Distributed Cross-Layer Optimization of MANETs in Composite Fading

John Papandriopoulos, Subhrakanti Dey and Jamie S. Evans

ARC Special Research Centre for Ultra-Broadband Information Networks (CUBIN),
Department of Electrical and Electronic Engineering,
University of Melbourne, VIC 3010, Australia

Abstract— **Cross-layer design can significantly improve the performance of mobile ad-hoc networks (MANETs), as indicated by the flurry of recent results in the literature. Much of this work stems from the Kelly network utility maximization (NUM) framework, where convexity is crucial for developing algorithms that reach the global optimum. Unfortunately many problems are nonconvex in nature, so convex approximations are abundant.**

In this paper, we consider the joint optimization of source data-rates and link transmitter powers in a MANET, specifically dealing with the statistical variations of the wireless channel. In this paradigm we show that the commonly applied high-SIR convex approximation is unrealistic, so we seek to find solutions of the unmodified NUM problem. Our first result shows that the canonical formulation (previously thought to be nonconvex) is indeed a convex problem for logarithmic TCP-Vegas utilities; we then derive an algorithm reaching the global optimum. Our main result caters for the general case of strictly concave utilities, where we derive an algorithm that provably converges to the global solution of the underlying nonconvex NUM problem.

I. INTRODUCTION

THE traditional design of communication networks feature a layered network stack where each slice utilizes only the services of those below it, and each are optimized independently. Recent research efforts have shown that significant performance benefits exist in undertaking a cross-layer design; that is, by optimizing functionality across the layers.

Nonlinear optimization has been instrumental in this process. For example, in a multi-hop wired network, TCP algorithms are used to address problems of congestion at the transport layer. These algorithms have recently been shown to implement approximate solutions to an underlying Network Utility Maximization (NUM)

$$\max_{\mathbf{x} \geq 0} \quad \sum_s U_s(x_s) \qquad (1)$$
$$s.t. \sum_{s:l \in L(s)} x_s \leq c_l, \qquad \forall l,$$

where each source s attains some nonlinear utility $U_s(x_s)$ by transmitting at data-rate x_s, doing so without overwhelming the set of intermediate links $L(s)$ along its route, and where each link contributes a shared capacity of c_l. While this problem has been well studied in the context of wired networks having fixed link capacities [1], [2], more recent efforts have

focused on a cross-layer design of source-rate allocation and capacity provisioning [3].

In a wireless mobile ad-hoc network (MANET), we can conveniently optimize the link capacities c_l through power control: by increasing the transmitter power on link l, its capacity also increases and vice-versa. It is therefore of great benefit to consider maximizing the utility (1) over both source-rates *and* link-powers through a cross-layer design. In an interference limited system such as CDMA, this optimization becomes non-trivial since the power allocation is coupled across the entire network.

In [4], Chiang considered such a problem where the link capacities c_l in the basic NUM problem (1) became a function of the radio link quality, given by the signal-to-interference ratio (SIR). A high-SIR approximation was taken in order to formulate a convex problem that was solved via an iterative algorithm, made distributed through message passing. These messages carried congestion state and link noise measurements that were repeatedly broadcast to all or nearby nodes. Nodes receiving such messages would figure the transferred state into their own power-allocation.

In his model, Chiang assumed the wireless channel was fixed or very slowly varying. As such, it becomes conceivable that such an algorithm is able to track any changes in the channel, since one would have the luxury of iterating much faster than the fading dynamics. Now consider a more realistic scenario where the fading rate is increased, due to mobility of the nodes and/or environment. The iteration rate would also need to increase to track of the dynamic fading state. Either the message passing overhead would become excessive, or the instantaneous channel state varies too quickly to track properly – at this point the scheme would collapse.

The focus of this paper is on more realistic situations where the channel variation falls into this fast-varying dynamic category. Rather than finding the optimum source rates based on the *instantaneous* link capacities, we allow the network to experience a limited amount of *fading-induced congestion*, and by doing so, we avoid the fast-update problems outlined above. This concept is made rigorous through a *rate-outage probability*: the probability of experiencing fading-induced congestion. We reformulate the NUM problem to jointly allocate power and source rates such that the rate-outage probability is within some arbitrarily small tolerance.

We also do away with the high-SIR assumption. Under the

rate-outage metric, we show that such an assumption cannot be justified. Our first contribution involves the underlying canonical problem, previously thought to be nonconvex. In Sec. III-C we show that it is actually convex for logarithmic TCP-Vegas utilities, through a suitable transformation. Our results can then be directly back-substituted into [4]–[6] and related work to achieve true global optimality. Sec. IV-A introduces the resulting optimal algorithm A.

