43rd IEEE Conference on Decision and Control
December 14-17, 2004
Atlantis, Paradise Island, Bahamas

Satellite Dynamic Bandwidth Allocation by Neural Open Loop-Feedback Control Strategies

Marco Baglietto*, Franco Davoli*, Mario Marchese°, and Maurizio Mongelli*
* DIST - Department of Communications, Computer and Systems Science
° CNIT - Italian National Consortium for Telecommunications

University of Genoa, Via Opera Pia 13, 16145 Genova, Italy
Email*: {mbaglietto,franco,mopa}@dist.unige.it, Email°: mario.marchese@cnit.it

Abstract - **In this paper, a novel optimization methodology is investigated to optimize the resource allocation in a satellite system where variations of fading conditions are added to those of traffic load. A neural approximation technique is studied to exploit on line optimal reallocation laws, as functions of the state of the network. No closed-form expressions for the system dynamic equations and the functional cost are needed. Simulation results show how the proposed technique outperms two other optimization strategies, based on a *certainty equivalent* assumption and on the application of *Perturbation Analysis*, respectively.**

I. Introduction

OPTIMIZATION techniques for telecommunication networks usually follow the minimization of a proper functional cost as outlined in (1):

$$\boldsymbol{\theta}^* = \arg\min_{\boldsymbol{\theta}\in\Theta} \underset{\boldsymbol{\omega}}{E}\left\{ L[\boldsymbol{\theta},\boldsymbol{\omega}] \right\}. \qquad (1)$$

We denote by $L[\boldsymbol{\theta},\boldsymbol{\omega}]$ the performance index of interest (e.g., blocking probability of connection requests, packet loss probability, packets' mean delay or delay jitter) and by $\boldsymbol{\theta}$ the vector of the decision variables. The expectation $\underset{\boldsymbol{\omega}}{E}[\cdot]$ is over all the feasible sample paths $\boldsymbol{\omega}\in\Omega$ of the system. When the domain of the decision variables is discrete, (1) expresses a *NP-hard Stochastic Discrete Resource Allocation Problem* (SDRAP), whose optimal solution can be found only through a strong consumption of computing power. When closed-form expressions of $L[\boldsymbol{\theta},\boldsymbol{\omega}]$ are available, a *parameter adaptive certainty equivalent control* can be applied. To do this, it is often necessary to assume the *Markovian* hypothesis concerning the stochastic environment. The major drawbacks of such approach stem from the related "*curse of dimensionality*" problem (when *Dynamic Programming* is applied) and from the limited availability of closed-form expressions when real life contexts are taken into account.

In this perspective, *Perturbation Analysis* (PA) techniques [1, 2] can be used to derive sensitivity estimators for the performance metrics under investigation, thus providing on-line gradient descent algorithms capable to optimally distribute the available resources among the users (see, e.g., [3, 4]). By applying a "*surrogate*" relaxation of the discrete functional cost and without exploiting any closed-from expression of $L[\boldsymbol{\theta},\boldsymbol{\omega}]$, PA obtains reliable shorter-term gradient estimators, only by looking at the current realization of the stochastic variables involved in the problem. The forms of such estimators are obtained on real data, by tracking the behaviour of the network and seeking to continuously improve its performance.

Unfortunately, PA-based techniques do not have any guarantee concerning the time needed to reach the optimal resource allocation.

Let $\gamma(\cdot)$ be the "instantaneous" performance index of interest (i.e., $L[\boldsymbol{\theta}(t),\boldsymbol{\omega}(t)] = \int_0^t \gamma[\boldsymbol{\theta}(\tau),\boldsymbol{\omega}(\tau)]\, d\tau$). When the stochastic environment is not stationary and the problem is explicitly formulated over time as a *Dynamic Stochastic Discrete Resource Allocation Problem* (D-SDRAP):

$$\underset{\boldsymbol{\omega}}{E} \lim_{T\to\infty} \frac{1}{T} \int_t^{t+T} \gamma[\boldsymbol{\theta}(t),\boldsymbol{\omega}(\tau)]\, d\tau . \qquad (2)$$

the decision variables, driven by the PA estimator, need some transient periods to "learn" optimality [4]. Therefore, the study of control techniques capable to track the sensitivity estimation, together with overcoming such suboptimal transient periods reveals to be an hot topic of research.

In this perspective, we study in this paper a novel control methodology, suitable for the resource allocation in a satellite network. We face the related D-SDRAP by exploiting a proper functional optimization problem. Following the principles of PA, we do not explicit any closed-form expression for the system dynamic equations and for the functional cost.

To this aim, we properly modify the control methodology of [5], in which a neural approximating technique is studied to tackle the curse of dimensionality problem arising for the structure of the decision functions as the dimension of state vector increases. The key idea is to adopt PA to train the neural decision functions off line, thus providing on line optimal reactions to variable system conditions and with a small computational effort.

