Cornell University

Lecture 17 - Networked Infrastructures, Transportation, Communication, Pipelines

Models for real networks

ECE 5260 – ORIE 5735

Anna Scaglione

Cornell Tech

April 6, 2026

Content

Cornell University

What we covered in Module 2 so far

Cornell University

1. **Random Walks**: $\boldsymbol{x}(t)$ is a coordinate vector indicating the current location of the walk, and **Consensus dynamics**: synchronous and asynchronous: $\boldsymbol{x}(t)$ is the state due to linear dynamics
2. **Epidemics**: $x(t)$ is the probability that agents are infected at t. **Social Networks**: $x(t)$ is the opinion held by an agent on a certain fact θ. **Financial Networks**: $x(t)$ is the value or a firm or money flow.
3. **Biological Networks**: We talked about modeling $\mathcal{G}(\mathcal{V}, \mathcal{E})$ nodes and edges from measurements of brain activity or gene-expression, used to find edges.

What we covered in Module 2 so far

Cornell University

1. **Random Walks**: $\boldsymbol{x}(t)$ is a coordinate vector indicating the current location of the walk, and **Consensus dynamics**: synchronous and asynchronous: $\boldsymbol{x}(t)$ is the state due to linear dynamics
2. **Epidemics**: $\boldsymbol{x}(t)$ is the probability that agents are infected at t. **Social Networks**: $x(t)$ is the opinion held by an agent on a certain fact θ. **Financial Networks**: $x(t)$ is the value or a firm or money flow.
3. **Biological Networks**: We talked about modeling $\mathcal{G}(\mathcal{V}, \mathcal{E})$ nodes and edges from measurements of brain activity or gene-expression, used to find edges.

Model abstractions seen..

Cornell University

- Early on we analyzed some general cases:
  1. **Linear and Linear affine dynamics**: continuous, discrete, randomized
  2. **General non-linear dynamics**, where we did a perturbation analysis to understand what to expect locally around stationary points. For non-smooth interactions (polarization via HK model) the analysis is not possible.
  3. **Flow networks**, a special class of non-linear dynamics with positive states.

Cyber-Physical Networked Sysems

These abstractions become especially useful when infrastructures are instrumented, networked, and operated through software.

Outline

Cornell University

Review of affine linear dynamics

Cornell University

Linear affine equations in discrete and ODE form respectively are

$$\boldsymbol{x}(t+1) = \boldsymbol{W}\boldsymbol{x}(t) + \boldsymbol{b}, \quad \dot{\boldsymbol{x}}(t) = \boldsymbol{M}\boldsymbol{x}(t) + \boldsymbol{b}$$

Affine discrete dynamics

If and only if $\boldsymbol{W}$ has all $\lambda_k(\boldsymbol{W}) <1$ the affine discrete dynamics $\boldsymbol{x}$ converge to:

$$\boldsymbol{x}^* = (\boldsymbol{I} - \boldsymbol{W})^{-1}\boldsymbol{b}$$

If $w_{ij} \geq 0$ then $\boldsymbol{W}$ has to be **sub-stochastic**, i.e. $\mathbf{1}^\top \boldsymbol{W} < 1$.

Affine continuous dynamics

If and only if $\boldsymbol{M}$ is Hurwitz if all its eigenvalues have negative real part, i.e. $\mathbb{R}\{\lambda_k\} < 0 \;\; \forall k = 1, N$ the continuous dynamics converge to:

$$\boldsymbol{x}^* = -\boldsymbol{M}^{-1}\boldsymbol{b}$$

If $\boldsymbol{M}$ is metzler and $\boldsymbol{b} \geq 0$ then all $\boldsymbol{x}(t)$ including $\boldsymbol{x}^*$ are non-negative.

Review of positive systems

Cornell University

Definition of Metzler matrices

An $N \times N$ square matrix $\boldsymbol{M}$ is a Metzler matrix if:

1. The off-diagonal elements of $\boldsymbol{M}$ are non-negative, i.e. $[\boldsymbol{M}]_{ij} \geq 0$ for $i \neq j$, the diagonal elements can be negative
2. $\boldsymbol{M}$ has an associated directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with edge weights $[\boldsymbol{M}]_{ij} > 0$ and with no self-loops
3. For any Metzler matrix, there exist a constant $c > 0$ such that $\boldsymbol{M} + c\boldsymbol{I}$ has all non negative eigenvalues

- An example of a Metzler matrix is the complement of the Laplacian $-\boldsymbol{L}$, whose dominant eigevalue is 0

