Optimizing Proportionally Fair Prices

Laura Wynter

▶ **To cite this version:**

Laura Wynter. Optimizing Proportionally Fair Prices. [Research Report] RR-4311, INRIA. 2001. inria-00072276

HAL Id: inria-00072276
https://inria.hal.science/inria-00072276

Submitted on 23 May 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Optimizing Proportionally Fair Prices

Laura Wynter

N° 4311

October 2001

THÈME 4

ISSN 0249-6399 ISRN INRIA/RR--4311--FR+ENG

Optimizing Proportionally Fair Prices

Laura Wynter*

Thème 4 — Simulation et optimisation
de systèmes complexes
Projet METALAU

Rapport de recherche n° 4311 — October 2001 — 12 pages

Abstract: We demonstrate the non-uniqueness of Proportionally Fair prices, and show that the model of [1] falls into the category of models which can be generalized to allow optimization of the link prices, with a view toward profit maximization. However, through this framework, we also show that the total revenue that can be obtained through Proportional Fairness pricing on a network is unique.

Key-words: Internet pricing, Bandwidth allocation, Lagrange multipliers

* Laboratoire PRISM, Université de Versailles and INRIA, Rocquencourt

Unité de recherche INRIA Rocquencourt
Domaine de Voluceau, Rocquencourt, BP 105, 78153 Le Chesnay Cedex (France)
Téléphone : +33 1 39 63 55 11 — Télécopie : +33 1 39 63 53 30

L'optimisation de la Tarification des Télécommunications basée sur la Politique de *Proportional Fairness*

Résumé : On montre la non-unicité des prix obtenus par la politique d'allocation de ressources *"Proportional Fairness"* et comment généraliser le modèle de [1] afin d'optimiser les prix sur les arcs, avec une vue à la maximisation des profits. Cependant, on montre également que revenu total qui pourrait être obtenu par cette politique de tarification est unique.

Mots-clés : Tarification de l'Internet, Allocation de Resources, Equité, Multiplicateurs de Lagrange

1 Introduction

The concept of Proportional Fairness Pricing [1], was motivated by the desire to incorporate the notion of fairness into the allocation of network resources. In that scheme, one interpretation is that a resource allocation is fair if it is in proportion to the users willingness to pay. Under this interpretation, the model can be seen as one that generalizes standard resource allocation models by incorporating user utility functions into the objective of the resource allocation optimization problem. Prices for bandwidth allocation are then provided by certain Lagrange multipliers of the model.

Alternatively, another interpretation is that an allocation x^* is proportionally fair if it satisfies the following inequality

$$\sum_{s \in S} \frac{x_s - x_s^*}{x_s^*} \leq 0, \tag{1}$$

for all feasible allocations x_s, where s is a user, or demand for communication, and S the set of all such demands on the network[1]. That is, an increase in the allocation of network resources for one user must be compensated by corresponding decreases in the allocations of one or more other users.

Letting $F(x) = \nabla f(x) = \nabla \log(x)$, $x \in X \subset \Re^{|S|}$, so that $F_s(x) = 1/x_s$, we can rewrite (1) as follows:

$$F(x^*)^T(x^* - x) \geq 0, \quad x \in X, \tag{2}$$

where X is the set of all feasible resource allocations, x. The expression (2) is then a standard *variational inequality*, which generalizes the class of optimization problems. Note then that to add a route-dependent weight to the inequality (1), as in [1], one needs only to define the function $f_s(x) = w_s log(x_s)$; more complex forms are then clearly possible as well.

When $F : |Re^{|X|} \mapsto \Re^{|X|}$ can be expressed as the gradient of some function $f : \Re^n \mapsto \Re$, then the variational inequality (2) is precisely the first order optimality condition of the following optimization problem[2]:

$$\max_{x \in X} f(x). \tag{3}$$

Mathematical program (3) then gives the utility form of the problem, when $f(x) = \log(x)$, or $f(x) = w \log(x)$, as defined above. In the case of Proportional Fairness Pricing, these two interpretations are therefore equivalent, as can be seen through the use of the variational inequality. The first interpretation is that of a network flow optimization problem, with a logarithmic objective (utility) function; the variational interpretation is precisely the optimality condition to that optimization problem. The variational inequality formulation allows, furthermore, the use of much more general utility functions, including differentiated service, asymmetric response functions, etc., as well as deeper insights into the mathematical properties and different solution methods for the problem.

