Dynamic Coding and Rate-Control for Serving Deadline-Constrained Traffic over Fading Channels

Harsha Gangammanavar and Atilla Eryilmaz

Abstract—We study the problem of optimal dynamic coding and rate-control for broadcasting deadline-constrained traffic with average delivery ratio constraints over time-varying wireless channels. In particular, we propose and analyze a novel policy that utilizes a combination of *pricing* and *finite-horizon dynamic programming* strategies to jointly optimize the operation of the following two components: (i) a dynamic rate allocation policy, which manages the incoming traffic flow rates so as to maximize their weighted sum; (ii) and a dynamic coding window selection policy, which satisfies the flows' individual delivery ratio requirements. Our heuristic fluid analysis of the resulting stochastic network operation indicates that our joint policy maximizes the weighted sum of the deadline-constrained flow throughput subject to heterogeneous delivery ratio requirements imposed on them. We also apply these general results to an important cellular downlink scenario with and without network coding capabilities to study its behavior under various conditions. Our simulations reveal that the dynamic coding strategy outperforms the optimal static coding strategy by opportunistically exploiting the statistical variations in the arrival and channel processes.

Index Terms—Delay-aware dynamic coding, stochastic control, network coding, deadline-constrained throughput optimization.

I. INTRODUCTION

While the traditional performance measure of a communication system is *throughput*, many real-world applications also have a range of delay-sensitivities and Quality-of-Service (QoS) requirements that are not accounted for. In particular, real-time media broadcasting or two-way voice/video communication applications possess very stringent deadline constraints and differing tolerance levels to dropped bits. Moreover, different flows entering the system may have different degrees of importance, necessitating prioritization of certain flows over the others.

Information theory reveals that the reliable transmission rate over a channel may be increased towards its capacity by block coding over increasing *coding window* (or *coding block*) *sizes*. Yet, increasing the coding window size also causes larger decoding delay, which is detrimental to the service quality in the presence of the aforementioned delay sensitive traffic. Thus, coding decisions must account for the diverse and multi-timescale flow requirements for satisfactory performance.

With this motivation, we study the optimal service of the aforementioned deadline-constrained flows in the presence of various block coding strategies that the transmitter can employ. Each such strategy is described by a *completion time*

distribution, that models the random time it takes for a coding window of a given size to be successfully transmitted over the communication channel[1]. The main objective of the work is to design optimal rate-control and coding strategies for serving deadline-constrained flows with varying delivery ratio requirements and priorities. Our main contributions are:

- We provide a generic communication system model (in Section II-A) with a multi-timescale operation that accounts for the heterogenous priorities and the multi-timescale requirements of the incoming traffic. We then formulate (in Section II-B) an infinite-horizon stochastic control problem of coding and rate-control for maximizing a weighted-sum of the flow rates subject to both their short-term deadline constraints and their long-term delivery ratio requirements.

- We propose (in Section III) a novel dynamic algorithm that jointly determines the incoming flow rates and the coding operation to solve the above stochastic problem. A novelty of this algorithm is its use of *dynamic pricing* at a slow timescale together with *finite-horizon dynamic programming* at a fast timescale to achieve the formulated infinite-horizon global objective. The global asymptotic optimality of a heuristic fluid model of this algorithm is provided in [6].

- We apply (in Section IV) the developed algorithm to an important application in cellular downlink scenario whereby a base station broadcasts multiple streaming deadline-constrained flows to N receivers over randomly varying erasure channels. We study the performance of our dynamic algorithm with and without *network coding* capabilities, and compare the dynamic policy performance to a static one to see strict improvements, even for small scenarios.

We note that queueing systems under impatient customers have been studied in the literature (e.g. [2], [10], [15]) for various cases of preemption, arrival/service rate distributions, etc. Yet, these works do not model the priorities and tolerance levels of applications, and do not account for possible coding parameters, and hence are not applicable to our problem. Also, recent works ([7], [8]) have studied the congestion-control and scheduling problem for similar deadline constrained traffic with delivery ratio constraints. However, they also do not allow for coding flexibilities, which fundamentally changes the shape of the achievable rate region and calls for a dynamic strategy for optimizing over coding decisions. Our work also differs from [14], [11] in considering multiple flows, a different quality of metric, and presence of a rate controller.