Review of positive systems

Cornell University

Definition of Metzler matrices

An $N \times N$ square matrix $\boldsymbol{M}$ is a Metzler matrix if:

1. The off-diagonal elements of $\boldsymbol{M}$ are non-negative, i.e. $[\boldsymbol{M}]_{ij} \geq 0$ for $i \neq j$, the diagonal elements can be negative
2. $\boldsymbol{M}$ has an associated directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with edge weights $[\boldsymbol{M}]_{ij} > 0$ and with no self-loops
3. For any Metzler matrix, there exist a constant $c > 0$ such that $\boldsymbol{M} + c\boldsymbol{I}$ has all non negative eigenvalues

- An example of a Metzler matrix is the complement of the Laplacian $-\boldsymbol{L}$, whose dominant eigevalue is 0

Review of positive systems

Cornell University

Definition of Metzler matrices

An $N \times N$ square matrix $\boldsymbol{M}$ is a Metzler matrix if:

1. The off-diagonal elements of $\boldsymbol{M}$ are non-negative, i.e. $[\boldsymbol{M}]_{ij} \geq 0$ for $i \neq j$, the diagonal elements can be negative
2. $\boldsymbol{M}$ has an associated directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with edge weights $[\boldsymbol{M}]_{ij} > 0$ and with no self-loops
3. For any Metzler matrix, there exist a constant $c > 0$ such that $\boldsymbol{M} + c\boldsymbol{I}$ has all non negative eigenvalues

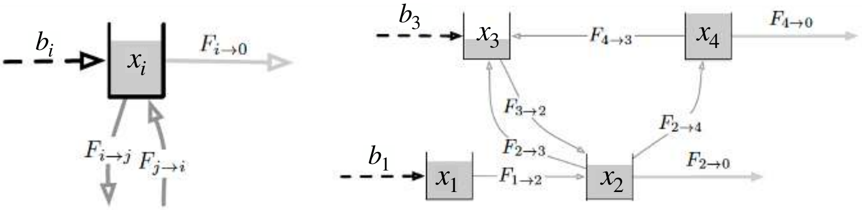
- An example of a Metzler matrix is the complement of the Laplacian $-\boldsymbol{L}$, whose dominant eigevalue is 0

Outline

Cornell University

Dynamical flow systems equations

Cornell University



(a) A compartment with inflow b_i, outflow $F_{i\to 0}$, and inter-compartmental flows $F_{i\to j}$

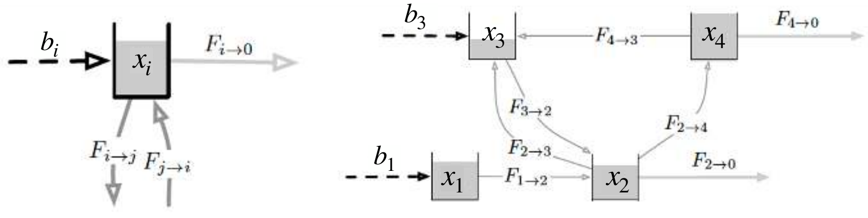
(b) A compartmental system with two inflows and two outflows

- The state of each node is $x_i(t)$; it is replenished by the input $b_i \geq 0$ and depleted by the flow $F_{i\to 0}(x_i)$
- In general, the flow depends on the state of the buffer, i.e. $F_{i\to j} = F_{i\to j}(x_i)$ and the input b_i could also depend on it, but usually it is the outflow. Specifically, flow functions must satisfy:

$$F_{i\to j}(x) \geq 0, \quad F_{i\to j}(0) = 0$$

Dynamical flow systems equations

Cornell University



(a) A compartment with inflow b_i, outflow $F_{i\to 0}$, and inter-compartmental flows $F_{i\to j}$

(b) A compartmental system with two inflows and two outflows

- The state of each node is $x_i(t)$; it is replenished by the input $b_i \geq 0$ and depleted by the flow $F_{i\to 0}(x_i)$
- In general, the flow depends on the state of the buffer, i.e. $F_{i\to j} = F_{i\to j}(x_i)$ and the input b_i could also depend on it, but usually it is the outflow. Specifically, flow functions must satisfy:

$$F_{i\to j}(x) \geq 0, \quad F_{i\to j}(0) = 0$$

Dynamical flow systems equations

Cornell University

- The dynamics of the buffer state are:

$$\dot{x}_i = \overbrace{b_i - F_{i\to 0}(x_i)}^{f(x_i)} + \overbrace{\sum_{j\in\mathcal{N}_i^{\text{in}}} F_{j\to i}(x_j) - \sum_{j\in\mathcal{N}_i^{\text{out}}} F_{i\to j}(x_i)}^{\sum_{j\in\mathcal{N}_i} g(x_i,x_j)}$$