Our second contribution involves a generalization to strictly concave utilities in Sec. IV-B. There we derive algorithm B that utilizes a novel scheme to attain the global solution of the underlying nonconvex problem in a distributed manner. For the special-case of logarithmic utility, we show that this scheme preserves the existing TCP-Vegas stack.

II. SYSTEM MODEL

We consider a MANET having $\mathcal{L} = \{1, \ldots, L\}$ logical links, shared by $\mathcal{S} = \{1, \ldots, S\}$ sources.

A. Network Stack Decomposition

1) Network Layer: For each source $s \in \mathcal{S}$ there exists a destination node and we denote the path (route) from this source to the destination as an ordered set of links $L(s) \subseteq \mathcal{L}$. In this work, we assume that this layer is fixed. i.e. static routes. In practice, routing may be adapted and optimized, with these improved routes communicated to the transport layer below. Provided that the time-scale of such updates is much slower than the relevant time-scale of the cross-layer optimization presented here, any such adaption will not adversely interfere with the results provided.

2) Transport Layer: We make the usual assumption that each source s attains a utility $U_s(x_s)$ when allocated a data-rate $x_s \geq 0$, where $U_s(\cdot)$ is an increasing concave function.

3) Physical Layer: We extend the physical-layer model considered in [4] to incorporate composite fading. This model utilizes CDMA so that each link in the system may simultaneously communicate within the same spectrum allocation, at the expense of multiple-access interference.

This model has an underlying assumption that nodes are able to transmit and receive simultaneously. From an information-theoretic perspective, such a mode of operation is indeed possible: *two-way channels* have double the capacity region of the corresponding one-way channel [7]. Moreover, real implementations of simultaneous transceivers have been demonstrated in practice [8], [9], where the self-interference problem has been mitigated through RF isolators and echo-cancelers, coupled with base-band digital filtering (e.g. utilizing a CDMA spreading gain).

The alternative, that we will not consider in this paper for simplicity, is to emulate full-duplex operation with distinct transmit and receive transmission modes (time-division duplex). This is only possible if each node in the network is globally synchronized through a *link-activity schedule* and is akin to a hybrid CDMA-TDMA system. Finding optimal schedules is a very difficult problem as it is combinatorial in nature, and the topic is an important area of research in its own

right. Our results can easily be applied to such a paradigm, through a similar time-slotted system model as outlined in [6], [10] and references therein.

We consider fixed CDMA spreading sequences $\mathbf{s}_l$ having length N and unit energy (i.e. $\mathbf{s}_l^T \mathbf{s}_l = 1$). These sequences are preassigned to each link l where matched filtering is utilized. Additionally, we make the simplifying assumption that the self-interference cancellation at each node is perfect.

The instantaneous capacity of each link $l \in \mathcal{L}$ is modeled as $c_l(\mathbf{P}) = W \log (1 + K SIR_l(\mathbf{P}))$ where W is the base-band (unspread) bandwidth of the transmitted signal and K is a constant 'SIR gap' to capacity for a particular modulation and coding scheme [11].

The signal-to-interference ratio (SIR) is defined as

$$SIR_l(\mathbf{P}) = \frac{P_l F_{ll} H_{ll} (\mathbf{s}_l^T \mathbf{s}_l)^2}{\sum\limits_{j \neq l} P_j F_{lj} H_{lj} (\mathbf{s}_l^T \mathbf{s}_j)^2 + \sigma^2},$$

where $\mathbf{P} = [P_1, \ldots, P_L]$ is a vector of transmitter powers, $F_{ij} H_{ij}$ is the *instantaneous channel gain* from the transmitter on link j to the receiver on link i, and we assume thermal noise power σ^2 at each receiver.

Without loss of generality, we will absorb the $(\mathbf{s}_i^T \mathbf{s}_j)^2$ terms into the gain terms H_{ij} using the 'effective gains' $G_{ij} \equiv H_{ij} (\mathbf{s}_i^T \mathbf{s}_j)^2$ to simplify the SIR expression

$$SIR_l(\mathbf{P}) = \frac{P_l F_{ll} G_{ll}}{\sum\limits_{j \neq l} P_j F_{lj} G_{lj} + \sigma^2}.$$

B. Fading Model

We can decompose each instantaneous channel gain $F_{ij} G_{ij}$ into fast- and slowly-varying components F_{ij} and G_{ij} respectively. We assume Rayleigh fast-fading, where F_{ij} are iid random variables having an exponential distribution with unit mean. The terms G_{ij} model slow-fading (such as distance-dependent path-loss and/or log-normal shadowing) and are assumed constant over the time-scale of interest.

The independence assumption on each F_{ij} is justified since they represent the fading on *distinct* paths between nodes in the network. The unity mean is without loss of generality, as any non-unity value can be absorbed into the corresponding G_{ij} component.