[1]We note that this is an alternative description of the coding performance to the traditional one that describes the probability of decoding error for a fixed transmission duration. We find that this alternative model is more useful for our design and analysis.
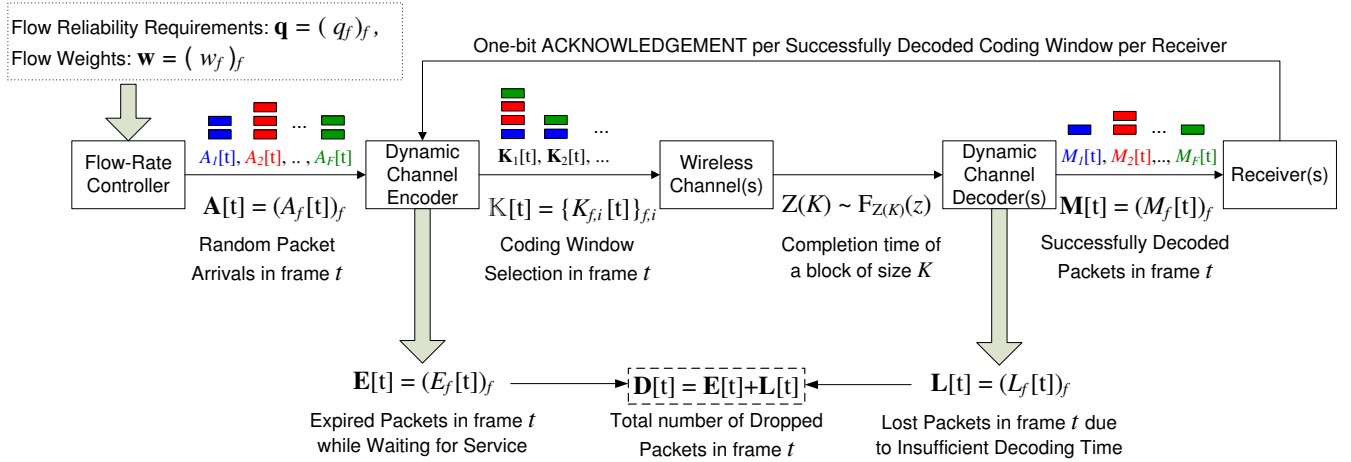
Fig. 1. Block diagram of a generic communication system for serving deadline-constrained flows with varying delivery ratio requirements and priorities.

II. SYSTEM MODEL AND PROBLEM FORMULATION

We study the general communication system depicted in Figure 1, whereby a transmitter serves a set $\mathcal{F}$ of deadline-constrained flows over unreliable wireless channel(s) subject to long-term delivery ratio constraints. Our work concerns the optimal design of the flow-rate controller and the dynamic channel encoder-decoder pair that operate at different timescales. Next, we describe the system components and operation in detail.

A. Detailed Description of the System Components

We start with the description of the multi-timescale system operation before discussing the details of the flow-rate controller and dynamic channel encoder/decoder operation along with the hard deadline constraints and soft delivery ratio constraints imposed by the flows.

Multi-timescale Operation: We setup a multi-timescale system operation whereby the flow-rate controller is allowed to operate in the slower time scale of flow-level deadline constraints. And the dynamic channel encoder/ decoder operates at a faster time-scale of channel variations.Accordingly, we use *time slots* as the smallest time unit in which these channel variations occur. In comparison, the rate controller operates in a slower timescale of what we call a *frame* of τ time slots, which is also the *hard deadline* within which each incoming packet must be successfully delivered to the destination(s), or otherwise is dropped (see Figure 2).

Flow-Rate Controller: The packets of each flow are generated at the beginning of each frame according to the *Flow-Rate Controller* mechanism (see Figure 1) that is to be designed. In particular, at the beginning of Frame t, the rate controller generates a random number, $A_f[t]$, of Flow-f packets for transmission within the frame. We also use $A_f(T) \triangleq \sum_{t=1}^{T} A_f[t]$ to denote the cumulative sum of all Flow-f arrivals in the first T frames of the system operation, and use $\mathbf{A}[t] = (A_f[t])_f$ to denote the vector of all flow packets generated in Frame t (see Figure 2).