- $[\boldsymbol{f}(\boldsymbol{x})]_e = F_e(x) \geq 0 \;\; \forall e = (i \to j) \in \mathcal{E} \; j \neq 0$, $[\mathbf{f}_{\text{out}}]_i = F_{i\to 0}$ and $[\boldsymbol{b}]_i = b_i$

$$\Rightarrow \quad \dot{\boldsymbol{x}} = \boldsymbol{b} - \mathbf{f}_{\text{out}}(\boldsymbol{x}) + \boldsymbol{B}\boldsymbol{f}(\boldsymbol{x})$$

- The vector $\boldsymbol{b}$ is called the **in-flow** and $\mathbf{f}_{\text{out}}$ the **out-flow**
- The equations are consistent with **conservation of mass**:

$$\mathbf{1}^\top\dot{\boldsymbol{x}} = \mathbf{1}^\top\boldsymbol{b} - \mathbf{1}^\top\mathbf{f}_{\text{out}}(\boldsymbol{x}) + \overbrace{\mathbf{1}^\top\boldsymbol{B}\boldsymbol{f}(\boldsymbol{x})}^{=0}$$

$\mathbf{1}^\top\dot{\boldsymbol{x}} = 0$ if the inflow matches the outflow.

Special case: Linear flow functions

Cornell University

- We note that these are non linear dynamics like the ones we defined in general:
  1. The node state x_i is x_i
  2. The local function $f_i(x)$ is $(b_i - F_{i\to 0}(x))$
  3. The edge function $g_{ij}(x_i, x_j)$ is $(F_{j\to i}(x_j) - F_{i\to j}(x_i))$ is
- A special case is when the functions are such that:

$$F_{i\to j}(x_i) = w_{ji}x_i \propto x_i \quad \text{outflow to node } j \text{ proportional to the state}$$

$$F_{i\to 0}(x_i) = w_{0i}x_i \propto x_i \quad \text{the external outflow is proportional to the state}$$

$$u_i \text{ not a function of the queue state}$$

$$\dot{x}_i = -\Big(w_{0i} + \sum_{j\in\mathcal{N}_i^{\text{out}}} w_{ji}\Big)x_i + \sum_{j\in\mathcal{N}_i^{\text{in}}} w_{ij}x_j + b_i \;\Rightarrow\; \dot{\boldsymbol{x}} = -(\text{diag}(\boldsymbol{w}_o) + \boldsymbol{L})\boldsymbol{x} + \boldsymbol{b}$$

Special case: Linear flow functions

Cornell University

- Introducing:
  1. the weighted adjacency matrix containing all flow rates $[\boldsymbol{A}]_{ij} = w_{ij}, i, j \in \mathcal{V}, \quad i \neq j$ and such that $[\boldsymbol{A}]_{ii} = 0$
  2. The out-degree weight vector $\boldsymbol{d}_{\text{out}} = \boldsymbol{A}\boldsymbol{1}$ (i.e. $[\boldsymbol{d}_{\text{out}}]_i = \sum_{j \in \mathcal{N}_i} w_{ij}$)
  3. The out-flow rate vector $[\boldsymbol{w}_o]_i = w_{i0}$
- The equation can be written as a function of the Laplacian matrix $\boldsymbol{L} = \text{diag}(\boldsymbol{d}_{\text{out}}) - \boldsymbol{A}$ as follows:

$$\dot{\boldsymbol{x}} = -(\text{diag}(\boldsymbol{w}_o) + \boldsymbol{L})\boldsymbol{x} + \boldsymbol{b}$$

- The matrix $-\text{diag}(\boldsymbol{w}_o) + \boldsymbol{L}) = \boldsymbol{M}$ the Metzler matrix for this system also called the **compartmental matrix**