It will be useful to define the 'average SIR',

$$\overline{SIR}_l(\mathbf{P}) = \frac{E[P_l F_{ll} G_{ll}]}{E\left[\sum_{j \neq l} P_j F_{lj} G_{lj} + \sigma^2\right]} = \frac{P_l G_{ll}}{\sum\limits_{j \neq l} P_j G_{lj} + \sigma^2} \tag{2}$$

where we have made use of the unity mean of each F_{ij}. To simplify the notation, we will occasionally drop this functional form and denote the average SIR as simply $\overline{SIR}_l$.

C. Rate-Outage Probability

We define the *rate-outage probability* as the probability that the ingress rate R_l to a link l exceeds its randomly time-varying capacity c_l, resulting in *fading-induced congestion*.

3880

Mathematically, we write

$$Pr\left\{R_l > c_l(\mathbf{P})\right\} = Pr\left\{F_{ll} < \frac{R_l^{th}}{P_l G_{ll}}\left[\sum_{j \neq l} P_j F_{lj} G_{lj} + \sigma^2\right]\right\}$$

where we define the *data-rate threshold* $R_l^{th} = \left[\exp\left(R_l/W\right) - 1\right]/K$.

In our Rayleigh model, we can bound the rate-outage probability above in terms of the average SIR (cf. [12], [13]),

$$Pr\left\{R_l > c_l(\mathbf{P})\right\} \leq 1 - \exp\left(-R_l^{th}/\overline{SIR}_l(\mathbf{P})\right) \quad (3)$$

where this bound was derived in [12] and has been shown to be very tight.

III. CROSS-LAYER DESIGN: OPTIMIZATION PROBLEMS

A. Nonconvex Problem Formulation

The joint power and congestion control optimization is written as the following NUM

$$\min_{\mathbf{x},\mathbf{P}\geq 0} (1-\omega)\sum_{l\in\mathcal{L}} P_l - \omega\sum_{s\in\mathcal{S}} U_s(x_s) \quad (4)$$

$$s.t.\ Pr\left\{\sum_{s:l\in L(s)} x_s > c_l(\mathbf{P})\right\} \leq \Omega_l, \quad \forall l \in \mathcal{L},$$

where we jointly optimize over the vector of nonnegative source rates $\mathbf{x} = [x_1,\ldots,x_S]$ and link transmitter powers $\mathbf{P} = [P_1,\ldots,P_L]$. The rate-outage probability targets $\Omega_l \in (0,1)$ for each link l are fixed inputs to the problem. The fixed scalar weight $\omega \in [0,1]$ is included to trade-off the conflicting goals of maximizing network utility (implying higher source-rates requiring higher link powers) and low power (to reduce interference and prolong the life of nodes relying on batteries).

To simplify the analysis, we make use of the bound (3) to re-formulate the constraint,

$$Pr\left\{\sum_{s:l\in L(s)} x_s > c_l(\mathbf{P})\right\} \leq \overbrace{1 - \exp\left(\frac{-R_l^{th}}{\overline{SIR}_l(\mathbf{P})}\right)}^{\text{new rate-outage constraint}} \leq \Omega_l.$$

Taking the logarithm of both sides of this new constraint and simplifying, we form a new optimization problem:

$$\min_{\mathbf{x},\mathbf{P}\geq 0} (1-\omega)\sum_{l\in\mathcal{L}} P_l - \omega\sum_{s\in\mathcal{S}} U_s(x_s) \quad (5)$$

$$s.t.\ \sum_{s:l\in L(s)} x_s \leq W \log\left(1 + \overline{K}_l \overline{SIR}_l(\mathbf{P})\right), \quad \forall l \in \mathcal{L},$$

where we have modified the SIR-gap

$$\overline{K}_l = K \times M_l(\Omega_l) \quad (6)$$

by the positive constant

$$M_l(\Omega_l) = -\log(1 - \Omega_l) \quad (7)$$

that is a function of the fixed rate-outage target Ω_l.

This optimization is of the same canonical form as considered in [4] with $\omega = 1$, however we deal with *average SIR*. Here we are not concerned with the instantaneous fading
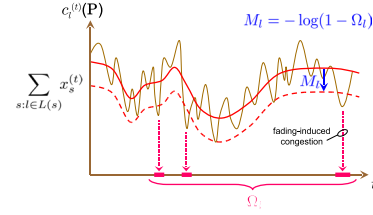

Fig. 1. Rate outage probability. Margin $M_l(\Omega_l)$ ensures fading-induced congestion is avoided with probability Ω_l.

