Non-convex Optimization and Resource Allocation in Communication Networks

Ravi R. Mazumdar
Professor of ECE
Purdue University, IN, USA

Motivation

- In communication networks
  - Scarce resource
  - Services with diverse QoS requirements
  - Dynamic time-varying environment
- Resource allocation need
  - Allocate resource efficiently
  - Treat services with diverse QoS requirements in a unified way
  - Adapt to dynamic environment
- A promising solution: Utility (and pricing) framework

Utility (and pricing) framework

- Utility
  - Degree of user's satisfaction by acquiring a certain amount of resource
  - Different QoS requirements can be represented with different utility function
- Price
  - Cost for resource
  - Device to control user's behavior to achieve the desired system purpose

Our work

- Allow non-concave utility function
  - Non-convex optimization problem
- Simple (and distributed) algorithm
- Asymptotic optimal resource allocation
- Downlink power allocation in CDMA networks
- Rate allocation in the Internet

Basic problem

$$\max \sum_{i=1}^{N} U_i(x_i)$$

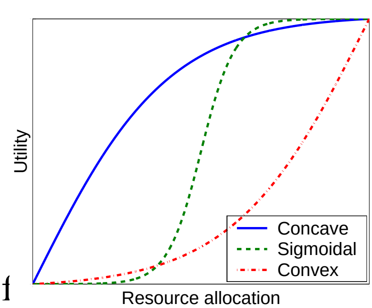
$$s.t. \quad \sum_{i=1}^{N} x_i \le C$$

$$0 \le x_i \le M_i, \; i = 1,2,..N$$

- N: Number of users in the system
- U_i: Utility function of user i
- x_i: The amount of resource that is allocated to user i
- M_i: The maximum amount of resource that can be allocated to user i
- C: Capacity of resource

Non-convexity in resource allocation

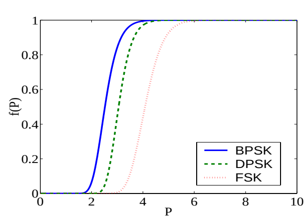
- If all utility functions are concave functions
  - Convex programming
  - Can be solved easily using KKT conditions or duality theorem
- But, in general, three types of utility functions



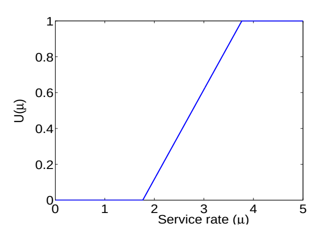
Non-convexity in resource allocation

- Concave function: traditional data services in the Internet
- "S" function: real-time services in the Internet and some services in wireless networks
- Convex function: some services in wireless networks
- Cannot be formulated as a convex programming
  - Cannot use KKT conditions and duality theorem
  - Need a complex algorithm for a global optimum

Examples of non-concave utility function



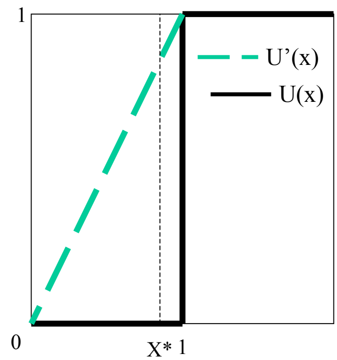
Example of non-concave utility function for wireless systems: packet transmission success probability



Example for non-concave utility function for wireline systems: the ratio of the highest arrival rate that makes the probability that delay of a packet is greater than T be less than P_{th} to the maximum arrival rate in M/M/1 queue

Inefficiency of the naïve approach

- 11 users and a resource with capacity 10
- Each user has a utility function $U(x)$
- Allocate 1 to 10 users and 0 to one user
  - 10 total system utility
- Concave hull $U'(x)$
- With $U'(x)$, each user is allocated x* = 10/11 and $U'(x^*) = 10/11$
- But $U(x^*) = 0$
  - Zero totally system utility



Need resource allocation algorithm taking into account the properties of non-concave functions