Special case: Linear flow functions

Cornell University

- Introducing:
  1. the weighted adjacency matrix containing all flow rates $[\boldsymbol{A}]_{ij} = w_{ij}, i, j \in \mathcal{V}, \quad i \neq j$ and such that $[\boldsymbol{A}]_{ii} = 0$
  2. The out-degree weight vector $\boldsymbol{d}_{\text{out}} = \boldsymbol{A}\mathbf{1}$ (i.e. $[\boldsymbol{d}_{\text{out}}]_i = \sum_{j \in \mathcal{N}_i} w_{ij}$)
  3. The out-flow rate vector $[\boldsymbol{w}_o]_i = w_{i0}$
- The equation can be written as a function of the Laplacian matrix $\boldsymbol{L} = \text{diag}(\boldsymbol{d}_{\text{out}}) - \boldsymbol{A}$ as follows:

$$\dot{\boldsymbol{x}} = -(\text{diag}(\boldsymbol{w}_o) + \boldsymbol{L})\boldsymbol{x} + \boldsymbol{b}$$

- The matrix $-\text{diag}(\boldsymbol{w}_o) + \boldsymbol{L}) = \boldsymbol{M}$ the Metzler matrix for this system also called the **compartmental matrix**

Outline

Cornell University

Modern infrastructures as cyber-physical systems

- Traditional infrastructures move **physical flows**: vehicles, packets, water, gas, electric power.
- Modern infrastructures are also **sensed, networked, and software-operated**.
- They combine:
  1. a **physical layer** (roads, routers, pipes, feeders),
  2. a **sensing layer** (detectors, GPS, traffic counters, packet traces, smart meters),
  3. a **communication/computing layer**,
  4. a **decision layer** (routing, scheduling, pricing, congestion control, dispatch).
- This is why infrastructure operation is increasingly a problem of estimation, prediction, optimization, control, and resilience.

Modern infrastructures as cyber-physical systems

Cornell University

- Traditional infrastructures move **physical flows**: vehicles, packets, water, gas, electric power.
- Modern infrastructures are also **sensed, networked, and software-operated**.
- They combine:
  1. a **physical layer** (roads, routers, pipes, feeders),
  2. a **sensing layer** (detectors, GPS, traffic counters, packet traces, smart meters),
  3. a **communication/computing layer**,
  4. a **decision layer** (routing, scheduling, pricing, congestion control, dispatch).
- This is why infrastructure operation is increasingly a problem of estimation, prediction, optimization, control, and resilience.

Modern infrastructures as cyber-physical systems

- Traditional infrastructures move **physical flows**: vehicles, packets, water, gas, electric power.
- Modern infrastructures are also **sensed, networked, and software-operated**.
- They combine:
  1. a **physical layer** (roads, routers, pipes, feeders),
  2. a **sensing layer** (detectors, GPS, traffic counters, packet traces, smart meters),
  3. a **communication/computing layer**,
  4. a **decision layer** (routing, scheduling, pricing, congestion control, dispatch).
- This is why infrastructure operation is increasingly a problem of estimation, prediction, optimization, control, and resilience.

Modern infrastructures as cyber-physical systems

- Traditional infrastructures move **physical flows**: vehicles, packets, water, gas, electric power.
- Modern infrastructures are also **sensed, networked, and software-operated**.
- They combine:
  1. a **physical layer** (roads, routers, pipes, feeders),
  2. a **sensing layer** (detectors, GPS, traffic counters, packet traces, smart meters),
  3. a **communication/computing layer**,
  4. a **decision layer** (routing, scheduling, pricing, congestion control, dispatch).
- This is why infrastructure operation is increasingly a problem of estimation, prediction, optimization, control, and resilience.

Flow networks examples

Cornell University

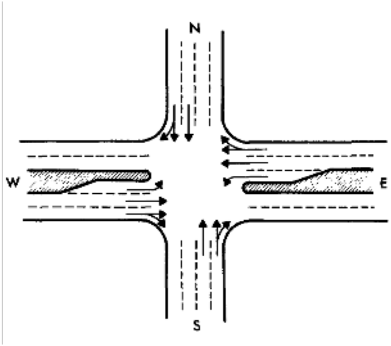
- We are going to go through key equations for network infrastructures:
  1. Transportation networks
  2. The Internet
  3. Pipelines (for water, gasoline, sewage, anything fluid and gas)
- Gas pipelines have a harder description because their storage is inside the edges, i.e. the pipes.
- Power System could be but there is no storage on the lines

Outline

Cornell University

1. Review of abstract models
2. Review of Flow Networks
3. Special cases of flow networks
4. Transportation networks
5. The Internet and Packet Switched Networks
6. Conclusions