Dynamic Channel Encoding/Decoding: The vector $\mathbf{A}[t] = (A_f[t])_f$ of packet arrivals enter the *Dynamic Channel Encoder* (see Figure 1) that performs block coding for reliable transmission. In particular, the encoder distributes the available

packets into groups, called *coding windows* (or simply *blocks*), for sequential transfer over the channel.
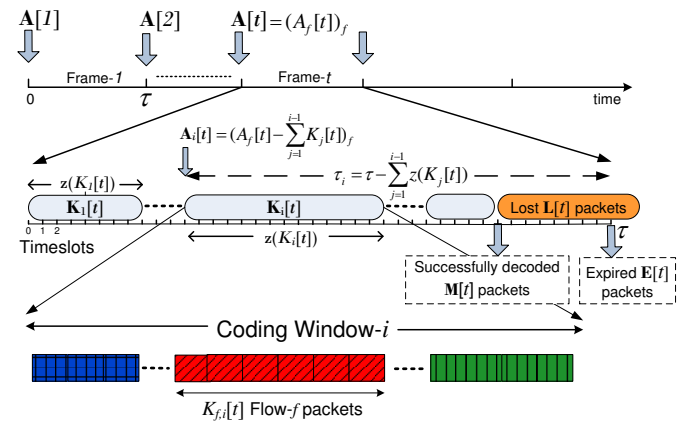

Fig. 2. Operation of the communication system of Figure 1 over time.

Let the vector $\mathbf{K}_i[t] = (K_{f,i}[t])_f$ denote the composition of the ith coding window in Frame t, where $K_{f,i}[t]$ gives the number of Flow-f packets in the ith coding window. Also, let $K_i[t] \triangleq \sum_f K_{f,i}[t]$ denote the total number of packets in the ith block. The blocks are constructed and transmitted sequentially such that the $(i+1)^{st}$ block is constructed and transmitted only after the ith block is successfully decoded by all the intended receivers. The successful decoding of each block is indicated through a single bit ACK signal by each intended receiver (see Figure 1).

We emphasize that the amount of time required for the successful decoding of a coding block of size K packets, henceforth called the *completion time*, is a random variable that depends on K, the channel statistics, and the coding strategy employed by the encoder. We capture this randomness by letting $Z(K)$ denote the completion time (in time slots) of a block of size K packets, which is generated independently according to a given distribution function $F_{Z(K)}(z)$. In Section IV, we shall provide examples of such distribution functions in downlink broadcast setup over erasure channels, but our construction applies to any distribution.

We also note that the encoder only knows the statistics of $Z(K)$ at the outset of its block construction, but can deduce the realization of *earlier block completion times* through the

acknowledgements it receives. To distinguish the two, we let $z(K)$ denote the realization of the random completion time $Z(K)$ with the convention that $z(K) = \infty$ if the frame ends before the block completion. Then, the time left in Frame t at the beginning of the ith block transmission is given by $\tau_i \triangleq \tau - \sum_{j=1}^{i-1} z(K_j[t])$. Similarly, the vector $\mathbf{A}_i[t] \triangleq (A_f[t] - \sum_{j=1}^{i-1} K_{f,j}[t])_f$ denotes the remaining packets of each flow at the beginning of the ith coding window construction within Frame t (see Figure 2). Thus, the construction of $\mathbf{K}_i[t]$ will depend on the remaining time τ_i in the frame, the remaining packets $\mathbf{A}_i[t]$ awaiting service, and the distribution of $Z(K)$.

For notational convenience, we use the matrix $\mathbb{K}[t] \triangleq [\mathbf{K}_1[t] \ \mathbf{K}_2[t] \ \cdots]$ to compactly refer to the sequence of such coding window decisions made in Frame t. Also, we use $K^f[t]$ to denote the total number of Flow-f packets over the constructed coding blocks in Frame t.