The remainder of the paper is organized as follows. In the next Section we summarize the network environment under investigation and formulate the D-SDRAP we deal with. In Section III, we detail a possible certainty equivalent approach and, in Section IV, we propose a control algorithm based on PA. In Section V, we introduce the functional optimization problem underlying the D-SDRAP under investigation and formulate our neural approximating technique. Simulation results are presented in Section VI to compare the proposed optimization approaches and, in Section V, we finally conclude by summarizing the obtained results and emphasizing directions for future research.

II. D-SDRAP IN A SATELLITE ENVIRONMENT

A. The Model of the Satellite Network

In satellite networks, variable fading conditions over the channel can heavily affect the transmission quality, especially when working in *Ka band*, where the effect of rain over the quality of transmission is more significant [6].

The satellite environment under investigation consists of a fully meshed satellite network that uses bent-pipe geostationary satellite channels, joining N traffic stations. With a notation that slightly differs from [2], each station i is assumed to have a buffer of fixed size Q_i with a single server of capacity θ_i. The stochastic processes associated with this model and useful in our optimization framework are: $\alpha_i(t)$ (the *input flow*), and $\gamma_i(\theta_i, t)$ (the *loss rate* process due to a full buffer).

B. The traffic model

We shall adopt a specific traffic model in order to specify a closed-form functional cost of the performance index. Hence, we introduce the model adopted for each station's input flow process $\alpha_i(t),\ i = 1,...,N$.

We suppose that each $\alpha_i(t)$ is composed by a self-similar stochastic process, due to the aggregation of some variable bit rate sources. For each station i, the statistical parameters that describe such process are: the peak bit rate B_p^i of the on-off sources and the burst arrival rate $\lambda_{burst}^i = M^i \big/ (\bar{\tau}^i + \bar{\varsigma}^i)$ of the aggregated flow, namely, the average number of *bursts* "seen" by station i. We denote by $\bar{\tau}^i$ and $\bar{\varsigma}^i$ the mean time durations of the burst (i.e., the period of activity of a source) and the silence periods, respectively, and are both *Pareto* distributed in order to achieve a self similar behaviour of the aggregated flow [7], M^i is the maximum number of on-off sources in the flow $\alpha_i(t)$.

C. The performance measure

The performance measure of interest is the *loss volume* $L(\cdot)$. For each station i it is given by:

$$L_i(\theta_i^d) = \lim_{T\to\infty} \frac{1}{T} \int_0^T \gamma_i(\theta_i^d(\tau), \alpha_i(\tau))\, d\tau,\ i = 1,...,N\ . \quad (3)$$

Where $\boldsymbol{\theta}^d(t) = \left[\theta_1^d(t),...,\theta_N^d(t)\right]$ is the vector of the service capacities allocated to each station at time t. From here on, we shall use the superscript d to stress that a quantity belongs to a "discrete" set. As to $\boldsymbol{\theta}^d(t)$, we have $\boldsymbol{\theta}^d(t) \in \Theta_d$:

$$\Theta_d = \left\{ \boldsymbol{\theta}^d(t) \in \mathbb{N}^N : \theta_i^d(t) = h_i(t)\cdot MAU, h_i \in \mathbb{N},\ \sum_{i=1}^{N} \theta_i^d(t) = K \right\} \quad (4)$$

namely, the allocated service rate for each station is a discrete number of "*Minimum Allocation Units*" (MAUs), i.e., the smallest portion of bandwidth that can be assigned to a station. K is the total service capacity available for the satellite system. The total loss volume of the overall system becomes the sum of the contributions of each station:

$$L(\theta_1^d,...,\theta_N^d) = \lim_{T\to\infty} \frac{1}{T} \sum_{i=1}^{N} \int_0^T \gamma_i(\theta_i^d(\tau), \alpha_i(\tau))\, d\tau\ . \quad (5)$$

D. The fading effect and problem formulation

The effect of fading is modeled as a reduction in the bandwidth actually "seen" by a traffic station. The fading effect is represented by a variable $\phi_i(t)$, which shows how the bandwidth $\theta_i^d(t)$ is reduced. For each station i, at time t, the "real" $\theta_i(t)$ is:

$$\theta_i(t) = \phi_i(t) \cdot \theta_i^d(t);\ \ \phi_i(t) \in [0,1];\ i = 1,...,N\ . \quad (6)$$

The reduction of the bandwidth actually seen by the station is due to the increase in the bandwidth required to maintain a fixed *Bit Error Rate* (BER) through *Forward Error Correction* (FEC) codes (and, possibly, transmission bit rate reduction). Whenever the fading effect lowers the signal strength, an adaptive control (located at the *Physical Layer*) monitors the *Carrier/Noise Power* factor and, on the basis of this measure, increases the redundancy of the packets sent, introduced by the FEC and bit rate adaptation [6]. In this way, since more bits are necessary to transmit a single packet, and the bit duration may be increased, the information bit rate is reduced, as if a smaller bandwidth would be seen by the station.

