacity at the same time. Finally, the proposed dynamic metric d_l includes the congestion prices, and takes the form:

$$d_l = p_l + \alpha \beta e^{-\frac{c_l}{c^{\max}}}, \quad \alpha \in \langle 0, 1 \rangle, \quad (17)$$

where c^{\max} is the highest link capacity in the network, $\beta = \max_{i \in S} x_i^{\max}$. Equation (17) takes into account, that at the initial iterations the network is “empty”, and none links are congested – thus $p_l = 0$. The exponential term reflects the fact, that capacity is a non-additive metric, that must be dealt with by shortest-path algorithm. In this case the routes with the highest capacities are occupied in the first place by the source-destination pairs with the highest transmission rates.

6. Results of Experiments

The simulation experiments were performed for a simple network configuration. Six sources were connected to the network presented in Fig. 1. All were assumed to trans-

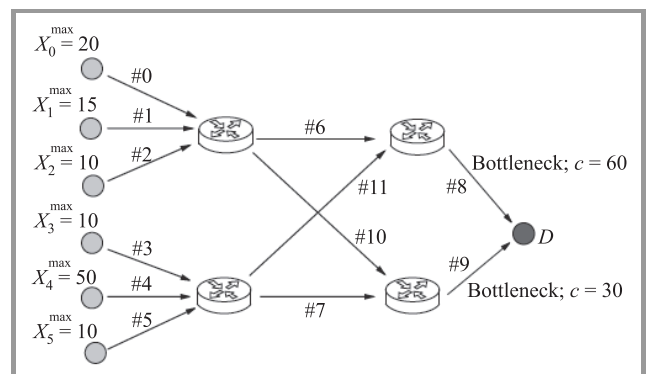


Fig. 1. Test network.

mit to the same destination, symbolized with a dark circle. The sources compete for the bandwidth of two bottlenecks (link 8 and link 9). Two algorithms for the bandwidth broker were compared, one that gives an optimal solution of the bandwidth allocation problem – mixed integer programming (MIP), and price-based scheme described in the previous section. The aggregate utilities were calculated.

The response of bandwidth broker is presented in Table 1. Both the MIP and price-based algorithm gave the same

result in – value of the utility function in both cases was 1659.4. This result was obtained under the assumption that all sources in the initial stage were sorted. For comparison in Table 1 (right-hand side) the performance of a modified price-based algorithm is presented, in which

Table 1

Links and rates, MIP, and price-based algorithm with sorting (left), price-based algorithm without sorting (right)

s	With sorting		Without sorting	
	link numbers	rate (x)	link' numbers	rate' (x)
0	6; 8	15.00	6; 8	19.00
1	9; 10	11.25	6; 8	14.00
2	9; 10	6.25	6; 8	9.00
3	7; 9	6.25	8; 11	9.00
4	8; 11	45.00	9; 7	30.00
5	7; 9	6.25	8; 11	9.00
$U(x)$	1659.4		1510.0	

only the Lagrange multipliers are used to represent link costs (metrics). No sorting was performed in the initial stage in this case. The obtained value of the utility function was 1510.0.

7. Final Remarks

In this paper the objective was to present various selected approaches to pricing, in particular concerned with the classical role of prices in balancing markets and systems in view of limited resources. The formulation and the main properties of the Price Method were reviewed. This method, developed with various modifications in the 1970-ties, still receives much attention. To illustrate this point possible application of the price-based mechanisms to Internet congestion control was discussed. In fact the size and the problems with operation of the Internet motivate the renewed interest in those mechanisms. The pricing “technology” is very much needed to control modern communications because there do not exist other optimization based technologies that could cope with the challenges offered by the Network.

Acknowledgments

The paper was partially supported by Polish Ministry of Science and Higher Education grant N N514 416934.

References

[1] L. A. DaSilva, “Pricing for QoS-enabled networks: a survey”, *IEEE Commun. Surv.*, 2000 [Online]. Available: <http://www.comsoc.org/pubs/surveys>

[2] M. Falkner, M. Devetsikiotis, and I. Lambadaris, “An overview of pricing concepts for broadband IP networks”, *IEEE Commun. Surv.*, 2000 [Online]. Available: <http://www.comsoc.org/pubs/surveys>

[3] G. Armitage, *Quality of Service in IP Networks: Foundations for a Multi-Service Internet*. Indianapolis: Macmillan, 2000.

[4] W. Findeisen, F. N. Bailey, M. Brdys, K. Malinowski, P. Tatjewski, and A. Wozniak, *Control and Coordination in Hierarchical Systems*. London: Wiley, 1980.

[5] L. S. Lasdon, *Optimization Theory for Large Systems*. New York: Macmillan, 1970.