[1] Note that the definitions of x and x^* are opposite to those of [1]; we maintain the standard custom in the optimization literature to refer to x^* as the optimal solution over all feasible x.

[2] This equivalence is notably not present when the vector function F has an asymmetric jacobian matrix. In this case, there is no equivalent convex (or concave) optimization formulation for the variational inequality.

Proportionally Fair prices posses certain desirable traits, including the fact that the prices are zero when the network is uncongested, and thereafter are proportional to the degree of oversaturation of the network. This framework then provides some intrinsic congestion control, should the prices be applied in practice. That was the premise of the work of [1]. Similar approaches can be found in Low and Lapsley [2], and many others since.

Existence of these prices was shown by the Kelly in [3], but their uniqueness has up to now remained an open question. We shall demonstrate when the prices are unique and not, and further, in the latter case, describe a strategy for optimizing a secondary criterion over the set of Proportionally Fair prices.

We begin by summarizing this model and necessary notation for our developments of it. Let $G(N, A)$ denote a strongly connected network, where N is the set of nodes, of which there are n, and A the set of links, a, of which there are m. Each link has associated with it a given capacity $u_a \in R_+ \cup \{\infty\}$, $a \in A$.

For certain node pairs, $s = (n_1, n_2) \in S \subset N \times N$, there is a demand for communication, known as d_s. The flow associated with the set of all such demands over the network can be expressed as a flow on each link a, or on each origin-to-destination route. We shall assume here that several routes serve the same origin-to-destination (od) pair. The set of all routes on the network will be referred to as R. This may represent the case of several physical routes that a request may take, or the presence of more than one network operator serving that pair.

Indicator matrices can be used to convert from the flow on links to flow on routes. Let Δ be the link-route indicator matrix, where $\Delta_{ar} = 1$ if link a is present along on route $r \in R$, and 0 otherwise. Let Γ be the od-route indicator matrix, where $\Gamma_{rs} = 1$ if route r serves the od-pair $s \in S$, and 0 otherwise.

In this paper, we shall suppose that the users are homogeneous in their quality of service requirements. Adding user differentiation within each node pair is straightforward, but complicates the notation.

The Proportional Fairness Pricing model associates with each od-pair a user utility function that takes as decision variable the amount of resource allocated to that user over his entire path. The utility function for each user s, $U_s(d_s)$, is assumed to be increasing, strictly concave, additive, and continuously differentiable with respect to d_s. Denoting $d = (d_s, s \in S)$, and $U(d) = \sum_{s \in S} U_s(d_s)$, the concave maximization problem of determining a Proportionally Fair resource allocation is given below:

$$\max \quad \sum_{s \in S} U_s(d_s) \tag{4}$$

subject to

$$\Gamma y = d \tag{5}$$

$$\Delta y \leq u \tag{6}$$

$$d, y \geq 0. \tag{7}$$

The variable $y \in \Re^{|R|}$ is the vector of route flows. The constraint (5) indicates that the aggregate flow over the routes serving each od-pair must be equal to the demand of that pair. The constraints (6) state that the aggregate flow on each link must be less or equal to the capacity of the link.

Since the objective function is assumed strictly concave in d, the amount of resource allocated to each user, that is to each od-pair, is unique. In [1] the use of a Lagrangian method was proposed to solve (4)–(7), in which the capacity constraints (6) are relaxed. Then the authors interpreted the Lagrangian multipliers, λ_a, $a \in A$ the *Proportionally Fair* price for a unit flow through link a; in other words, λ_a is the shadow price of additional capacity at resource a.