The optimization problem can now be stated.

Problem D-SDRAP (*Dynamic-Stochastic Discrete Resource Allocation Problem*): find $\boldsymbol{\theta}^{d*}(t) \in \Theta^d,\ t \ge 0$ in a such a way that the cost function:

$$\underset{\omega_1,...,\omega_N}{E} \left[\lim_{T\to\infty} \frac{1}{T} \sum_{i=1}^{N} \int_0^T \gamma_i(\theta_i^d(\tau), \phi_i(\tau), \alpha_i(\tau))\, d\tau \right] \quad (7)$$

is minimized. We denote by ω_i the generic sample path for station i, namely, a realization of the stochastic processes that characterize the temporal evolution of station i: $\phi_i(t),\ \alpha_i(t),\ i=1,...,N$.

III. A CERTAINTY EQUIVALENT APPROACH

Also in the presence of a self-similar behaviour of the traffic sources, it is possible to obtain analytical models for the computation of the loss probability performance. We now describe in some detail such optimization strategy. Following [6], we adopt the *Tsybakov-Georganas* formula [7] for the cell loss probability $PLoss_i$ of each station i:

$$PLoss_i(\theta_i^d)=\begin{cases}\dfrac{\chi\cdot\lambda_i\cdot R^a}{a\cdot(a-1)\cdot(X_i-\lambda_i\cdot R\cdot\overline{\tau})}\cdot(Q_i)^{-a+1} & \text{if } X_i>\lambda_i R\overline{\tau}\\ 1 & \text{otherwise}\end{cases} \quad (8)$$

where $\hat{T}$ is a reference time interval (slot), and L is the number of bits in a cell. Then, $R=\left\lceil\dfrac{\hat{T}\cdot B_p}{L}\right\rceil$ is the number of cells generated by an active burst in a slot and $X_i=\left\lfloor\dfrac{\theta_i^d\cdot\phi_i}{L}\cdot\hat{T}\right\rfloor$ represents the bandwidth θ_i^d, assigned to station i (and degraded according to the current value of fading ϕ_i), expressed in cells per slot (disregarding cell overhead). Moreover, $\lambda_i=\lambda_{burst}^i\hat{T}$. For simplicity, we have assumed $B_p^i=B_p,\forall i$.

The D-SDRAP (7) can be now slightly modified by taking into account (8), thus stating the problem of the minimization of the overall loss probability at each time instant where a new bandwidth reallocation is performed:

$$\boldsymbol{\theta}^{d*}(k)=\arg\min_{\boldsymbol{\theta}^d(k)\in\Theta^d}\sum_{i=1}^{N}PLoss_i(\theta_i^d(k));\quad k=1,2,... \quad (9)$$

The index k denotes the reallocation time instants at which a new solution of (9) is computed, according to the current state of the network (in terms of traffic load and fading levels). The minimization of (9) can be performed through dynamic programming, whose computational burden limits its applicability in real time.

In the following, we shall denote the application of this technique as CF&DP (*Closed-form* and *Dynamic Programming* approach).

IV. THE ON-LINE SURROGATE ALGORITHM

In this section, we summarize the on-line "surrogate" optimization technique of [3] and detail its application to our problem. By means of the gradient of the functional cost $\dfrac{\partial L(\boldsymbol{\theta})}{\partial\theta_i}$, $i=1,...,N$, it is possible to obtain an on-line optimization descent, in order to optimally distribute the available channel capacity among the stations [4]. As regards the computation of such a gradient, the reader is referred to [2], where a PA technique is investigated in order to compute performance derivative estimators for a traffic buffer, as function of the sample paths ω_i, $i=1,...,N$ of the system. Once the discrete constraint set Θ_d is "relaxed" into a continuous one Θ_c:

$$\boldsymbol{\theta}^c\in\Theta_c,\ \Theta_c=\left\{\theta_i^c\in\Re^+;\sum_{i=1}^{N}\theta_i^c=K\right\}. \quad (10)$$

and letting $\hat{t}$ be the delay latency of the satellite system, each station i, for every $t=k\hat{t},\ k=1,\ 2,...$, must:

1) observe *the buffer temporal evolution during the time interval* $[(k-1)\hat{t},k\hat{t}]$ *according to the current sample path* ω_i *and bandwidth allocation* $\theta_i^d[(k-1)\hat{t}]$, $\boldsymbol{\theta}^d[(k-1)\hat{t}]\in\Theta_d$;