Measure of Transmission Success or Failure: If the frame ends during the completion of a coding block transmission, we assume that all the packets in that block are lost. The number of lost packets in Frame t is denoted by $\mathbf{L}[t] = (L_f[t])_f$. Similarly, the vector $\mathbf{M}[t] = (M_f[t])_f$ denotes the vector of successfully decoded packets of each flow in Frame t (see Figures 1 and 2). Since the completion time is random, $\mathbf{M}[t]$ is also random and can be described as a complicated function $g(\mathbf{Z}(\mathbb{K}[t]))$ of the completion times under the particular coding matrix $\mathbb{K}[t]$ (see [6] for a more detailed discussion on $g(\cdot)$). Thus, we denote the distribution of $\mathbf{M}[t]$ with $F_{g(\mathbf{Z}(\mathbb{K}[t]))}(m)$. Further, we let $M_f(T) \triangleq \sum_{t=1}^{T} M_f[t]$ denote the total Flow-f service in the first T frames of the system operation.

Note that some packets may not even get a chance to start their transmission before the end of their frame. These packets are assumed to *expire* at the end of the frame, with $\mathbf{E}[t] = (E_f[t])_f$ denoting the number of such expiries. Together with the losses, the total number of dropped packets $\mathbf{D}[t] = (D_f[t])$ in Frame t equals $\mathbf{E}[t] + \mathbf{L}[t]$ (see Figures 1 and 2).

Application Preferences and Requirements: So far, we have only assumed that packet transmissions of each flow is considered successful only if they are successfully decoded within a single frame, after their arrival. Next, we differentiate flows according to their *relative priority* and their *long-term delivery ratio requirements*. To model their relative priority, we assign a set of positive weights $\mathbf{w} = (w_f)_f$ to each flow, and try to maximize the weighted sum of their mean flow rates. On the other hand, their long-term delivery ratio requirement is measured by $\mathbf{q} = (q_f)_f$, where $q_f \in [0, 1)$ imposes an upper bound on the *fraction of dropped Flow-f packets in the long run*. In other words, we are forced to guarantee that

$$\liminf_{T \to \infty} \frac{\mathbb{E}[M_f(T)]}{\mathbb{E}[A_f(T)]} \geq (1 - q_f), \quad \forall f, \quad (1)$$

Hence, it is required that at least $(1-q_f)$ fraction of arriving Flow-f packets be served within their deadline on average. Thus, the requirement gets more stringent as q_f decreases towards 0. Real-time applications such as voice/video transfers can typically tolerate a certain fraction of packet losses, and hence have such *delivery ratio requirements*. This traffic modeling follows that of [7], [8], and is attractive for both practical modeling and theoretical analysis purposes.

B. Stochastic Control Problem Formulation

For the system model and traffic requirements presented in Section II-A, our main objective is to design a joint rate controller and dynamic coding strategy that maximizes $\mathbf{w}$-weighted sum of generated flow rates subject to the short-term deadline constraint τ of all packets and the long-term delivery ratio requirements $\mathbf{q}$ of the flows given in (1). In this section, we provide an *infinite-horizon optimal stochastic control* formulation of this qualitative objective.

For ease of exposition, we define expressions for the long-term flow rates and successful service rates as:
$\overline{\lambda}_f := \limsup_{T \to \infty} \mathbb{E}[A_f(T)]/T$, $\underline{\lambda}_f := \liminf_{T \to \infty} \mathbb{E}[A_f(T)]/T$ and $\underline{\mu}_f := \liminf_{T \to \infty} \mathbb{E}[M_f(T)]/T$, where the expectations are over the randomness of the packet generation process and the completion times , which are, in turn, functions of the coding strategy and the channel variations. We are now ready to pose our stochastic control problem (SCP) as:

Definition 1 (∞-Horizon Stochastic Control Problem):

(SCP): $\underset{\{\mathbf{A}[t],\mathbb{K}[t]\}_{t \geq 1}}{\text{Maximize}} \quad \sum_f w_f \underline{\lambda}_f$

subject to
$$K^f[t] \leq A_f[t], \quad \forall f, \forall t \geq 1, \quad (2)$$
$$\overline{\lambda}_f(1 - q_f) \leq \underline{\mu}_f, \quad \forall f, \quad (3)$$
$$\mathbf{M}[t] \sim F_{g(\mathbf{Z}(\mathbb{K}[t]))}(m), \ \forall t \geq 1, \quad (4)$$

where (2) assures that the coding is limited to packets available in that frame, (3) assures that long-term delivery ratio requirements are satisfied, and (4) indicates that the successful packet transmissions are random as a function of the completion-time distributions and coding decisions.