[6] M. D. Mesarovic, D. Macko, and Y. Takahara, *Theory of Hierarchical, Multilevel, Systems*. New York: Academic Press, 1970.

[7] *Large-Scale Systems Control and Decision Making*, H. Tamura and T. Yoshikawa, Eds. New York: Marcel Dekker, 1990.

[8] K. Malinowski, “Applicability of the Lagrange multiplier method to optimization of quadratic systems”, *IEEE Trans. Autom. Contr.*, vol. AC-22, no. 4, pp. 648–651, 1977.

[9] X-R. Cao and H-X. Shen, “Internet pricing: comparison and examples”, in *39th IEEE Conf. Decis. Contr.*, Sydney, Australia, 2000.

[10] F. P. Kelly, A. K. Maulloo, and D. K. H. Tan, “Rate control for communication networks: shadow prices, proportional fairness and stability”, *J. Oper. Res. Soc.*, vol. 49, no. 3, pp. 237–252, 1998.

[11] R. J. La and V. Anantharam, “Charge-sensitive TCP and rate control in the Internet”, Rep., 2000 [Online]. Available: <http://www.eecs.berkeley.edu>

[12] S. H. Low and D. E. Lapsley, “Optimization flow control. I: Basic algorithm and convergence”, *IEEE/ACM Trans. Netw.*, vol. 7, no. 6, pp. 861–874, 1999.

[13] K. Malinowski, “Optimization network flow control and price coordination with feedback; proposal of a new distributed algorithm”, *Comput. Commun.*, vol. 25, pp. 1028–1036, 2002.

[14] I. C. Paschalidis and J. N. Tsitsiklis, “Congestion-dependent pricing of network services”, *IEEE/ACM Trans. Netw.*, vol. 22, no. 2, pp. 171–184, 2000.



Krzysztof Malinowski received his Ph.D. in 1974, D.Sc. in 1978 from the Warsaw University of Technology (WUT), Poland. He works as a Professor of control and information engineering at the same University. He is the Head of the Control and Systems Division. He is also a Professor at the Research and Academic Computer

Network (NASK) and a member of the Polish Academy of Sciences. He is the author or co-author of four books and over 130 journal and conference papers. For many years he was involved in research on hierarchical control and management methods. His current interests are control, simulation and optimization of communication networks. He was a visiting Professor at the University of Minnesota in 1976 and in 1978/1979, and then, from 1981 till 2001, served as a consultant to the Decision Technologies Group of UMIST in Manchester, UK. He was the Director of Institute of Control and Computation Engineering in 1984–1996 and Dean of the Faculty of Electronics and Information Technology in 1996–1999. From 2003 to 2009 he was the Director for Research of NASK. From 2000 to 2002 he was the project manager, on the Polish

side, of the QOSIPS project in the IST Programme of the FP 5. The QOSIPS project was concerned with development of the system consisting of the Quality Module and the Pricing Module, for tariff establishment and monitoring of telecommunication services of differentiated quality of service.

e-mail: K.Malinowski@ia.pw.edu.pl
 Institute of Control and Computation Engineering
 Warsaw University of Technology
 Nowowiejska st 15/19
 00-665 Warsaw, Poland
 e-mail: Krzysztof.Malinowski@nask.pl
 Research Academic Computer Network (NASK)
 Wąwozowa st 18
 02-796 Warsaw, Poland



Ewa Niewiadomska-Szynkiewicz received her Ph.D. in 1996, D.Sc. in 2006 from the Warsaw University of Technology. She works as a Professor of control and computation engineering at the Warsaw University of Technology. She is the Head of the Complex Systems Group. She is also an associate professor at the Research and

Academic Computer Network (NASK), and the Director for Research of NASK since 2009. She is the author or co-author of three books and over 90 journal and conference papers. Her research interests focus on complex systems modeling and control, computer simulation, global

optimization, parallel calculations and computer networks. She was involved in a number of research projects including three EU projects, coordinated the Groups activities, managed organization of a number of national-level and international conferences.

e-mail: ens@ia.pw.edu.pl
 Institute of Control and Computation Engineering
 Warsaw University of Technology
 Nowowiejska st 15/19
 00-665 Warsaw, Poland
 e-mail: ewan@nask.pl
 Research Academic Computer Network (NASK)
 Wąwozowa st 18
 02-796 Warsaw, Poland



Przemysław Jaskóła received his M.Sc. in computer science from the Warsaw University of Technology, Poland, in 1999. Currently he is a Ph.D. student in the Institute of Control and Computation Engineering at the Warsaw University of Technology. Since 2005 he works for Research and Academic Computer Network (NASK). His re-

search area focuses on hierarchical optimization and computer networks.

e-mail: Przemyslaw.Jaskola@nask.pl
 Research Academic Computer Network (NASK)
 Wąwozowa st 18
 02-796 Warsaw, Poland