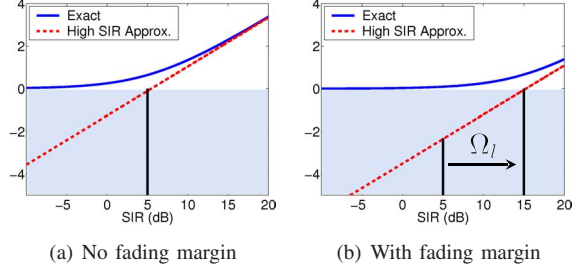(a) No fading margin　　　　(b) With fading margin

Fig. 2. High-SIR approx. requires excessively large SIR with fading margins.

state of the channel. Rather, we allocate a fading-margin (7) for the rate-allocation on each link. This margin is 'just the right amount' in a Rayleigh fading environment to meet the rate-outage probability target Ω_l and is illustrated in Fig. 1.

B. Prior Treatment: High SIR Approximations

The NUM problem above is not jointly convex in $\{\mathbf{x},\mathbf{P}\}$. Prior treatment made use of a high-SIR approximation on the link capacities. Coupled with the transformation $\widetilde{P}_n = \log P_n$, a convex problem resulted [4]–[6], [14]. Such a procedure is equivalent to the lower-bound approximation:

$$\sum_{s:l\in L(s)} x_s \leq \overbrace{W \log\left(\overline{K}_l \overline{SIR}_l(\mathbf{P})\right)}^{\text{lower-bound approximation}} \leq W \log\left(1 + \overline{K}_l \overline{SIR}_l(\mathbf{P})\right)$$

Fig. 2 shows that this approximation is far too loose for typical operating parameters, where $\overline{SIR} > 15$dB is required to realize a non-zero link rate.[1] Without margins, reasonable approximation accuracy is obtained at $SIR > 10$dB, jumping to $\overline{SIR} > 20$dB with margins employed. Such high operating points are undesirable for CDMA systems, as we must limit the interference by employing extremely large spreading gains, or complex interference suppression schemes so channels appear near-orthogonal. The high-SIR approximation clearly cannot be justified, nor is it required, as we now demonstrate.

C. A New Convex Approach for Logarithmic Utilities

Let us first consider utility functions having the form $U_s(x_s) = \nu_s \log x_s$. Such functions correspond to the TCP-Vegas protocol, where $\nu_s \equiv \alpha_s d_s$ is a scaled base round-trip-time (RTT) of source s [16]. We return to the general case of concave utilities in Sec. IV-B.

Our development makes use of the convexity results below. Proofs for Results 1 and 2 are straightforward and are omitted.

[1] $\Omega_l = 10\%, BER = 10^{-3}, K = \frac{-1.5}{\log(5BER)}$ (see [11], [14], [15])

3881

For notational convenience, we write the average SIR as

$$\overline{SIR}_l(\widetilde{\mathbf{P}}) = \frac{\exp(\widetilde{P}_l)G_{ll}}{\sum\limits_{j \neq l} \exp(\widetilde{P}_j)G_{lj} + \sigma^2},$$

where $\widetilde{P}_n = \log P_n$ is a logarithmic transformation.

Result 1 The function $h(x) = -\log\left(\log(1 + e^x)\right)$ is non-increasing and convex in $x \in \mathbb{R}$.

Result 2 The function $g_l(\widetilde{\mathbf{P}}) = \log\left(\overline{SIR}_l(\widetilde{\mathbf{P}})\right)$ is concave.

Result 3 The function $f_l(\widetilde{\mathbf{P}}) = -\log\left(\log(1 + \overline{SIR}_l(\widetilde{\mathbf{P}}))\right)$ is convex.

Proof: Write $f_l(\widetilde{\mathbf{P}}) = h(g_l(\widetilde{\mathbf{P}}))$ where $h(\cdot)$ and $g_l(\cdot)$ are given as in Results 1 & 2. The proof follows immediately from scalar composition [17, pg.84] and Results 1 & 2. ∎

We now show that the (nonconvex) formulation (5) is really convex with a transformation of coordinates $\widetilde{x}_s = \log x_s$ and $\widetilde{P}_l = \log P_l$. In these new variables, the problem is written

$$\min_{\widetilde{\mathbf{x}}, \widetilde{\mathbf{P}}} \quad (1-\omega)\sum_{l \in \mathcal{L}} e^{\widetilde{P}_l} - \omega \sum_{s \in \mathcal{S}} \nu_s \widetilde{x}_s \tag{8}$$

$$s.t. \quad \log\left(\frac{1}{W}\sum_{s:l \in L(s)} e^{\widetilde{x}_s}\right) - \log\log\left(1 + \overline{K}_l\overline{SIR}_l(\widetilde{\mathbf{P}})\right) \leq 0, \forall l.$$

The objective is a sum of convex exponentials and linear terms. Each constraint comprises a convex log-sum-exp term and a convex function of the powers $\widetilde{\mathbf{P}}$ (Result 3). Therefore, (8) is in standard convex form and the high-SIR approximation is not required.