We emphasize that solving SCP through standard control theoretic methods is extremely difficult, as it is an infinite-horizon problem. Also it contains instantaneous constraints (2), randomness described by convoluted distributions (4) and long-term average requirements (3). Here, we propose a combination of *dynamic pricing* and *finite-horizon dynamic programming* strategies operating in different timescales that yields an implementable solution to this complex problem.

III. Algorithm Description and Analysis

In this section, we propose a practical stochastic algorithm for solving SCP (see Definition 1). Essentially, the algorithm introduces and maintains a time-varying price vector $\mathbf{X}[t] = (X_f[t])_f$, where $X_f[t]$ measures the experienced delivery ratio requirement violation for Flow f by Frame t. Both the rate controller and the dynamic coding strategy reacts to $\mathbf{X}[t]$ to determine the amount of new injections $\mathbf{A}[t]$ and the coding window composition $\mathbb{K}[t]$ in Frame t.

Definition 2 (Joint Rate Control and Coding Algorithm): For a given set $\mathcal{F}$ of τ-deadline-constrained flows, their relative weights $\mathbf{w} = (w_f)$, and their long-term delivery ratio requirements $\mathbf{q} = (q_f)$, the joint dynamic algorithm performs the following operations:
▶ **Price Update**: We maintain a price variable $\mathbf{X}[t] = (X_f[t])_f$, where $\{X_f[t]\}_t$ for each f is initiated at $X_f[0] = 0$

and is updated at each frame according to:

$$X_f[t] = (X_f[t-1] + \beta(D_f[t-1] - q_f A_f[t-1]))^+, \quad (5)$$

where $(y)^+ = max(0, y)$, and $\beta > 0$ is a small stepsize parameter. We recall that $A_f[t-1]$ and $D_f[t-1]$ denotes the number of generated and dropped Flow-f packets in Frame $(t-1)$, and hence are known at the beginning of Frame t.

▶ **Rate Controller**: Given $\mathbf{X}[t]$, the rate controller updates the vector $\lambda[t] = (\lambda_f[t])_f$ of mean number of packets to be generated in Frame t as:

$$\lambda_f[t] = (\lambda_f[t-1] + \alpha(w_f - X_f[t](1 - q_f)))^+, \quad \forall f, \quad (6)$$

where $\alpha > 0$ is a small parameter. Then, a random vector, $\mathbf{A}[t]$, of packets are generated with the computed mean vector $\lambda[t]$ and with a finite variance[2].

▶ **Dynamic Coding Strategy**: The coding strategy is based on the following *Finite-Horizon Dynamic Programming* construction: for any non-negative valued price vector $\mathbf{X}$, we define the optimal reward-to-go function $J_{\mathbf{X}}^\star(\mathbf{B}, s)$ as the maximum value of the $\mathbf{X}$−weighted total mean success rates when there is a vector of $\mathbf{B} = (B_f)_f$ packets waiting for transmission and when there are $s \in \{0, \cdots, \tau\}$ slots left until the end of the frame. Then, $J_{\mathbf{X}}^\star(\mathbf{B}, s)$ satisfies Bellman's equation ([1]):

$$J_{\mathbf{X}}^\star(\mathbf{B}, s) = \max_{\{\mathbf{K}_1 : K_{f,1} \le B_f, \forall f\}} \{\mathbb{E}[J_{\mathbf{X}}^\star((\mathbf{B} - \mathbf{K}_1), (s - Z(K_1)))$$
$$+ (\sum_f K_{f,1} X_f) \cdot \mathcal{I}(Z(K_1) \le s)]\},$$

where $\mathcal{I}(B)$ is the indicator function of event B. This is solved through backward recursion with the initial conditions: $J_{\mathbf{X}}^\star(\mathbf{B}, s) = 0$, for all $\mathbf{B}$ and all $s \le 0$.