The authors further decomposed (4)–(7) into a coupled pair of problems: one for the users, who seek to maximize their individual utilities without regard to the capacity constraints, and the other for the network manager, who seeks to allocate resources so as to maximize revenue, and satisfy capacity restrictions. The decision variable in both sets of problems is the same – that of the amount of resource to allocate to each user, d_s. The link between the two problems is the price paid by each user λ_s, which is the sum of the link prices λ_a over all links on the paths of all routes used by the user s, that is,

$$\lambda_s = \sum_{a \in A} \sum_{r \in R} \Delta_{ar} \Gamma_{rs} \lambda_a.$$

The pair of linked problems is:

$$\max_d \quad \textstyle\sum_{s \in S} U_s(d_s) - \lambda_s d_s \tag{8}$$

subject to

$$d_s \geq 0, \tag{9}$$

and

$$\max_d \sum_{s \in S} \lambda_s d_s \tag{10}$$

subject to

$$\Gamma y = d, \tag{11}$$

$$\Delta y \leq u, \tag{12}$$

$$d, y \geq 0. \tag{13}$$

The network problem attempts to choose the resource allocation so as to maximize revenue *as a function of flow only.*

Since the user problem (8)–(9) has no constraints linking the different users, and the objective function is additive, it can be decomposed and a solution obtained individually for each user, s; that is, one can solve the individual user problems separately:

$$\max_{d_s} \quad U_s(d_s) - \lambda_s d_s \tag{14}$$

subject to

$$d_s \quad \geq \quad 0. \tag{15}$$

It was shown by Kelly in [3] that *there exists a multiplier vector, or shadow price, λ such that the solution to the two problems, (8)–(9) and (10)–(13) is the same*, and further is the same as that which one would obtain from solving (4)–(7).

2 A restricted bilevel programming problem

The above model from [1] is of significant interest for the allocation of telecommunication resources, as it provides an economic foundation to the resulting allocation, when a utility function is used to describe user preferences in the place of the traditional (linear) cost minimization objective.

It further provides a starting point for modeling link pricing and its relation to the demand for resource usage on a telecommunications network. That is, it combines marketing ideas of pricing with technical preoccupations of flow levels on the network.

However, this model of Proportionally Fair prices and resource allocation does not go as far as permitting the network manager to treat pricing as a goal in itself. In the above models, the decision variables are all resource allocation levels, and prices are outputs of the algorithmic strategy.

In this section, we recall the framework proposed by Larsson and Patriksson in [4] that will allow us to extend and generalize the model above to one which explicitly considers the network manager's objective of maximizing revenue.

In particular, in this section we provide a formulation of the above problems as a single saddle-point problem, which is an important special case of the general bilevel programming problem. It permits optimizing the network manager's objective function over a restricted set of prices, and, further, draws a parallel between the method of marginal-cost pricing, and operator's profit maximization.

In addition, this approach has the advantage of providing a model having a unique optimal solution, and is therefore computationally very easy to solve, as opposed to its more general bilevel programming counterpart.

The method works as follows. Consider a canonical network flow optimization problem with capacity constraints:

$$\min \quad f(x) \tag{16}$$

subject to

$$\Gamma y = d \quad (17)$$
$$\Delta y \leq u \quad (18)$$
$$d, y \geq 0, \quad (19)$$

where the flow on a link, a, is $x_a = \sum_{r \in R} y_r \Delta_{ar}$.

Then, the linear system obtained from the Karush-Kuhn-Tucker (KKT) optimality conditions for the problem (16)–(19) can be expressed after some manipulation, as was done in Larsson and Patriksson [4], as follows:

$$\Delta^T [\nabla f(x^*) + \lambda] - \Gamma \pi \geq 0, \quad (20)$$
$$[\nabla f(x^*) + \lambda]^T x^* - d^T \pi = 0, \quad (21)$$
$$\lambda^T (x^* - u) = 0, \quad (22)$$
$$\lambda \geq 0, \quad (23)$$

where π is the Lagrange multiplier vector for constraints (17), and λ is the Lagrange multiplier vector for constraints (18), as before. Equation (21) can be obtained by writing the following linear programming formulation of (16)–(19), solving for x^*:

$$\min_x \quad \nabla f(x^*)^T x \quad (24)$$

subject to constraints (17)–(19). The function $\nabla f(x^*)$ is the value of the gradient of the function f *evaluated at the optimal value*, x^*. While model (24) is clearly not solvable in practice, since the function value $f(x^*)$ at x^* is not known, it permits writing a linear programming dual to the original nonlinear problem (16)–(19). Taking the Lagrangian relaxation of the capacity constraints (18) and adding that into the linear programming objective function (24), we obtain

$$\begin{array}{ll} \min_x & \nabla f(x^*)^T x + \lambda^T (x - u) = \\ \min_x & \nabla f(x^*)^T x + \lambda^T x. \end{array} \quad (25)$$

Then, the linear programming dual of (25) with constraints (17) and (19) is:

$$\max_\pi \quad d^T \pi \quad (26)$$

subject to

$$\Gamma^T \pi \leq \nabla f(x^*) + \lambda. \quad (27)$$

Strong (linear programming) duality gives the term analogous to that of equation (21), that is:

$$[\nabla f(x^*) + \lambda]^T x^* - d^T \pi = 0.$$

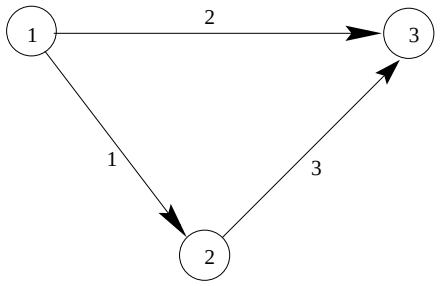

Figure 1: A network with 3 nodes and 3 links.

For a fixed, optimal, value of x^*, the system (20)–(23) describes a polyhedron in the Lagrange multiplier vectors λ and π.

Along with the system of (20)–(23), the authors in [4] observed that the multiplier vector λ will not necessarily be unique in network flow problems.

Theorem 1 *[Karush-Kuhn-Tucker Necessary Conditions][5] Let x^* be a local minimum of the problem* $\min f(x)$ *subject to $h(x) = 0$ and $g(x) \leq 0$, where f, h_i, and g_j, are continuously differentiable functions from $\Re^m \mapsto \Re$, for each $i \in I_1$, $j \in I_2$. Let $A(x)$ be the set of active inequality constraints at x. If the equality constraint gradients $\nabla h_i(x)$, $i = 1, \ldots I_1$ and the active inequality constraint gradients $\nabla g_j(x)$, $j \in A_{I_2}(x)$, are linearly independent (i.e., the point x is* regular*), then the Lagrange multiplier vectors for constraints $h(x) = 0$ and $g(x) \leq 0$ are unique.*

Indeed, the requirement of linear independence of binding constraint gradients is not usually satisfied in networks, as the constraint gradients take the form of indicator matrices for the graph, and in particular the arc-route indicator matrix, which are not constructed so as to have linearly independent rows.

The following example illustrates, for simple network structures, cases where the linear independence, and therefore the uniqueness of the multiplier vectors (shadow prices), holds and does not hold.

Example 1 *Consider the network of Figure 1. The network has 3 nodes, 3 links, and one origin-destination pair, from node 1 to node 3. There are 2 routes joining the origin-*

destination pair. The link-route 0-1 incidence matrix, Δ, is given below.

$$\Delta = \begin{vmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{vmatrix}$$

and the route-od pair incidence matrix, Γ, is given by the following:

$$\Gamma = \begin{vmatrix} 1 & 1 \end{vmatrix}$$

where Γ has only one row since there is only one od-pair in the network. The combined jacobian matrix of the constraints, (5)–(6), M, is then

$$M = \begin{vmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{vmatrix}$$

The Theorem 1 states that when the rows of the matrix, M, corresponding to active constraints at an optimal solution are extracted into a submatrix, and that those rows are linearly independent, then the multiplier vectors will be unique.