IV. CROSS-LAYER DESIGN: IMPLEMENTATION

A. Algorithm A: Logarithmic Utilities

The convex optimization problem (8) is readily solved through its Lagrangian dual formulation, from which an iterative algorithm can be derived and convergence proved. We are assured that the global optimum is attained since there is no 'duality gap': Slater's condition is satisfied since a strictly feasible (yet suboptimal) solution is clearly possible.

Due to space limitations, we present only the final algorithm; full details can be found in [18].

1) Link Algorithm – Congestion Price Update: at each iteration t, link l updates a dual variable λ_l according to

$$\lambda_l^{(t+1)} = \left[\lambda_l^{(t)} + \epsilon\$_l(\mathbf{x}, \mathbf{P})\right]^+ \tag{9}$$

where ϵ is a sufficiently small step size, $[\cdot]^+ = \max(0, \cdot)$ and

$$\$_l(\mathbf{x}, \mathbf{P}) = \log\left(\frac{1}{W}\sum_{s:l \in L(s)} x_s^{(t)}\right) - \log\log\left(1 + \overline{K}_l\overline{SIR}_l^{(t)}\right)$$

is a measurable quantity local to each link. In a similar spirit to other related work in congestion control, these dual variables λ_l are interpreted as 'congestion prices' [16], here in logarithmic form, where $\$_l(\cdot, \cdot)$ intuitively represents the change in price. At equilibrium, $\$_l(\cdot, \cdot) = 0$ and the resulting power and source-rate allocation $\{\mathbf{P}^*, \mathbf{x}^*\}$ are optimal.

2) Link Algorithm – Power Update: power is allocated on each link l according to

$$P_n^{(t+1)} = \frac{\mu_n^{(t)}}{(1-\omega) + \sum\limits_{l \neq n} G_{ln}\Phi_l^{(t)}} \tag{10}$$

where

$$\mu_n^{(t)} = \frac{\lambda_n^{(t)}\overline{K}_n\overline{SIR}_n^{(t)}}{(1 + \overline{K}_n\overline{SIR}_n^{(t)})\log(1 + \overline{K}_n\overline{SIR}_n^{(t)})},$$

and $\Phi_l^{(t)} = \mu_l^{(t)}\dfrac{\overline{SIR}_l^{(t)}}{G_{ll}P_l^{(t)}} \in \mathbb{R}_+$ are messages comprising local information from each link l. Note the coupling of the congestion prices λ_l into the power allocation for each link n. They act as the glue between the layers as noted in [4].

This power update is made distributed through message passing, similar in concept to [4]: each **receiver** on link l broadcasts their message Φ_l. In turn, **transmitters** on link n receive it and estimate G_{ln} though training sequences. They then utilize the quantity to update their power through (10).

3) Source Algorithm – Rate Update: each source updates its rate according to

$$x_s^{(t+1)} = \frac{\omega\nu_s}{\sum\limits_{l \in L(s)} \Lambda_l^{(t)}} \tag{11}$$

where $\Lambda_l^{(t)} = \lambda_l^{(t)}/\sum_{m:l \in L(m)} x_m^{(t)}$ are congestion prices normalized by the ingress rate into each link l.

This update is also distributed through message passing: a reserved field in each acknowledgment (ACK) packet header (sent from the receiver back to the source) is used to accumulate the normalized congestion prices Λ_l at each intermediate link l. When the ACK reaches the source, this reserved field forms the summation in the denominator of (11).

B. Algorithm B: General Concave Utilities

We now develop an optimal scheme catering for *any increasing concave utility* $U_s(\cdot)$. Under the *special-case* of logarithmic utilities, the resulting algorithm makes use of the existing TCP-Vegas stack, employing delay measurements at each source for distributed rate-allocation. In turn, power control is based on the congestion prices, similar to the development above.

We make no assumptions on the $\overline{SIR}$ regime and therefore aim to solve the nonconvex problem (5) for the global optimum, and doing so in a distributed fashion. Analogous results dealing with concave utilities in [4] are based on a suboptimal convex approximation in a high-SIR regime.

Taking a slightly different direction to the efforts of Sec. III-B, we make use of the (new) bound

$$\alpha \log z + \beta \leq \log(1+z) \begin{cases} \alpha = \frac{z_0}{1+z_0} \\ \beta = \log(1+z_0) - \frac{z_0}{1+z_0}\log z_0 \end{cases} \tag{12}$$

that is tight with equality at a chosen value z_0 when the constants $\{\alpha, \beta\}$ are chosen appropriately.