Recall that $\mathbf{A}_i[t] = (A_{f,i}[t])_f$ and τ_i respectively denote the vector of remaining packets and the number of remaining time slots in Frame t at the beginning of the ith block construction (see Figure 2). Then, the ith block of Frame t for $i = 1, 2, \cdots$ is selected as follows until the frame ends:

$$\mathbf{K}_i[t] = \underset{\{\hat{\mathbf{K}}_i \ge 0 : \hat{K}_{f,i} \le A_{f,i}[t], \forall f\}}{\arg\max} \mathbb{E}[J_{\mathbf{X}[t]}^\star(\mathbf{A}_i[t] - \hat{\mathbf{K}}_i, \tau_i - Z(\hat{K}_i))$$
$$+ (\sum_f \hat{K}_{f,i} X_f[t]) \cdot \mathcal{I}(Z(\hat{K}_i) \le \tau_i)].$$

We note that the price update rule (5) uses an *instantaneous* measure of the long-term fractional packet drop requirement to change $X_f[t]$ with increments of β. The key observation is that if $\{X_f[t]\}_t$ is guaranteed to be stable in the long run, the long-term delivery ratio requirement of Flow f will be met. Accordingly, the rate controller of (6) discourages packet generation for flows if their scaled price value $X_f[t]$ exceeds their weights. Also, the coding strategy in each frame weighs the successful service rates of flows with their existing prices, $\{X_f[t]\}_f$, effectively prioritizing the service of those flows whose delivery ratio requirement is violated more severely.

We provide a description and an analysis of a heuristic fluid approximation of this stochastic algorithm in [6], where we establish that the fluid approximation is globally asymptotically

stable ([9], [13]). It is well-known (e.g. see [3], [5], [12]) that such fluid analysis provides the foundation for proving the optimality of its stochastic counterpart, which is part of our future work.

IV. APPLICATION: CELLULAR DOWNLINK SCENARIO

In this section, we study an important application of our generic construction whereby a Base Station (BS) served multiple flows by broadcasting their packets to N receivers over time varying channels (see Figure 3), where each packet is a vector of length l over a finite field $\mathbb{F}_d$, for some $d \in \mathcal{Z}_+$.
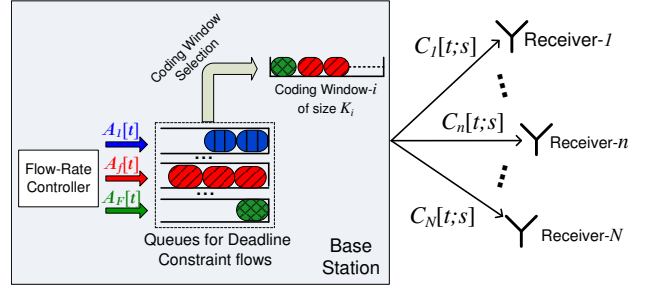

Fig. 3. Cellular downlink model for broadcasting deadline-constrained flows to N users over erasure channels.

We assume that the channel conditions of each receiver independently vary in each time slot. Specifically, $C_n[t; s]$ which denotes the state of the nth receiver's channel in the sth time slot of Frame t, is an independent Bernoulli process with $c_n = P(C_n[t; s] = 1) = 1 - P(C_n[t; s])$. If $C_n[t; s] = 1$ in a time slot, then the packet transmitted by the BS is assumed to be successfully received by Receiver-n. Yet, the BS is not allowed to know $(C_n[t; s])_n$ but only knows $(c_n)_n$.