2) compute *the gradient estimation* $\dfrac{\partial L(\boldsymbol{\theta}[(k-1)\hat{t}])}{\partial\theta_i}$;

3) adjust *the value of its "bandwidth allocation need" using the gradient method:*

$$\theta_i^c[k\hat{t}]=\theta_i^d[(k-1)\hat{t}]-\eta\ \frac{\partial L(\boldsymbol{\theta}[(k-1)\hat{t}])}{\partial\theta_i};$$

4) communicate *such* $\theta_i^c[k\hat{t}]$ *to each master station;*

(for each station that has the role of master station) by looking at the information received by the other stations (i.e., $\theta_\kappa^c[(k-1)\hat{t}]$, $\kappa=1,...,N;\ \kappa\neq i$*),* ***convert*** $\boldsymbol{\theta}^c[(k-1)\hat{t}]$ *to the nearest discrete feasible neighbor* $\boldsymbol{\theta}^d[k\hat{t}]\in\Theta_d$.

Following [3], the nearest feasible neighbor $\boldsymbol{\theta}^d[k\hat{t}]\in\Theta_d$ of $\boldsymbol{\theta}^c[(k-1)\hat{t}]$ can be determined, at step 4, by an $O(N+1)$ algorithm based on the $N+1$ discrete neighbors of $\boldsymbol{\theta}^c[(k-1)\hat{t}]$, not necessarily all feasible, and on the selection of one of them, which satisfies the discrete constraint set Θ_d.

In the following, we shall denote the application of this technique as SE&GD (*Sensitivity Estimation* and *Gradient Descent* optimization approach).

V. A FUNCTIONAL OPTIMIZATION APPROACH

As we shall show in the simulation results, the SE&GD technique can be successfully applied on line if the gradient stepsize η is properly refined by means of an off-line performance evaluation. However, the investigation necessary to evaluate the best value for the gradient stepsize η requires a deep simulation inspection [4]. On the other hand, if the behaviour of the stochastic processes does not allow the decision variables to reach a steady state, it is difficult to verify if the on-line surrogate technique guarantees the best performance. In order to avoid these drawbacks and since it is impossible to solve (7) analytically, we now study a closed-loop control strategy, based on a functional optimization approach. The idea is to provide resource reallocations as functions of a suitable "information vector" $\boldsymbol{I}(t)$ which summarizes the "recent history" of the system. In particular, new resources reallocations are performed for $t = k\hat{t},\ k = 0,\ 1,\ldots$. Let:

$$\begin{aligned}\boldsymbol{I}(k\hat{t}) = col\{\ &\boldsymbol{\theta}^d((k-\Xi)\hat{t}),\ldots,\boldsymbol{\theta}^d((k-1)\hat{t},\\ &\boldsymbol{\varphi}((k-\Xi)\hat{t}),\ldots,\boldsymbol{\varphi}((k-1)\hat{t}),\\ &\boldsymbol{\alpha}((k-\Xi)\hat{t}),\ldots,\boldsymbol{\alpha}((k-1)\hat{t}))\ \};\end{aligned} \quad (11)$$

$$\boldsymbol{\varphi}(k\hat{t}) = \{\phi_1(k\hat{t}),\ldots,\phi_N(k\hat{t})\};\boldsymbol{\alpha}(k\hat{t}) = \{\alpha_1(k\hat{t}),\ldots,\alpha_N(k\hat{t})\}$$

be an aggregate vector that maintains a finite horizon memory (of depth Ξ) over the values assumed by the decision variables and the measurable stochastic variables during the time interval $[(k-\Xi)\hat{t},(k-1)\hat{t}]$. Let $J_{k\hat{t}}(\boldsymbol{\theta}^d(k\hat{t})) = \underset{\boldsymbol{\varphi},\boldsymbol{\alpha}}{E}\ L_{k\hat{t}}[\boldsymbol{\theta}^d(k\hat{t}),\boldsymbol{\varphi},\boldsymbol{\alpha}]$ be the functional cost after such bandwidth reallocation (over an infinite time horizon), namely $L_{k\hat{t}}[\boldsymbol{\theta}^d(k\hat{t}),\boldsymbol{\varphi},\boldsymbol{\alpha}] =$

$$\lim_{T\to\infty}\frac{1}{T}\sum_{i=1}^{N}\int_{k\hat{t}}^{k\hat{t}+T}\gamma_i(\boldsymbol{\theta}^d(k\hat{t}),\phi_i(\tau),\alpha_i(\tau))\,d\tau \quad (12)$$

Let $\boldsymbol{f}(\boldsymbol{I}(k\hat{t}))$ be a reallocation law, which provides a surrogate continuous bandwidth allocation $\boldsymbol{\theta}^c(k\hat{t}) \in \Theta_c$ as a function of the current information vector:

$$\boldsymbol{\theta}^c(k\hat{t}) = \boldsymbol{f}(\boldsymbol{I}(k\hat{t}))\ . \quad (13)$$

Then, by converting $\boldsymbol{\theta}^c(k\hat{t})$ into its discrete feasible neighbour $\boldsymbol{\theta}^d(k\hat{t}) \in \Theta_d$ through the algorithm mentioned in the Section IV, Problem D-SDRAP now becomes a functional optimization problem.