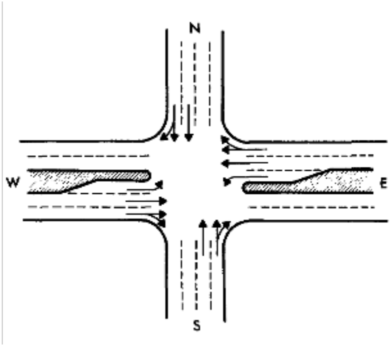
Urban transportation networks



- The roadway network includes intersections and streets through which traffic moves.
- These elements can be translated into a structure of nodes and directed links (including weights), respectively
- The weight measures a *link impedance*, as the time it takes to transverse the link, the energy cost etc.
- A two-way street is represented by two opposite-direction links

Urban transportation networks

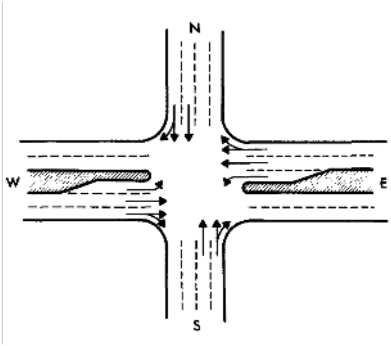
Cornell University



- The roadway network includes intersections and streets through which traffic moves.
- These elements can be translated into a structure of nodes and directed links (including weights), respectively
- The weight measures a *link impedance*, as the time it takes to transverse the link, the energy cost etc.
- A two-way street is represented by two opposite-direction links

Urban transportation networks

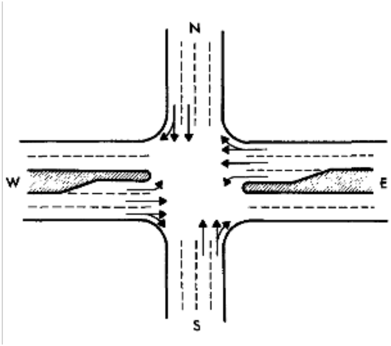
Cornell University



- The roadway network includes intersections and streets through which traffic moves.
- These elements can be translated into a structure of nodes and directed links (including weights), respectively
- The weight measures a *link impedance*, as the time it takes to transverse the link, the energy cost etc.
- A two-way street is represented by two opposite-direction links

Urban transportation networks

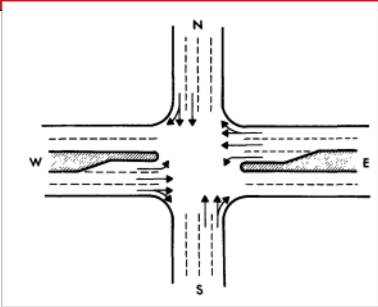
Cornell University

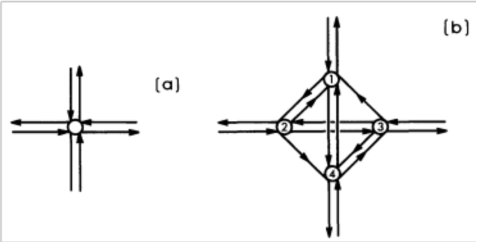


- The roadway network includes intersections and streets through which traffic moves.
- These elements can be translated into a structure of nodes and directed links (including weights), respectively
- The weight measures a *link impedance*, as the time it takes to transverse the link, the energy cost etc.
- A two-way street is represented by two opposite-direction links

Transportation networks representation

Cornell University

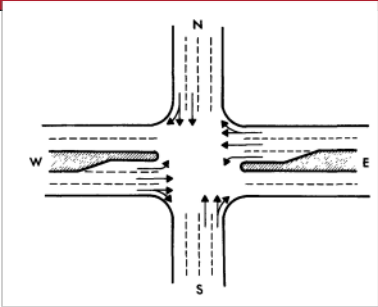


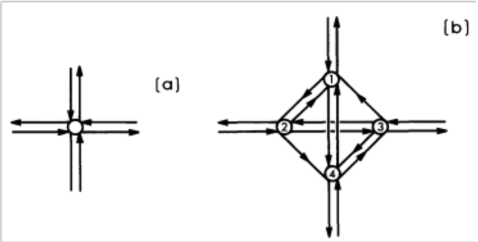


- As seen in the figure, the network mapping is not unique
- The intersection can be represented
  (a) simply as a node with the streets leading to and from it represented by links connected to that node;
  (b) by a network of four nodes and 10 links so that each permissible movement through the intersection is a separate link.