It is easy to check that this particular system agrees with the descriptions provided in Section II-A except that we can now express the completion time $Z(K)$ more precisely for different coding strategies. We use two key strategies, namely *Randomized Broadcast Coding (RBC)* and *Round Robin Scheduling (RRS)*, that possess attractive optimality characteristics (see [4]) and lends themselves to tractable analysis. RBC is a network coding strategy over a block of K packets where in a slot, say s, any linear combination (over the finite field) of the K packets in the block can be transmitted. Specifically, if $\mathbf{P}(s)$ denote the packet chosen for transmission in slot s, we have $\mathbf{P}[s] = \sum_{k=1}^K \alpha_k[s] P_k$, where $\{\alpha_k[s]\}_k$ are chosen uniformly at random from the field $\mathbb{F}_d \backslash \{0\}$ for every time slot $[s]$. RRS is a scheduling strategy where Packet-k for $\{k = 1, \ldots, K\}$ is transmitted in time-slots $(rK + k)$ for $r = 0, 1, \ldots$ until all the receivers receive the whole block.

We provide the derivation of completion time distributions under RBC and RRS in [6], and instead present several simulation results that study the performance of our joint algorithm under different requirements and coding strategies. We make a reasonable assumption that the channels between the base station and each of the users are independent and identical in distribution. We let the arrival distribution to be uniform with domain 0 to $2\lambda_f$ and, hence, with mean λ_f. The delivery ratio metric is kept symmetric across the flows for simplicity of exposition.

---

[2]For example, among many others, $A_f[t]$ may be uniform, two-point Bernoulli, or Poisson distributed with the updated mean $\lambda_f[t]$.

Fixed vs. Dynamic Coding: Figure 4 compares the developed dynamic coding algorithm to a strategy where the coding window size is kept constant.The figure indicates that the fixed coding window of 3 performs the best among the various fixed coding window sizes. The dynamic algorithm also chooses $K = 3$ most frequently (92.96%) but also dynamically adapts its window size, when the time to deadline is decreasing, thereby outperforming any fixed window strategy.
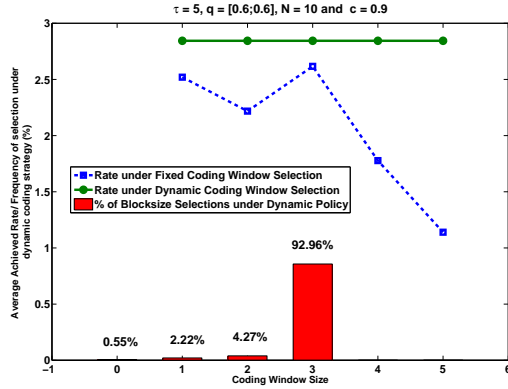

Fig. 4.　Fixed Coding Window vs Dynamic Coding Window

Effect of the Transmission Strategy and N: In Figure 5 we plot the achieved rates for two flows with variations in the number of receivers in the cellular downlink system and the transmission strategies. It can be seen that the achieved rates are lower when the number of receivers is increased and also when the strategy is changed from RBC to RRS. Both these can be attributed to the larger expected completion times.
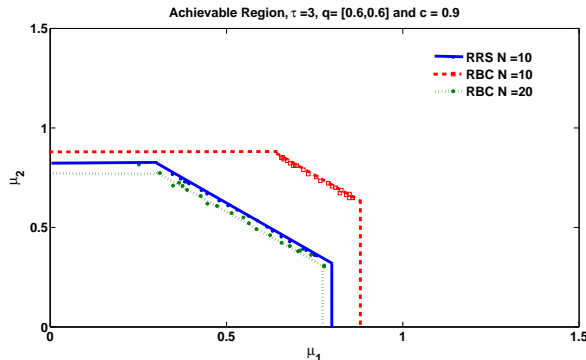

Fig. 5.　Achieved Rates for varied number of receivers and transmission strategies

Effect of Reliability Metric, q: The Figure 6 illustrates the failure of strategies to serve the flows with very high demands ($q_f < 0.55$), while providing the maximum achievable rate to the lower demand flows ($q_f > 0.9$). The lower expected completion time of the RBC strategy allows service of flows with more stringent delivery ratio requirement ($0.55 \geq q_f$) at a higher rate than the RRS strategy ($0.75 \geq q_f$). The comparison for asymmetric reliabilities is provided in [6].