Problem FD-SDRAP (*Functional Dynamic-Stochastic Discrete Resource Allocation Problem*): find the optimal bandwidth reallocation function $\boldsymbol{f}^*(\cdot)$ such that $\boldsymbol{f}^*(k\hat{t}) \in \Theta_d,\ k = 0,\ 1,\ldots$, and the cost:

$$J_{k\hat{t}}(\boldsymbol{f}(\boldsymbol{I}(k\hat{t}))) = \underset{\boldsymbol{\varphi},\boldsymbol{\alpha}}{E}\ L_{k\hat{t}}[\boldsymbol{f}(\boldsymbol{I}(k\hat{t})),\boldsymbol{\varphi},\boldsymbol{\alpha}] \quad (14)$$

is minimized. Such formulation yields a closed-loop resource allocation law, able to perform on line dynamic control reactions to variable traffic and fading levels conditions. However, at time $k\hat{t}$, new reallocations are performed "as if" they were to be applied without changes in the future, i.e., possible changes in the statistics of the random variables are disregarded. Then, the control strategy resulting from the solution of Problem FD-SDRAP for successive time instants $k\hat{t}, k = 0,\ 1,\ldots$ can be considered as an *Open Loop Feedback Control* (OLFC) one.

A. The Modified Extended Ritz method

In order to approximate the optimal OLFC allocation law $\boldsymbol{f}^*(\cdot)$, we investigate an extension of the *Extended Ritz* method ([5]). The Extended Ritz method is a technique that consists of approximating the solution of a functional optimization problem, by fixing the structure of the decision functions; namely, such decision functions are constrained to take on the structure of one hidden layer networks, i.e., linear combinations of basis functions containing free parameters to be optimized:

$$\bar{\boldsymbol{f}}(\boldsymbol{I},\boldsymbol{w}) = col\left\{\sum_{l=1}^{\upsilon} c_l^i\, p(\boldsymbol{I},\tilde{\boldsymbol{w}}) + c_0^i; i = 1,\ldots,N\right\} \quad (15)$$

$$\boldsymbol{w} = \left\{\tilde{\boldsymbol{w}}_l,\ l = 1,\ldots,\upsilon;\ c_l^i,\ l = 1,..,\upsilon,\ i = 1,\ldots,N;\ c_0^i,\ i = 1,\ldots,N\right\}$$

where $p(\cdot,\cdot)$ represents a suitable basis function.

Among the possible choices of structures of the form (15), we choose *one hidden layer feedforward neural networks*, in virtue of their powerful approximation properties to face the possible exponential growth in the number of free parameters needed to obtain a growing degree of accuracy. Then, we have:

$$\bar{\boldsymbol{f}}(\boldsymbol{I},\boldsymbol{w}) = col\left\{\sum_{l=1}^{\upsilon} c_l^i \cdot \sigma(\tilde{\boldsymbol{w}}_{1l}^T\boldsymbol{I} + \tilde{\mathrm{w}}_{0l}) + c_0^i;\ i = 1,\ldots,N\right\} \quad (16)$$

where $\sigma(x) = 1/(1+e^{-x})$ is a sigmoidal activation function. With such a choice, we have $\tilde{\boldsymbol{w}}_l = col\{\tilde{\boldsymbol{w}}_{1l},\tilde{\mathrm{w}}_{0l}\},\ l = 1,\ldots,\upsilon$. In order to guarantee the fulfilment of the "continuous" channel constraints (10) (i.e., $\boldsymbol{\theta}^c(t) \in \Theta_c, \forall t$), we compose the output of the neural network with a "normalization operator" $\boldsymbol{n}(\cdot)$. We thus obtain the surrogate continuous bandwidth allocation $\boldsymbol{\theta}^c(k\hat{t})$ at any time $k\hat{t}$ as:

$$\boldsymbol{\theta}^c(k\hat{t}) = \boldsymbol{n}\left[\bar{\boldsymbol{f}}(\boldsymbol{I}(k\hat{t}),\boldsymbol{w})\right] \quad (17)$$

$$\boldsymbol{n}(\bar{\theta}) = col\left\{K\cdot\sigma(\bar{\theta}_i)\Big/\sum\nolimits_{j=1}^{N}\sigma(\bar{\theta}_j);\ i = 1,\ldots,N\right\}.$$

