Mathematical Modelling of Computer Networks: Part II

Module 2: Wireless Scheduling

Lecture 3: Large Deviations and Wireless Scheduling
3rd December 2013

Laila Daniel and Krishnan Narayanan
Dept. of Computer Science, University of Helsinki, Finland

Large Deviations and Wireless Scheduling - Today's lecture

- What is 'Large Deviation'
- Cramer theorem for empirical averages
- Basic results of Ying et al paper on Large Deviations and Wireless Scheduling
- Large Deviations optimality and Lyapunov drift minimization
- Summary

What is Large Deviation? - (1/3)

- $X_1, X_2, \dots$ iid random variables
- $S_n = X_1 + X_2, \dots + X_n$, partial sums
- $\mu n\, E[X_1]$ and $\sigma^2 = Var(X_1)$
- Strong Law of Large Numbers (SLLN)

$$lim_{n\to\infty} \frac{1}{n} S_n \xrightarrow{a.s} \mu$$

- Central Limit Theorem (CLT)

$$lim_{n\to\infty} \frac{S_n - nE[X_1]}{\sigma\sqrt{n}} \xrightarrow{Dist} \mathbb{Z}, \quad \mathbb{Z} \sim N(0,1)$$

- CLT quantifies the probability that S_n differs from $nE[X_1]$ by an amount of order $\sqrt{n}$
- Deviations of order (size) $\sqrt{n}$ are called 'Normal'

What is Large Deviation? - (2/3)

- In 'Large Deviations consider events where S_n differs from $nE[X_1]$ by an amount of order n
- Deviations of this size are called 'large'
- E.g. Event $\{S_n - \mu n \geq an\}, \quad a > 0$
- The probability of this event tends to zero as $n \to \infty$
- Such events are called 'rare events'
- we want to quantify the rate at which this occurs

What is Large Deviation? - (3/3)

- The decay rate is exponential in n.:
- (Under certain conditions on the tail distribution of n:)

$$lim_{n \to \infty} - \frac{1}{n} log \mathbb{P}(S_n - \mu n \geq an) = I(a), \quad a > 0$$

- Rate function I(.) is the map that maps a into I(a): $a \mapsto I(a)$
- Knowledge of the rate function is crucial for evaluation of integrals of exponential functionals of S_n as $n \to \infty$

Cramer Theorem for empirical averages -(1/2)

- First basic result of Large Deviation theory
- Cramer Theorem:
- Let X_1, X_2, ... iid random variables
- Let $\phi(t) = E[e^{tX_1}] < \infty, \forall t \in \mathbb{R}$
- Let $S_n = \sum_{i=1}^{n} X_i$
- Then for all $a > E[X_1]$,

$$lim_{n \to \infty} - \frac{1}{n} log \mathbb{P}(S_n > a_n) = I(a)$$

Here $I(z) = sup_{t \in \mathbb{R}}[Zt - M(t)]$ where $M(t) = log(\phi(t))$

Cramer Theorem for empirical averages -(2/2)

- $M(t)$ can be recovered from $I(z)$ as follows

  $M(t) = sup_{z \in \mathbb{R}}[tZ - I(z)]$

  The above equation for $M(t)$ is the dual of the formula for I(Z)

  $I(z) = sup_{t \in \mathbb{R}}[Zt - M(t)]$ where $M(t) = log(\phi(t))$
- This is an instance of Legender-Fenchel duality
- So ϕ and I are in a 1-1 relation in that each is determined by the other

Large Deviation analysis of Wireless Scheduling - (1/3)

- These results are based on Ying et al paper
- Setting: A cellular network consisting of abase station and n receivers
- The channel states of the receivers are assumed to be independent of each other
- A multi-state channel model is assumed where each channel is assumed to be in one of L states
- Assume that the time is slotted and each of the channel states of the receivers at each time slots are known to the base station
- Then the base station can decide which queue to serve according to the channel states
- Here it is assumed that the base station operates in a TDMA fashion, i.e., the base station can serve only one queue in each time slot

Large Deviation analysis of Wireless Scheduling - (2/3)

- The goal is to compare two different scheduling policies:
- Queue Length Based (QLB) policy
- Greedy (G) policy
- under the given constraint stated as upper bound of the queue overflow probability (QOP) or delay violation probability (DVP)
- QOP and DVP can be thought of Quality of Service (QoS) constraints
- Note that throughput optimality (Maximum network throughput) is not always the best criterion as it may come at the price of excessive delay for some users

Large Deviation analysis of Wireless Scheduling - (3/3)

- The following results are obtained in this paper based on Large Deviation analysis
- Result 1: The total network throughput under QLD policy is no less than the throughput of the Greedy policy for all N where N is the number of receivers
- Result 2: This result deals with the lower bound on the throughput of the QLB policy
- For sufficiently large N, the lower bound of the throughput under QLB is shown to be tight, strictly increasing with N and strictly larger than the throughput of the Greedy policy
- For the simple multi-state channel model, on-off-model, the lower bound obtained is tight for all N

Basic Model - (1/3)

- QLB policy
- Choose user i^* to transmit if

$$i^* = \in arg\ max_i \gamma_i[t] Q_i[t]$$

where $_i[t]$ is the queue length of user i in slot t and $\gamma_i[t]$ is the channel state process of user i at slot t

- Greedy policy
- Choose user i^* to transmit if

$$i^* = \in arg\ max_i \gamma_i[t]$$

- Queue overflow Constraint

$$Pr(max_i Q_i(0) > B) \leq \epsilon$$

where $Q_i(0)$ is the stationary queue length

Basic Model - (2/3)

- Instead of of studying this constraint as above we study the approximation to the constraint given by

$$\theta_B(N, A) := lim_{n \to \infty} - \frac{1}{B} log Pr(max_i Q_i(0) > B) \leq \delta$$

  where the large deviations exponent θ_B is a function of number of users and the total arrival rate.
- The exponent δ can be related to ϵ for large B using the approximation $e^{-\delta B} = \epsilon$
- Notice the crucial element of *scale invariance* (size of the buffer B) character of the quantity $\theta_B(N, A)$

Basic Model - (3/3)

- Let D(t) to be the maximum delay experienced so far by any bit in any of the queues in slot t, and D(0) to be the stationary maximum delay.
- The steady state Delay violation constraint D(t) can be expressed as

$$Pr(D(0) > D) \leq \epsilon$$

- Since the arrival rate is constant, it is easily seen that

$$Pr(D(0) > D) = Pr(max_i Q_i(0) > \frac{\lambda}{N} D)$$

- The delay violation constraint can be expressed as

$$Pr(max_i Q_i(0) > \frac{\lambda}{N} D) \leq \epsilon$$

$$\theta_D(N, \lambda) := lim_{D \to \infty} - \frac{1}{D} log\, Pr(max_i Q_i(0) > \frac{\lambda D}{N}) \geq \delta$$

$$\theta_D(N, \lambda) = \frac{\lambda \theta_B(N, \lambda)}{N}$$

- Thus we will primarily consider the queue overflow problem when we analyze the wireless systems using large deviations