V. CONCLUSION

In this paper we have presented an optimal rate control and coding operation for deadline constrained traffic over time varying wireless channels. A model was developed for a general communication system 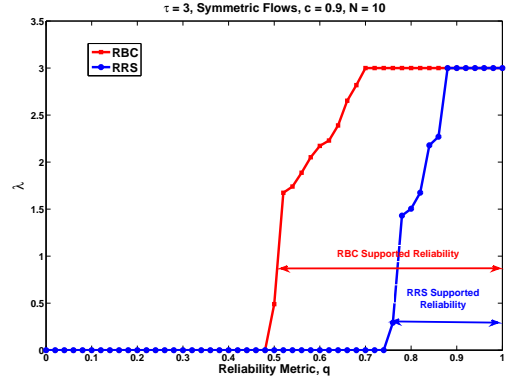with heterogeneous delivery ratio requirements and priorities of different deadline constrained flows. A novel algorithm operating at different timescales was presented that combines dynamic pricing to determine the incoming flow rates and finite horizon dynamic programming for coding operations. The developed generic algorithm is also applied to the important scenario of cellular broadcast over erasure channels with random network coding. Simulations corroborate our results and also show the effect of various system parameters and application requirements.

Fig. 6.　Achieved Rates for RBC and RRS Strategies

REFERENCES

[1] Dimitri P. Bertsekas. *Dynamic Programming and Optimal Control*. Athena Scientific, 1995.

[2] P.P. Bhattacharya and A. Ephremides. Optimal scheduling with strict deadlines. *Automatic Control, IEEE Transactions on*, 34(7):721–728, Jul 1989.

[3] J. G. Dai and S. P. Meyn. Convergence and moments of networks and their fluid models. *IEEE Transactions on Automatic Control*, 40:1889–1904, 1995.

[4] A. Eryilmaz, A. Ozdaglar, and M. Medard. On delay performance gains from network coding. In *Information Sciences and Systems, 2006 40th Annual Conference on*, pages 864–870, March 2006.

[5] A. Eryilmaz and R. Srikant. Fair resource allocation in wireless networks using queue-length based scheduling and congestion control. In *Proceedings of IEEE Infocom*, volume 3, pages 1794–1803, Miami, FL, March 2005.

[6] H.Gangammanavar and Atilla Eryilmaz. Optimal coding and scheduling techniques for broadcasting deadline constraint traffic over unreliable wireless channels. Technical report, The Ohio State University, 2010.

[7] I. H. Hou, V. Borkar, and P. R. Kumar. A theory of qos for wireless. *INFOCOM 2009, IEEE*, pages 486–494, June 2009.

[8] J. J. Jaramillo and R. Srikant. Optimal Scheduling for Fair Resource Allocation in Ad Hoc Networks with Elastic and Inelastic Traffic. *ArXiv e-prints*, July 2009.

[9] H. Khalil. *Nonlinear Systems*. 3rd ed., Prentice Hall, Upper Saddle River, NJ, 2002.

[10] D.G. Pandelis and D. Teneketzis. Stochastic scheduling in priority queues with strict deadlines. *Advances in Applied Probability*, 26:258–279, 1994.

[11] Dinesh Rajan, Ashutosh Sabharwal, and Behnaam Aazhang. Delay-bounded packet scheduling of bursty traffic over wireless channels. *IEEE Trans. Info. Theory*, 50:125–144, 2004.

[12] Srinivas Shakkottai and R. Srikant. Network optimization and control. *Found. Trends Netw.*, 2(3):271–379, 2007.

[13] R. Srikant. *The Mathematics of Internet Congestion Control*. Birkhäuser, Boston, MA, 2004.

[14] Raghava N. Swamy and Tara Javidi. Optimal code length for bursty sources with deadlines. In *ISIT'09: Proceedings of the 2009 IEEE international conference on Symposium on Information Theory*, pages 2694–2698, Piscataway, NJ, USA, 2009. IEEE Press.

[15] Z.-X. Zhao, S.S. Panwar, and D. Towsley. Queueing performance with impatient customers. In *INFOCOM '91. Proceedings. Tenth Annual Joint Conference of the IEEE Computer and Communications Societies. Networking in the 90s., IEEE*, pages 400–409 vol.1, Apr 1991.