The surrogate continuous bandwidth allocation $\boldsymbol{\theta}^c(k\hat{t})$ can be converted to the feasible bandwidth allocation $\boldsymbol{\theta}^d(k\hat{t}) \in \Theta_d$. We call "*neural approximators*" the functions (17) obtained as composition of the neural networks (16) and the normalization operators $\boldsymbol{n}(\cdot)$. Moreover, we shall call "*neural bandwidth allocation functions*" the mappings made up by the composition of the neural approximators together with the application of the algorithm that projects each $\boldsymbol{\theta}^c(k\hat{t})$ to its discrete feasible neighbour $\boldsymbol{\theta}^d(k\hat{t})$ and denote it as $\boldsymbol{\theta}^d(k\hat{t}) = \hat{\boldsymbol{f}}(\boldsymbol{I}(k\hat{t}), \boldsymbol{w})$. It follows that a cost function is obtained by substituting the fixed structure of such neural bandwidth allocation function into cost (14), depending on the parameter vector $\boldsymbol{w}$, thus leading to the following mathematical programming problem.

Problem FD-SDRAP$_{\mathrm{w}}$. Find the optimal parameter vector $\boldsymbol{w}^*$ such that the cost

$$\underset{\boldsymbol{\varphi},\boldsymbol{\alpha}}{E}\, L_{k\hat{t}}\left[\hat{\boldsymbol{f}}(\boldsymbol{I}(k\hat{t}),\boldsymbol{w}),\boldsymbol{\varphi},\boldsymbol{\alpha}\right] \qquad (18)$$

is minimized. In this way, the functional optimization Problem FD-SDRAP (14) has been reduced to an unconstrained nonlinear programming one.

B. The training algorithm

To solve such nonlinear programming problem, we apply a gradient-based algorithm:

$$\boldsymbol{w}^{h+1} = \boldsymbol{w}^h - \xi\, \nabla_{\boldsymbol{w}} \underset{\boldsymbol{\varphi},\boldsymbol{\alpha}}{E}\, L_{k\hat{t}}(\boldsymbol{w}^h,\boldsymbol{\varphi},\boldsymbol{\alpha}),\ h = 0,1,2,\ldots \quad (19)$$

However, the explicit computation of the expected cost and its gradient is a very hard task, even if closed-form formulas for the functional cost $L_{k\hat{t}}(\cdot)$ were available. We choose to compute a realization $\nabla_{\boldsymbol{w}} L_{k\hat{t}}(\boldsymbol{w}^h,\boldsymbol{\varphi}^h,\boldsymbol{\alpha}^h)$ instead of the gradient $\nabla_{\boldsymbol{w}} \underset{\boldsymbol{\varphi},\boldsymbol{\alpha}}{E}\, L_{k\hat{t}}(\boldsymbol{w}^h,\boldsymbol{\varphi},\boldsymbol{\alpha})$ and we apply the updating algorithm:

$$\boldsymbol{w}^{h+1} = \boldsymbol{w}^h - \xi_h\, \nabla_{\boldsymbol{w}} L_{k\hat{t}}(\boldsymbol{w}^h,\boldsymbol{\varphi}^h,\boldsymbol{\alpha}^h),\ h = 0,1,2,\ldots \quad (20)$$

where the index h denotes both the steps of the iterative procedure and the generation of the hth realization of the stochastic processes $\boldsymbol{\varphi}$ and $\boldsymbol{\alpha}$. The components of the gradient $\nabla_{\boldsymbol{w}} L_{k\hat{t}}(\boldsymbol{w}^h,\boldsymbol{\varphi}^h,\boldsymbol{\alpha}^h)$ can be obtained by using the classical backpropagation equations for the training of neural networks. The backpropagation procedure must be initialized by means of the quantities $\partial L / \partial \theta_i^c$, $i = 1,\ldots,N$ (i.e., the gradient $\nabla_{\boldsymbol{\theta}^c} L(\boldsymbol{\theta}^d,\boldsymbol{\varphi}^h,\boldsymbol{\alpha}^h)$). Unfortunately, in our case, such quantities cannot be obtained analytically as in [5], because no closed-form of the functional cost $L(\cdot)$ is available. In order to avoid such a heavy drawback, during the off-line training phase (20), we estimate the gradient $\nabla_{\boldsymbol{\theta}^c} L(\boldsymbol{\theta}^d,\boldsymbol{\varphi}^h,\boldsymbol{\alpha}^h)$ through the PA technique of [2], also employed to develop the on-line surrogate algorithm of Section IV.

In the following, we shall denote the application of this technique as "ExtRitz".

