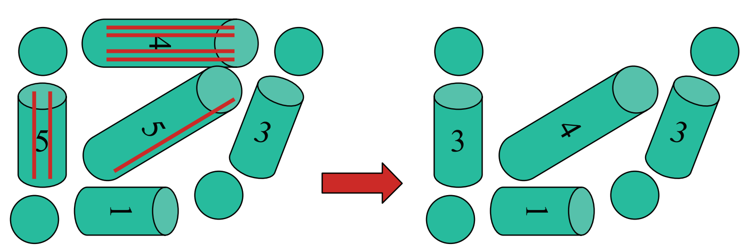
Outline

- Optimal routing based on flows
- First derivative length
- Optimal routing characterization
- Flow deviation: Frank-Wolfe

Optimal routing based on flows

- For path selection, we have considered minimum cost when we have a given price is paid per path
- What happens when we add capacity considerations?
- If we have circuits or virtual circuits, then we can create a topology that takes into account the presence of other users



- If we have a probabilistic description of traffic progression, then we could use dynamic programming

Routing based on flows

- When we have packets or fine granularity virtual circuits (VCs), we can assume that we have roughly fluid flows
- We will try to optimize a cost function related to the flow F(i,j) (arrival rate in terms of of what we may be considering) on the link (i,j)
- A reasonable metric is $\sum_{(i,j)} D(F(i,j))$ where D is some monotonically increasing function that grows very sharply when F(i,j) approaches the link capacity
- These type of models are called flow models - they look only at mean flow, they do not consider higher moments

Example based on Kleinrock's approximation

$$\lambda(i,j) = \sum_{\text{all paths p traversing link (i,j)}} \lambda(p)$$

$$\mu(i,j) = \text{service rate on link (i, j)}$$

$$N(i,j) = \text{average number of packets on link (i, j)}$$

- Assume all queues behave like M/M/1 with arrival rate λ(i,j), service rate μ(i,j), and service/propagation delay d(i,j)
- Then

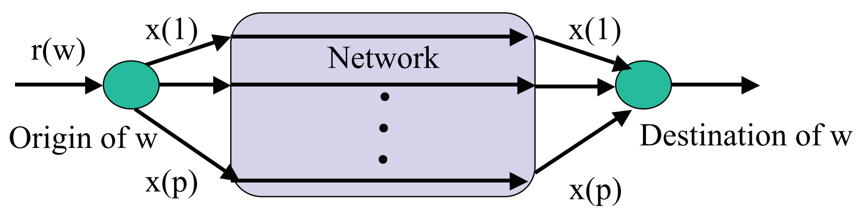
$$N_{i,j} = \frac{\lambda_{i,j}}{\mu_{i,j} - \lambda_{i,j}} + \lambda_{i,j} d_{i,j}$$

Increases sharply when we approach link capacity

$$\Rightarrow \quad D(F(i,j)) = \frac{F(i,j)}{C(i,j) - F(i,j)} + d(i,j)\,F(i,j)$$

Optimal flow routing problem

- For every OD pair w=(i,j), we are given the arrival rate r(w)
- We denote:
  - W: the set of all OD pairs w
  - P(w): a set of paths available for routing for OD pair w (possibly all paths)
  - x(p): the flow of path p (in units per second)



MIT

Problem statement

- We want to choose the flows so that we minimize

$$\sum_{(i,j)} D(F(i,j))$$

subject to the constraints

$$F(i,j) = \sum_{\text{all paths p containing (i,j)}} x(p)$$

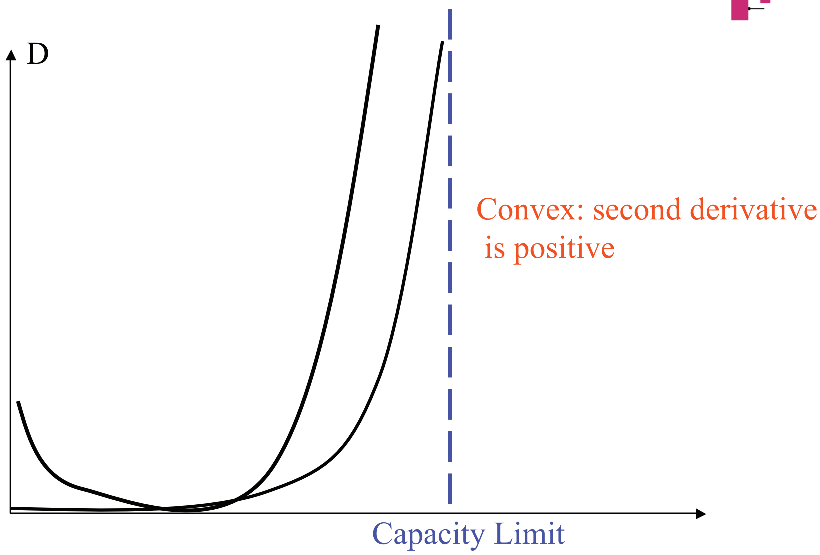
$$\sum_{p \text{ in } P(w)} x(p) = r(w) \text{ for al OD pairs w in W}$$

