

## Content Caching and Scheduling in Wireless Networks With Elastic and Inelastic Traffic

### Abstract:

The rapid growth of wireless content access implies the need for content placement and scheduling at wireless base stations. We study a system under which users are divided into clusters based on their channel conditions, and their requests are represented by different queues at logical front ends. Requests might be elastic (implying no hard delay constraint) or inelastic (requiring that a delay target be met). Correspondingly, we have request queues that indicate the number of elastic requests, and deficit queues that indicate the deficit in inelastic service. Caches are of finite size and can be refreshed periodically from a media vault. We consider two cost models that correspond to inelastic requests for streaming stored content and real-time streaming of events, respectively. We design provably optimal policies that stabilize the request queues (hence ensuring finite delays) and reduce average deficit to zero [hence ensuring that the quality-of-service (QoS) target is met] at small cost. We illustrate our approach through simulations.