VI. SIMULATION RESULTS

We developed a C++ simulator for the satellite network under investigation and for the optimization algorithms proposed in this paper. The width of the confidence interval over the overall loss probability of the system was less then 1% for 95% of the cases. The number of active stations in the system is 2. We suppose no fading attenuation acting over the system (i.e., $\phi_i(t) = 1;\ i = 1,\ldots,N;\ \forall t$). Similar results are obtained in variable fading conditions. This was validated by simulation results not reported here.

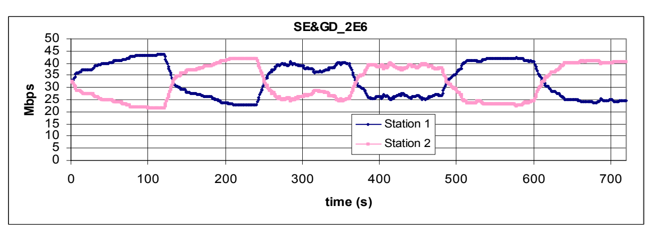
Each source is supposed to transmit at a peak bit rate B_p=1.0 Mbps, the total capacity K is fixed to 65.0 Mbps and the simulation time to 12 minutes. The number of connections for each station is fixed at $M^i = 70$ for all i. Variable systems conditions are taken into account according to Table 1. The maximum burst arrival rate is 35 bursts/s, but different lower traffic conditions can affect the satellite stations (14.0, 23.33, 17.5 bursts/s).

Time Interval (s)	**0.0 – 120.0**	**120.0 – 240.0**	**240.0– 360.0**
Station1, bursts/s	35.0	14.0	35.0
Station2, bursts/s	14.0	35.0	23.33
Time Interval (s)	**360.0 – 480.0**	**480.0 – 600.0**	**600.0– 720.0**
Station1, bursts/s	23.33	35.0	17.5
Station2, bursts/s	35.0	17.5	35.0

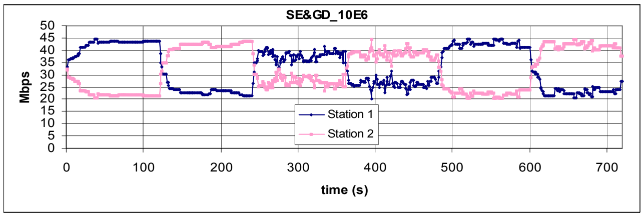
Table 1. Variable burst arrival rates.

Both stations are provided with a finite buffer of 100 ATM cells. The reallocation period $\hat{t}$ and the MAU are set to 1.0 and 100 Kbps, respectively (as, e.g., in [6]).

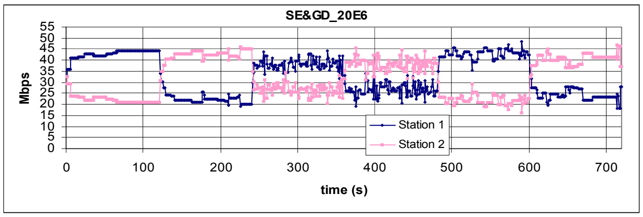
The depth Ξ of the time horizon of the information vector $\boldsymbol{I}(\cdot)$ in the ExtRitz technique was set to 5. The training phase took around 28 hours with an AMD Athlon @2.2GHz. As expected, different gradient step sizes yield different behaviours of the SE&GD technique (Fig. 1 a), b), c)). For each station, the fraction of the total system's capacity assigned by the SE&GD technique is visualized. It is clear how SE&GD is able to react to traffic variations. Sub-optimal transient periods are much more evident with $\eta = 2\cdot10^6$. It is worth noting that with too low gradient stepsize values (i.e., $\eta \in [10^1, 10^6]$) and with too high gradient stepsize values (i.e., $\eta \in [3\cdot10^7, 10^9]$) the SE&GD technique fails in optimizing the system performance.



a) SE&GD, $\eta = 2\cdot10^6$.



b) SE&GD, $\eta = 10 \cdot 10^6$.



c) SE&GD, $\eta = 20 \cdot 10^6$.

Figs. 1. SE&GD's allocations.

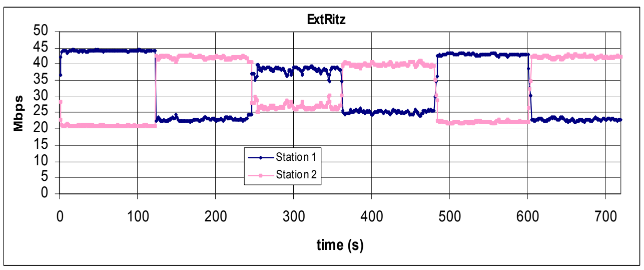


Fig. 2. ExtRitz's allocation.