Basic solution

- Define $$L(\bar{x},\lambda) = \sum_{i=1}^{N} U_i(x_i) + \lambda(C - \sum_{i=1}^{N} x_i)$$
- For a given $\lambda \geq 0$, $\bar{x}(\lambda)$ that maximizes $L(\bar{x},\lambda)$ is a global optimal solution of

$$\max \sum_{i=1}^{N} U_i(x_i)$$

$$s.t. \quad \sum_{i=1}^{N} x_i \leq \sum_{i=1}^{N} x_i(\lambda)$$

$$0 \leq x_i \leq M_i, \; i = 1,2,\ldots N$$

Basic solution

- If $\sum_{i=1}^{N} x_i(\lambda^*) = C$, $\overline{x}(\lambda^*)$ is a global optimal solution of the basic problem
- But, due to the non-concavity of the utility function, there may not exist such a λ^*
- If $\left| C - \sum_{i=1}^{N} x_i(\lambda^*) \right|$ is minimized, $\overline{x}(\lambda^*)$ could be a good approximation of the global optimal solution
- We will try to find such a λ^*

Basic solution

- $\overline{x}(\lambda)$ maximizes $L(\overline{x}, \lambda)$, if

$$x_i(\lambda) = \arg \max_{0 \le x \le M_i} \{U_i(x) - \lambda x\}, \; i = 1,2,\ldots, N$$

- If we interpret λ as price per unit resource, $x_i(\lambda)$ is the amount of resource that maximizes user i's net utility
- Define $\lambda_i^{\max}$ for each user i

$$\lambda_i^{\max} = \min \left\{ 0 \le \lambda \le \infty \mid \max_{0 \le x \le M_i} \{U_i(x) - \lambda x\} = 0 \right\}$$

- Each user has a unique $\lambda_i^{\max}$

Basic solution

- We call $\lambda_i^{\max}$ the maximum willingness to pay of user i, since if $\lambda > \lambda_i^{\max}$, $x_i(\lambda) = 0$
- If U_i is convex or "s" function, $x_i(\lambda)$ is discontinuous at $\lambda_i^{\max}$
- Due to this property, we may not find a λ^* such that

$$\sum_{i=1}^{N} x_i(\lambda^*) = C$$

- For the simplicity, we assume that each user has a different maximum willingness to pay
- Further, assume that $\lambda_1^{\max} > \lambda_2^{\max} > \ldots > \lambda_N^{\max}$

Basic solution

- To minimize $\left| C - \sum_{i=1}^{N} x_i(\lambda) \right|$, while satisfying $\sum_{i=1}^{N} x_i(\lambda) \le C$,

allocate resource to users from 1 to K

$$K = \max_{1 \le j \le N} \left\{ \sum_{i=1}^{j} x_i(\lambda_j^{\max}) \le C \right\} \qquad (1)$$

- Hence, resource is allocated to users in a decreasing order of their maximum willingness to pay

Basic solution

- For the selected users, we can easily find a λ^* such that

$$\sum_{i=1}^{K} x_i(\lambda^*) = C, \qquad (2)$$

since the problem for the selected users is reduced convex programming

- Hence, resource is allocated to each user according to

$$\bar{x}(\lambda^*) = (x_1(\lambda^*), \ldots x_K(\lambda^*), 0, \ldots 0)$$

Basic solution

- We can show that
  - If

$$\sum_{i=1}^{K} x_i(\lambda_{K+1}^{\max}) < C \text{ and } \sum_{i=1}^{K+1} x_i(\lambda_{K+1}^{\max}) > C, \qquad (3)$$

    the proposed resource allocation may not be a global optimum
  - Otherwise, it is a global optimum
- Hence, the proposed resource allocation may not be a global optimum

Basic solution

- However,

$$\sum_{i=1}^{N} U_i(x_i^o) - \sum_{i=1}^{K} U_i(x_i(\lambda^*)) \le \max_{1\le i\le N} \{U_i(M_i)\},$$

where $(x_1^o, x_2^o, \ldots, x_N^o)$ is a global optimal allocation

- Hence,

$$\frac{\sum_{i=1}^{K} U_i(x_i(\lambda^*))}{\sum_{i=1}^{N} U_i(x_i^o)} \to 1 \text{ as } K \to \infty \ (N \to \infty)$$

- This implies that the proposed resource allocation is asymptotically optimal

Algorithm for wireless system

- If x_i is power allocation for user i,
- C is the total transmission power of the base station, and
- $C = M_i$ for all i,
- The basic problem is equivalent to the downlink power allocation problem for a single cell with total transmission power C
- The joint power and rate allocation problem for the downlink that maximizes the expected system throughput can be formulated by using the basic problem