With (12), we can form a new lower-bound approximation of the constraint in problem (5):

$$\overbrace{\sum_{s:l\in L(s)} x_s \leq W\alpha_l \log\left(\overline{K_l SIR_l}(\mathbf{P})\right) + W\beta_l}^{\text{new lower-bound approximation}} \quad (13)$$

$$\leq W\log\left(1 + \overline{K_l SIR_l}(\mathbf{P})\right)$$

where we have the additional freedom of making the bound exact for a given $z_0 = \overline{K_l SIR_l}(\mathbf{P})$ by choosing $\{\alpha_l, \beta_l\}$ as in (12). With this new bound, optimization (5) becomes

$$\min_{\mathbf{x},\mathbf{P}\geq 0} \quad (1-\omega)\sum_{l\in\mathcal{L}} P_l - \omega\sum_{s\in\mathcal{S}} U_s(x_s) \quad (14)$$

$$s.t. \quad \sum_{s:l\in L(s)} x_s \leq W\alpha_l \log\left(\overline{K_l SIR_l}(\mathbf{P})\right) + W\beta_l, \quad \forall l.$$

Taking a logarithmic transformation of co-ordinates $\widetilde{P}_l = \log P_l$ **in the link powers only** results in an equivalent convex problem for fixed $\{\alpha_l, \beta_l\}$: the objective is a sum of exponentials and negative-concave (thus convex) utilities; the constraint set is convex by Result 2.

Not surprisingly, we arrive at the same canonical form as Chiang [4] when $\{\alpha_l = 1, \beta_l = 0\}$. However, we have choice over the approximation constants and so the following procedure for adaptation then becomes natural:

1: initialize iteration counter $t = 0$
2: initialize all $\alpha_l^{(t)} = 1$, $\beta_l^{(t)} = 0$ (high-SIR approx.)
3: **repeat**
4:    *Minimize*: solve (14) for solution $\{\mathbf{x}^{(t)}, \mathbf{P}^{(t)}\}$
5:    *Tighten*: update $\alpha_l^{(t+1)}, \beta_l^{(t+1)}$ at $z_0 = \overline{K_l SIR_l}(\mathbf{P}^{(t)})$
6:    increment t
7: **until** convergence

Proposition 1 Each iteration t results in a monotonically improving objective. The sequence always converges, at which point the lower-bound approximation (13) becomes exact.

    Proof: See [18]; omitted due to space limitations. ∎

It turns out that each sub-problem need not be minimized fully: we require only an improved objective at each outer iteration t. This is the key to the formation of a distributed algorithm: each link l need not know when 'sub-problem t' has converged; each tightens at regular intervals with only local information of $\overline{K_l SIR_l}$.

A word of caution: our convex lower bounds are constructed with identical slope at the tightening point as their corresponding nonconvex bounding functions. For the sequence of sub-problems to converge, each partial minimization should be 'deeper' than a simple gradient descent step on the original nonconvex problem formulation. i.e. a Newton step, or a number $D > 1$ of descent steps should be performed at each minimization step before tightening.

We now summarize the solution to sub-problem (14). Alg. B comprises the solution procedure above, where only a small number $D > 1$ (typically 10) updates of each sub-problem are executed before the tightening step. The full development appears in [18].
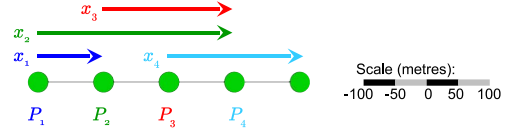

Fig. 3.   MANET topology with four sources.

1) Source Algorithm – Rate Update: Sources adapt their rate according to the well-known duality-model solution [16]

$$x_s^{(t+1)} = U_s'^{-1}\left(\sum_{l\in L(s)} \lambda_l^{(t)}\right), \quad (15)$$

where $U_s'^{-1}(\cdot)$ is the inverse of the first derivative.

For the logarithmic utilities $U_s(x_s) = \nu_s \log x_s$ of TCP-Vegas, the solution becomes: $x_s^{(t+1)} = \nu_s/\lambda^s$ where $\lambda^s = \sum_{l\in L(s)} \lambda_l^{(t)}$ is the accumulated congestion along each route. This can be implicitly obtained through delay measurements and thus explicit message passing becomes unnecessary [16]: we have effectively preserved the existing TCP-Vegas stack at each source. Further, it enables a MANET cloud to become transparently interconnected between other wired internet networks that make use of TCP-Vegas, and doing so without breaking end-to-end semantics.