$$x(p) \geq 0$$

we assume that D is convex and that its first and second derivatives exist

MIT

Examples of D



Flow problem

- We can perform a substitution in the following manner: consider the cost

$$D(x) = \sum_{(i,j)} D\left(\sum_{\text{all paths p containing (i,j)}} x(p) \right)$$

consider the partial derivatives

$$\frac{\partial D(x)}{\partial x(p)} = \sum_{\text{all links (i,j) on path p}} D'\underbrace{\left(\sum_{\text{all paths p containing (i,j)}} x(p) \right)}_{\text{Depends on (i,j) only}}$$

First derivative lengths

- Length interpretation:

Define $D'_{i,j} = D'\left(\sum_{\text{all paths p containing (i, j)}} x(p)\right)$ to be the first derivative length of the link (i, j)

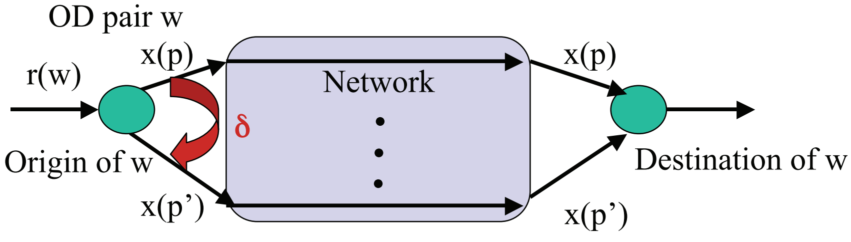
Note : the first derivative length depends on x crucially

The derivative of D with respect to x(p) is thus the sum of the first derivative lengths of the links on the path p

MIT

Characterization of optimal flows

- A set of path flows x is optimal if the cost cannot be improved by making a feasible change of flow
- Consider shifting an increment δ from path p to path p' of the OD pair w



- What is the effect of this transfer on the cost (assuming feasibility is maintained?)

MIT

Change in cost

- Convexity means that a decrease is possible only if there is a local decrease direction - no local minima that are non global
- Let us use the fact that D has first and second derivatives
- A Taylor series expansion approach yields that the difference in cost due to the infinitesimal transfer of flow must be

$$-\frac{\partial D(x)}{\partial x(p)}\delta - \frac{\partial^2 D(x)}{\partial^2 x(p)}\delta^2 + \frac{\partial D(x)}{\partial x(p')}\delta + \frac{\partial^2 D(x)}{\partial^2 x(p')}\delta^2 + o(\delta^3)$$

- To a first order approximation, the change in cost is

$$-\frac{\partial D(x)}{\partial x(p)}\delta + \frac{\partial D(x)}{\partial x(p')}\delta$$

Approximately linear in flow change

MIT

Characterization of optimal flows

- Cost change due to shifting traffic from p to p' must be detrimental in cost in the solution was optimal
- So

$$\frac{\partial D(x)}{\partial x(p)} \le \frac{\partial D(x)}{\partial x(p')}$$

- This condition is the same as rewriting for all OD pairs that for the optimal flow x*,

$$x^*(p) > 0 \text{ only if } \frac{\partial D(x^*)}{\partial x(p)} \le \frac{\partial D(x^*)}{\partial x(p')} \text{ for all } p' \text{ in } P(w)$$

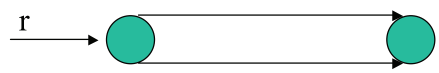
- So optimal path flows are positive only on paths that are shortest with respect to first derivative lengths

MIT

Solutions

- The link lengths depend on the link flows
- The first derivative lengths depend on the unknown optimal path flows, so the problem is harder than a shortest path problem
- Sometimes we can solve the problem using a closed form solution, but generally more involved methods may be necessary

High capacity link C_1



Low capacity link C_2

$$\text{cost } D(x(i)) = \frac{x(i)}{C_i - x(i)}$$

$C_1 \geq C_2$, so two cases are possible

Solution example

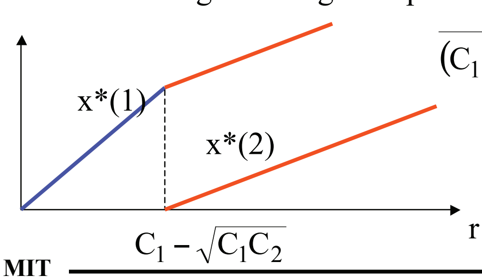
- Case 1: all the flow goes along 1

$$\frac{\partial D(x^*)}{\partial x(p)} \leq \frac{\partial D(x^*)}{\partial x(p')} \text{ for all } p' \text{ in } P(w)$$

$$\text{gives } \frac{C_1}{(C_1 - r)^2} \leq \frac{1}{C_2} \Rightarrow r \leq C_1 - \sqrt{C_1 C_2}$$