Algorithm for wireless system

- The base station can be a central controller that can
  - collect information for all users
  - select users and allocates power to the selected users
- The base station selects users according to Equation (1)
- The base station allocates power to the selected users according to Equation (2)

Algorithm for the Internet

- If x_i is allocated rate for user i,
- C is the capacity of the link, and
- M_i is the maximum rate that can be allocated to user i,
- The basic problem is equivalent to the rate allocation problem in the Internet with a single bottle-neck link
- However, in the Internet, no central controller
- Need a distributed algorithm

Algorithm for the Internet

- The following problems are solved iteratively by each user and the node
  - User problem for user *i*

$$x_i(\lambda^{(n)}) = \arg\max_{0 \le x \le M_i} \{U_i(x) - \lambda^{(n)} x\}$$

    - Each user *i* determines its transmission rate that maximizes its net utility with price $\lambda^{(n)}$
  - Node problem

$$\lambda^{(n+1)} = \left[\lambda^{(n)} - \alpha^{(n)}(C - \sum_{i=1}^{N} x_i(\lambda^{(n)}))\right]^+$$

    - Node determines the price for the next iteration $\lambda^{(n+1)}$ according to the aggregate transmission rate and delivers it to each user

Algorithm for the Internet

- $C - \sum_{i=1}^{N} x_i(\lambda^{(n)})$ is a subgradient of the dual of the basic problem
- Dual problem
  - Convex programming
  - Non-differentiable due to the non-convexity of the basic problem
- Hence, the algorithm solves the dual problem by using subgradient projection
- By taking

$$\alpha^{(n)} \to 0 \text{ and } \sum_{n=0}^{\infty} \alpha^{(n)} = \infty,$$

- $\lambda^{(n)}$ converges to the dual optimal solution λ^{o}

Algorithm for the Internet

- At the dual optimal solution λ^o

$$\sum_{i=1}^{N} x_i(\lambda^o) = C,$$

  i.e., global optimal rate allocation can be obtained, except when Equation (3) is satisfied
- In this case, $\lambda^o = \lambda^{K+1}$ and, thus, $\lambda^{(n)} \to \lambda^{K+1}$
- Hence, by Equation (3), the aggregate transmission rate oscillates between feasible and infeasible solutions causing congestion within the node
- To resolve this situation, we will use "self-regulating" property of users

Algorithm for the Internet

- *We call the property of a user that it does not transmit data even though the price is less than its maximum willingness to pay, if it realizes that it will receive non-positive net utility the "self-regulating" property*
- Assume that each user has the "self-regulating" property

Algorithm for the Internet

- Further assume that the node allocate rate to each user as

$$x'(\lambda) = \begin{cases} x_i(\lambda), \text{ if } \sum_{i=1}^{N} x_i(\lambda) \le C \\ f_i(\overline{x}(\lambda)), \text{ otherwise} \end{cases}$$

  - $x_i(\lambda)$ is transmission rate of user i at price λ
  - f_i is a continuous function such that

$$f_i(\overline{x}) < x_i \text{ and } \sum_{j=1}^{N} f_j(\overline{x}) = C$$

- A good candidate for f_i is $f_i(\overline{x}) = \dfrac{x_i}{\sum_{j=1}^{N} x_j} C$

Algorithm for the Internet

- Then, if Equation (3) is satisfied, there exists an iteration m_i for each i, $i = K+1, \ldots, N$ such that

$$U_i(x_i'(\lambda^{(n)})) - \lambda^{(n)} x_i'(\lambda^{(n)}) \le 0, \text{ for all } n \ge m_i, i = K+1, \ldots, N$$

- By "self-regulating" property, users from K+1 to N stop transmitting data
  - This is equivalent to selecting users from 1 to K as in Equation (1)
- After that, rate allocation for users from 1 to K converges to rate allocation that satisfies Equation (2)
- The proposed asymptotically optimal solution can be obtained with the "self-regulating" property

Summary

- Non-convex resource allocation problem allowing non-concave utility functions
- Simple solution that provides an asymptotical optimum
- The same problem and solution can be applied to both wireless and wireline systems
- However, for an efficient and feasible algorithm in each system
  - Must take into account a unique property of each system
  - Results in a different algorithm for each system even though two algorithms provide the same solution