2) Link Algorithm – Congestion Prices and Power Control: Congestion prices λ_l evolve according to the buffer process, implicitly given by the steepest ascent (9), where instead

$$\$_l(\mathbf{x},\mathbf{P}) = \sum_{s:l\in L(s)} x_s^{(t)} - \overbrace{W\left[\alpha_l \log\left(\overline{K_l SIR_l}(\mathbf{P}^{(t)})\right) + \beta_l\right]}^{\text{approximation of } c_l}.$$

Link transmitter powers are updated with

$$P_n^{(t+1)} = \frac{W\lambda_n \alpha_n}{(1-\omega) + \sum_{l\neq n} G_{ln}\Phi_l^{(t)}} \quad (16)$$

where $\Phi_l^{(t)} = W\lambda_l^{(t)}\alpha_l \frac{\overline{SIR_l}(\mathbf{P}^{(t)})}{G_{ll}P_l^{(t)}} \in \mathbb{R}_+$ are messages passed via broadcasts as in Sec. IV-A.

Theorem 1 Algorithm B solves the nonconvex problem (5) for the global optimum power- and rate-allocation.

    Proof: See [18]; omitted due to space limitations. ∎

V. ILLUSTRATIVE NUMERICAL EXAMPLE

Consider the network topology of Fig. 3. Each link is assigned a random CDMA spreading sequence of unit energy having length $N = 32$. The composite fading channel is modeled by a slowly-varying gain $G_{ij} = H_{ij}(\mathbf{s}_i^T \mathbf{s}_j)^2$ as outlined in Section II-B, where $H_{ij} = d_{ij}^{-4}$ is a loss depending on distance d_{ij} from transmitter on link j to the receiver on link i, assumed fixed. A transmission bandwidth of 1 MHz is selected, giving a $W = 31.25$ kHz baseband on each link. We model the SIR gap $K = -1.5/\log(5BER)$ for a BER of 10^{-3} corresponding to MQAM modulation [11].

For a fair comparison of Alg. A and B, we consider logarithmic utilities throughout. A scaled base-RTT of $\nu =$

(a) Convergence comparison. (b) Link Transmitter Powers (Alg. B)

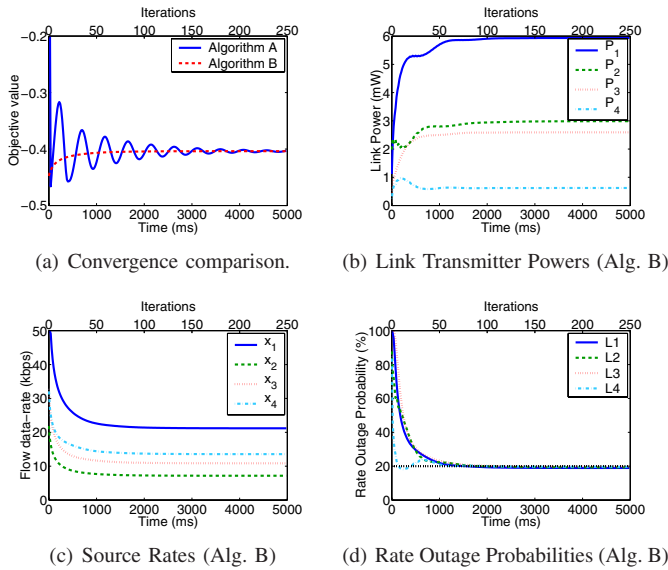(c) Source Rates (Alg. B) (d) Rate Outage Probabilities (Alg. B)

Fig. 4. Numerical results with perfect estimates of slowly-varying channel gains (fast-varying gains are iid Rayleigh unit-mean and are perfectly averaged over). Top-left graph shows the convergence of each algorithm. The remaining three graphs show Alg. B converge to the optimum powers and rates, while the rate outage-probability converges to the target of 20% (dotted). Each update occurs at 20ms intervals, corresponding to 250 iterations in total.

20 ms is arbitrarily chosen for each source. We consider an equal trade-off between power-allocation and network-utility with $\omega = 1/2$, and rate-outage targets $\Omega_l = 20\%$ for all links.

Fig. 4 summarizes numerical results under the assumption of perfect knowledge of slowly-varying gains G_{ij}. In all cases, rate-outage targets are met with equality, see Fig. 4(d). We plot the objective as a function of the iteration number in Fig. 4(a). Alg. B converges much faster than Alg. A, and both to the global optimum within 300 (Alg. A) and 75 (Alg. B) iterations respectively. This is significant, as we have effectively solved a nonconvex problem formulation for the global optimum with a distributed algorithm, preserving the existing TCP-Vegas stack in the process: no explicit message passing is required for source rate allocation.

We have assumed perfect knowledge of the slowly varying gains G_{ij}. In [18] we show that with *imperfect* channel state knowledge, these algorithms are found to be robust with desirable tracking and convergence properties.

VI. CONCLUSION

This paper has introduced a new cross-layer design paradigm for multi-hop mobile networks in interference-limited fading environments. Explicitly taking the channel variation into account, a new NUM problem was formulated that considered the joint source rate- and link power-allocation, without any of the high-SIR assumptions commonly found in the literature. Two schemes were developed that required only the large-scale fading state, thus avoiding the fast-update requirements and large overheads of message passing associated with other related work.