Noting that the constraint (5) is always active, since it is an equality constraint, and therefore that the fourth row of M is always present, it is easy to see that the rows of M will be linearly independent only when at least two capacity constraints are not active. In other words, at least two links must have flows strictly less than their capacity. In this case, the multipliers are naturally zero on the non-capacitated links. Therefore, as soon as any two links reach capacity, the matrix given by M will have linearly dependent rows, and the multipliers will not be unique. Consequently, the Proportionally Fair prices will not be unique.

While the uniqueness of the multiplier or price vector is clearly network dependent, we observe that even for very small and simple network structures, the Proportionally Fair Price vector will not be unique.

When these multipliers are not unique, then, Larsson and Patriksson [4] proposed optimizing a secondary objective over the set of those multiplier values. That is, given the system (20)–(23), they posed the following secondary optimization problem:

$$\min \quad \phi(\lambda) \tag{28}$$

subject to

$$\lambda \in T \cap P, \tag{29}$$

$$\tag{30}$$

where T is the set of λ given by (20)–(23), and $\phi : \Re^{|A|} \mapsto \Re$ is coercive, and preferably convex. Since the feasible set is polyhedral, when P is, this secondary, or network manager's optimization, problem is easy to solve using standard linear or nonlinear programming algorithms, depending on whether ϕ is a nonlinear or linear function of λ.

The network manager's objective can take various forms. Larsson and Patriksson [4], in the context of transport planning, discuss at length the network objective of trying to achieve a particular flow (of transport users) by adding prices judiciously; this includes trying to adjust link flows from their user-equilibrium values to a system optimum (using marginal cost pricing) or to some other flow (which then changes the prices that are produced by the model), or even allowing prices to be positive or negative, in which case, travel subsidies are considered (negative prices).

3 Profit maximization with proportionally fair pricing

In our context, we are primarily concerned with a telecommunications service provider, rather than a public authority, and so prices should be determined so as to maximize revenue.

Returning now to the Proportional Fairness pricing model [1] of (4)–(7), we shall express the corresponding polyhedral system of multipliers, or shadow prices, of the capacity constraints, which have been referred to as Proportionally Fair prices. While an existence result for these prices is immediate, and was included [3], the uniqueness of these prices has been an open issue. The issue is of very clear importance if the prices are to be used. Indeed, in the case of non-unique prices, depending on the algorithm used, a different price (multiplier) vector would be obtained.

Within the context of this paper, the interest is in obtaining a *particular* price vector, that is, one which satisfies an additional objective. Therefore, this question is first addressed for the model of (4)–(7).

Remark 1 *Recall that, according to Theorem 1, proving the uniqueness of the prices requires a study of the network topology at the optimal flow solution. If the active constraint gradients evaluated at that solution are linearly independent, then the multiplier vectors are known to be unique by the Karush-Kuhn-Tucker necessary optimality conditions. The active constraint gradients in this model are a subset of the route-arc indicator matrix, Δ, and the entire route-od indicator matrix, Γ. Depending upon which capacity constraints are active, the subset of Δ and the matrix Γ may or may not have linearly independent rows.*

Remark 2 (Pareto optimal solutions) *In order to satisfy the Pareto optimality of the optimal resource allocation, efficient links would likely not remain under-utilized with respect to their capacity limits. Therefore, a consequence of having many active capacity constraints is that Pareto optimal solutions would admit, in general, non-unique Proportionally Fair prices.*

We can study the system defining the multiplier vectors that arises from looking at the KKT conditions and the statement of strong duality of the corresponding linear program.

The model of (4)–(7) is different from that of (16)–(19) as the decision variable is the total demand between each pair of nodes, rather than the link flow.