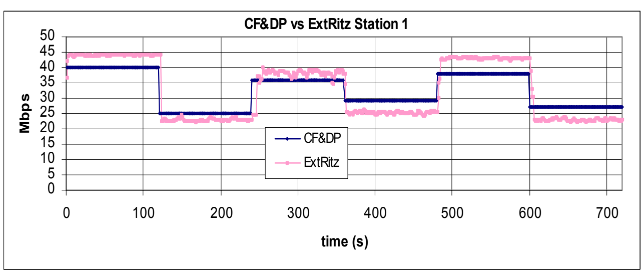


Fig. 3. ExtRitz versus CF&DP at station 1.

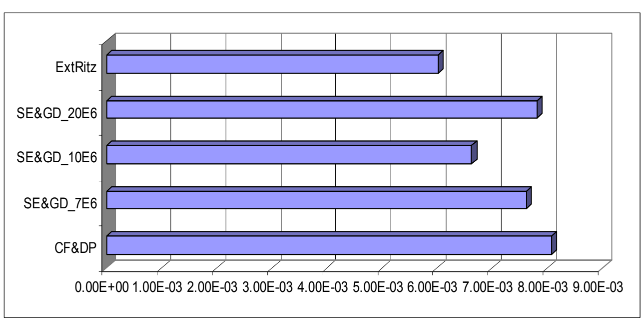


Fig. 4. Loss performance.

The best SE&GD's performance is achieved with $\eta = 10 \cdot 10^6$ (see Fig. 4), since good reactions to traffic changes are obtained without the strong bandwidth oscillations exhibited when $\eta = 20 \cdot 10^6$, especially in the middle of the simulation horizon. On the other hand, the ExtRitz technique (Fig. 2) always guarantees the best reactions in front of the variable traffic conditions. The CF&DP strategy does not achieve the optimal resource allocation, as is clear by comparing its allocations with respect to the ExtRitz's ones, for example at the first station (Fig. 3).

Looking at the ExtRitz bandwidth allocations, it is clear that (after the training phase) it provides the optimal bandwidth allocations, without any sub-optimal transient period. This fact, clearly, has an impact on the system performance (Fig. 4). The CF&DP does not maintain the best resource allocation: the bandwidth allocation to the station in heavier traffic load is lower than the one obtained by applying the ExtRitz technique (around 40.0 Mbps using CF&DP in front of the 45.0 Mbps obtained with ExtRitz).

The CF&DP technique, even if a perfect knowledge over the traffic sources' state is available, is not able to yield the optimal resource allocation. In fact, the *PLoss* formula (8) holds asymptotically in the number of sources [7] (i.e., the number of sources in the aggregated flows $\boldsymbol{\alpha}(t)$ should tend to infinity, together with $\bar{\varsigma}$). Hence, in a realistic scenario with a finite number of on-off sources, it can be seen only as a heuristic indication about the performance achieved by the system under the current state of the network. On the other hand, by exploiting sensitivity estimation, it is possible to capture the best resource allocation under variables systems conditions. The proposed ExtRitz technique makes use of PA to learn the best resource allocation off line, without the need to explicit any closed-form expression for the system dynamic equations and for the functional cost.

REFERENCES

[1] F. J. Vazquez-Abad, C. G. Cassandras, V. Julka, "Centralized and Decentralized Asynchronous Optimization of Stochastic Discrete-Event Systems," *IEEE Trans. Automat. Cont.*, vol. 43, no. 5, May 1998.

[2] Y. Wardi, Melamed, B., C.G. Cassandras , C.G.Panayiotou, "Online IPA Gradient Estimators in Stochastic Continuous Fluid Models," *J. of Optimization Theory and Applic.*, vol. 115, no. 2, pp. 369-405, Nov. 2002.

[3] K. Gokbayrak, C.G. Cassandras, "Generalized Surrogate Problem Methodology for Online Stochastic Discrete Optimization," *J. of Optimization Theory and Applic.*, vol. 114, no. 1, pp. 97-132, July 2002.

[4] F. Davoli, M. Marchese, M. Mongelli "Optimal Resource Allocation in Satellite Networks: Certainty Equivalent Approach versus Sensitivity Estimation Algorithms," *Int. J. of Commun. Sys.*, 2004, to appear.

[5] R. Zoppoli, M. Sanguineti, T. Parisini, "Approximating Networks and Extended Ritz Method for the Solution of Functional Optimization Problems," *J. of Optimization Theory and Applic.*, vol. 112, no. 2, pp. 403-439, Feb. 2002.

[6] N. Celandroni, F. Davoli, E. Ferro, "Static and Dynamic Resource Allocation in a Multiservice Satellite Network with Fading," *Internat. J. Satell. Commun. Netw.,* , vol. 21, no 4-5, pp. 469-487, July-Oct. 2003.

[7] B. Tsybakov, N.D. Georganas, "Self-Similar Traffic and Upper Bounds to Buffer-Overflow Probability in an ATM Queue," *Performance Evaluation*, vol. 32, pp. 57-80, 1998.