The first scheme involved a NUM problem previously thought to be nonconvex, involving logarithmic utilities. It was shown that through a suitable transformation, the problem is actually convex and an optimal algorithm was outlined.

A second scheme was given to solve a more general nonconvex NUM problem having concave utilities. At the core was a series of convex sub-problems requiring only partial solution, leading to a distributed algorithm without complicated and centralized branch-and-bound or cutting-plane methods that are typically used to solve such problems. An algorithm was provided that provably converges to the globally optimum joint power- and rate-allocation.

The general results presented in this paper may be directly back-substituted into other related problems found in the literature that had previously relied on a convex high-SIR approximation. Consequently, such problems may now be solved for the true global optimum with the results herein.

REFERENCES

[1] F. P. Kelly, "Charging and rate control for elastic traffic," *Euro. Trans. on Telecom.*, vol. 8, pp. 33–37, Jan. 1997.

[2] S. Low and D. Lapsley, "Flow control, I: Basic algorithm and convergence," *IEEE/ACM Trans. on Net.*, vol. 7, no. 6, pp. 861–875, Dec. 1999.

[3] J. Wang, S. H. L. L. Li, and J. C. Doyle, "Cross-layer optimization in TCP/IP networks," *IEEE/ACM Trans. on Net.*, vol. 13, no. 3, pp. 582–268, June 2005.

[4] M. Chiang, "Balancing transport and physical layers in wireless multi-hop networks: Jointly optimal congestion control and power control," *J. Sel. Areas in Comm.*, vol. 23, no. 1, pp. 104–116, 2005.

[5] M. Johansson, L. Xiao, and S. Boyd, "Simultaneous routing and power allocation in CDMA wireless data networks," in *IEEE Int. Conf. on Comm.*, May 2003.

[6] R. Madan, S. Cui, S. Lall, and A. Goldsmith, "Cross-layer design for lifetime maximization in interference-limited wireless sensor networks," in *IEEE INFOCOM*, Mar. 2005.

[7] K.-H. Pan, H.-K. Wu, R.-J. Shang, F. Lai, and Y.-W. Lin, "Communications over two-way waveform channels in wireless networks," in *Elec. and Computer Eng., 1999 IEEE Canadian Conf. on*, Edmonton, Alta., May 1999, pp. 45–50.

[8] K. Tsubouchi, H. Nakase, A. Namba, and K. Masu, "Full duplex transmission operation of a 2.45-GHz asynchronous spread spectrum using a SAN convolver," *IEEE Trans. on Ultrasonics, Ferroelectrics and Frequency Control*, vol. 40, no. 5, pp. 478–482, Sept. 1993.

[9] S. Chen, M. A. Beach, and J. P. McGeehan, "Division-free duplex for wireless applications," *Electr. Letters*, vol. 34, no. 2, pp. 147–148, Jan. 1998.

[10] U. C. Kozat, I. Koutsopoulos, and L. Tassiulas, "A framework for cross-layer design of energy-efficient communication with QoS provisioning in multi-hop wireless networks," in *IEEE INFOCOM*, vol. 2, Mar. 2004, pp. 1446–1456.

[11] A. Goldsmith, *Wireless Communications*. Cambridge Univ. Press, 2005.

[12] J. Papandriopoulos, J. S. Evans, and S. Dey, "Optimal power control for Rayleigh-faded multiuser systems with outage constraints," *IEEE Trans. on Wireless Comm.*, vol. 4, pp. 2705–2715, Nov. 2005.

[13] S. Kandukuri and S. Boyd, "Optimal power control in interference-limited fading wireless channels with outage-probability specifications," *IEEE Trans. on Wireless Comm.*, vol. 1, no. 1, pp. 46–55, 2002.

[14] D. Julian, M. Chiang, D. O'Neill, and S. Boyd, "QoS and fairness constrained convex optimization of resource allocation for wireless cellular and ad hoc networks," in *IEEE INFOCOM*, vol. 2, June 2002, pp. 477–486.

[15] X. Qiu and K. Chawla, "On the performance of adaptive modulation in cellular systems," *IEEE Trans. on Comm.*, vol. 47, no. 6, pp. 884–895, June 1999.

[16] S. Low, L. Peterson, and L. Wang, "Understanding TCP Vegas: a duality model," *J. ACM*, vol. 49, no. 2, pp. 207–235, 2002.

[17] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge Univ. Press, 2004.

[18] J. Papandriopoulos, *PhD Thesis*. The University of Melbourne, 2005. [Online]. Available: http://jpap.yi.org/research