To do so, we begin by formulating the linear programming form of the Proportional Fairness pricing problem. We obtain:

$$\max_{d,y} \quad \nabla U(d^*)^T d \tag{31}$$

subject to

$$\begin{aligned} \Gamma y &= d && (32)\\ \Delta y &\leq u && (33)\\ d, y &\geq 0. && (34) \end{aligned}$$

The linear programming dual to (31)–(34) can then be expressed as follows:

$$\min_{\lambda,\pi} \quad u^T \lambda \tag{35}$$

subject to

$$\begin{aligned} \Gamma^T \pi + \Delta^T \lambda &\geq 0 && (36)\\ d^T \pi &\geq \nabla U(d^*) && (37)\\ \lambda &\geq 0. && (38) \end{aligned}$$

By strong duality, we have that

$$\nabla U(d^*) = u^T \lambda. \tag{39}$$

Equation (39) says that the quantity $u^T\lambda$ is equal to a constant value, $\nabla U(d^*)$, and the term $u^T\lambda$ gives the revenue accrued from the network flow, when link prices λ_a, $a \in A$ are charged. Indeed, λ will be identically zero when the constraint (6) is not active, i.e. when capacity is not reached. For that link, the component of right-hand side of (39) makes no contribution. On the other hand, the price λ is collected precisely when capacity is reached. on each link, that is when $\Delta_{ar} y_r^* = u_a$, $a \in A$. Consequently, $u^T\lambda$ gives the total revenue obtained from the flow y^*. This is summarized in the following proposition.

Proposition 1 *The total revenue obtained by charging Proportional Fairness prices, that is* $\sum_{a\in A} \lambda_a x_a$, *where* $x_a = \Delta_{ar} y_r$ *is the total flow on link* $a \in A$, *is unique.*

Proof.

By strong duality on the linear programming formulation of (4)–(7), we have that (39) holds. Since $u^T\lambda$ represents the revenue associated with the optimal flow y^*, and $\nabla U(d^*)$ is constant for fixed d^*, the desired result follows.

Remark 3 *Although the total revenue that can be obtained through the Proportion Fairness pricing model is unique, the prices themselves are not necessarily unique, as discussed in Remark 1. Indeed, there are an infinite number of ways that the same total revenue* $\sum_{a \in A} \lambda_a x_a$, *can be obtained by distributing prices* λ_a *among the links* $a \in A$ *of the network, all giving the same optimal link flow* x *and user resource level* d. *This result also arose in the elastic-demand formulation of traffic equilibrium in [4].*

4 Conclusions

We have analyzed the Proportional Fairness Pricing model and presented it within a more general optimization framework, in which the prices, viewed as Lagrange multipliers on a convex (or concave) optimization problem, can be optimized. The set of those prices, or Lagrange multipliers, was expressed explicitly, using a linear-programming dual formulation of the nonlinear programming problem.

The uniqueness of the prices was discussed and it was demonstrated through a very simple example that the prices are not likely to be unique. The problem which allows those prices to be optimized while maintaining their characterization as "Proportionally Fair" was provided, and is in fact, not much more complex to solve than the original problem.

Further research along these lines could take the form of more general utility functions, including differentiated services with asymmetric response functions.

References

[1] F. Kelly, A. K. Maulloo, and D. K. H. Tan, "Rate Control for Communication Networks: Shadow Prices, Proportional Fairness, and Stability," *Journal of the Operational Research Society* 49, 1998, available from URL http://www.statslab.cam.ac.uk/~frank/rate.html, pp. 23752.

[2] S.H. Low and D.E. Lapsley, "Optimization flow control I: Basic algorithm and convergence," *IEEE/ACM Transactions on Networking*, vol. 7, pp.861–874, December, 1999.

[3] F.P. Kelly, "Charging and Rate Control for Elastic Traffic", *European Transactions in Telecommunications*, 8, 1997, pp.33-37.

[4] T. Larsson and M. Patriksson, "Side Constrained Traffic Equilibrium Models-Traffic Management Through Link Tolls," Chapter 7 in *Equilibrium and Advanced Transportation Modelling*, P. Marcotte and S. Nguyen (eds.), Kluwer Academic Publishers, Boston, MA, pp. 125-151, 1998.

[5] D. Bertsekas, *Nonlinear Programming*, Athena Scientific, Belmont, Mass., USA, 1999.

Unité de recherche INRIA Rocquencourt
Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38330 Montbonnot-St-Martin (France)
Unité de recherche INRIA Sophia Antipolis : 2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399