- Case 2 : Flow goes along both paths

$$\frac{C_1}{(C_1 - x(1))^2} \leq \frac{C_2}{(C_2 - x(2))^2}$$



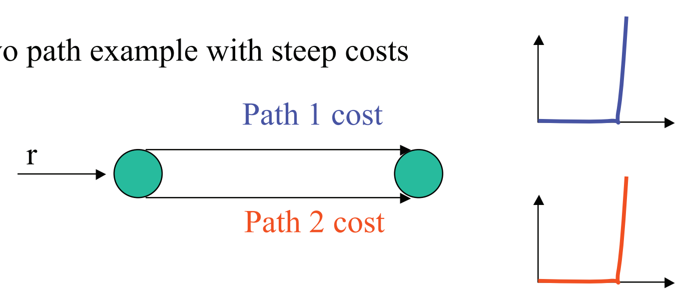
Solution methods

- In general finding solutions may be difficult and different types of methods may need to be applied
- Reduce cost while maintaining feasibility
- A common theme in the methods for improving solutions is to perturb current solution in a way that is feasible and reduces cost
- Why not just use minimum first derivative lengths (MFDLs)?
- Lengths are dependent on flow values x, so search is in general difficult and can lead to instability

MIT

Search problems

- Two path example with steep costs



- Paths alternate having zero cost, flow is thrown from one path to another unless step size is chosen accurately and we are in the right neighborhood

MIT

General flow deviation

- General formulation:

Feasibility: $\sum_{p \text{ in } P(w)} \Delta x(p) = 0$ for all w in W

overall flow conservation

Descent direction: $\sum_{w \text{ in } W} \sum_{p \text{ in } P(w)} \frac{\partial D(x)}{\partial x(p)} \Delta x(p) \leq 0$

cost must have local reduction possibilities if any global reduction opportunities exist

(use first derivative linear cost approximation)

Frank-Wolfe

- An example of flow deviation method
- Linearly mix the current solution with the solution using only minimum first derivative lengths (MFDLs)
- Start at some x(1), x(2), …, x(n) and find MFDL
- Let $\overline{x(1)}, \overline{x(2)}, \ldots, \overline{x(n)}$ be the solution when all the flow is sent along paths with MFDL
- We set convex combination

$x(p) := x(p) + \alpha\, (\overline{x(p)} - x(p))$ for all p in P(w), all w in W

α is selected so that it minimizes the cost for the x(p) assignment above

α is the same for all paths

Problem: very slow in general

Slow convergence of Frank-Wolfe

- Let us revisit the two path example
- Suppose we should be in case 1 (all flow along first path), but that we currently have some small flow along path 2 also
- How do we converge to a solution in which we get rid of the flow along path 2?

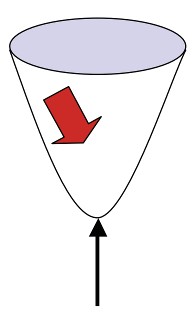
$$\text{MFDL for path 2}: \frac{C_2}{\left(C_2 - x(2)\right)^2}$$

- Suppose that we have very large C_2 - then the step taken by F-W is very small, and the behavior is dominated by C_2 rather than by the changing flow
- While each step yields an improvement, the convergence is very slow

MIT

Steepest descent

$$\text{Gradient of some function } f: \nabla f(x) = \begin{bmatrix} \frac{\partial f(x)}{\partial x(1)} \\ \cdots \\ \frac{\partial f(x)}{\partial x(n)} \end{bmatrix}$$



Descent direction

$$\text{Hessian}: \nabla^2 f(x) = \begin{bmatrix} \frac{\partial^2 f(x)}{\partial x(1)^2} \cdots & \frac{\partial^2 f(x)}{\partial x(1)\partial x(n)} \\ \cdots & \\ \frac{\partial^2 f(x)}{\partial x(n)\partial x(1)} \cdots & \frac{\partial^2 f(x)}{\partial x(n)^2} \end{bmatrix}$$

Optimum point

Convex function: the Hessian is positive semidefinite

MIT

Steepest descent

- Rather than have single parameter for all paths, we use a vector of values (the superscript denotes the step number)

$$\underbrace{x(p)^{k+1} = x(p)^{k} - \alpha\, \nabla D(x^{k+1})}_{\textbf{All vectors}}$$

- An improvement on this method is Newton's method, which also uses the second derivative

$$x(p)^{k+1} = x(p)^{k} - \alpha \left(\nabla^{2} D(x^{k+1})\right) \nabla D(x^{k+1})$$

MIT