Transportation networks representation

Cornell University



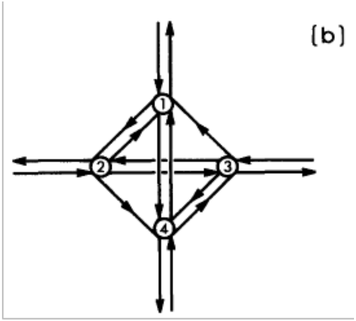


- As seen in the figure, the network mapping is not unique
- The intersection can be represented
  (a) simply as a node with the streets leading to and from it represented by links connected to that node;
  (b) by a network of four nodes and 10 links so that each permissible movement through the intersection is a separate link.

Intersection flows

Cornell University

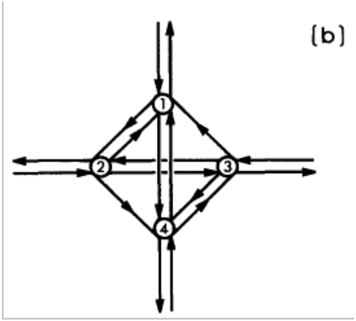
In this representation, each intersection movement can be associated with the appropriate delay



- Link (3,4), will be associated with left-turning delay experienced by the traffic turning from the eastern approach into the southern one.
- Link (3,1), will represent right-turning maneuvers for flow from the east to the north, with the appropriate delay.
- The links leading into the intersection will be associated only with the travel time on the incoming approaches, not with the intersection delay itself.

Intersection flows

In this representation, each intersection movement can be associated with the appropriate delay

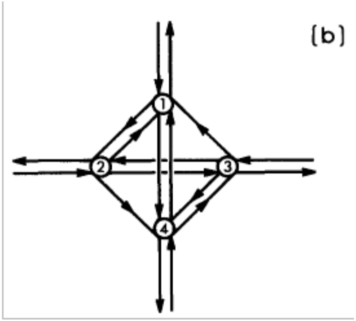


- Link (3,4), will be associated with left-turning delay experienced by the traffic turning from the eastern approach into the southern one.
- Link (3,1), will represent right-turning maneuvers for flow from the east to the north, with the appropriate delay.
- The links leading into the intersection will be associated only with the travel time on the incoming approaches, not with the intersection delay itself.

Intersection flows

Cornell University

In this representation, each intersection movement can be associated with the appropriate delay

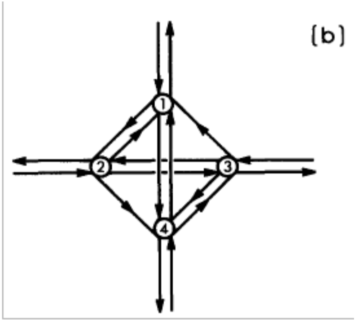


- Link (3,4), will be associated with left-turning delay experienced by the traffic turning from the eastern approach into the southern one.
- Link (3,1), will represent right-turning maneuvers for flow from the east to the north, with the appropriate delay.
- The links leading into the intersection will be associated only with the travel time on the incoming approaches, not with the intersection delay itself.

Intersection flows

Cornell University

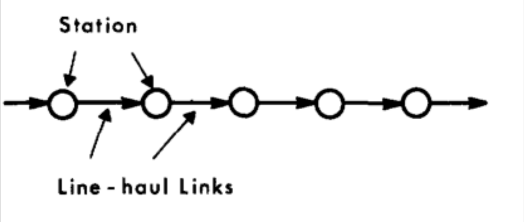
In this representation, each intersection movement can be associated with the appropriate delay



- Link (3,4), will be associated with left-turning delay experienced by the traffic turning from the eastern approach into the southern one.
- Link (3,1), will represent right-turning maneuvers for flow from the east to the north, with the appropriate delay.
- The links leading into the intersection will be associated only with the travel time on the incoming approaches, not with the intersection delay itself.

Transit networks

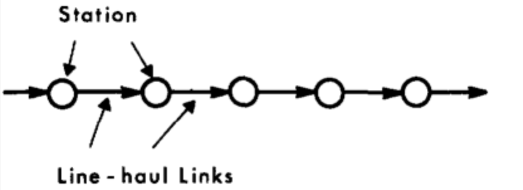
Cornell University



- Vehicular traffic through streets and intersections is not the only flow in the urban area. Passenger transit-lines capture the movement of people
- A **transit line** can be represented by a simple linear network in which the transit stations (or bus stops) are nodes and the line-haul portions are links
- The link impedance can include not just the travel time but also the transit fare, weight time etc.

Transit networks

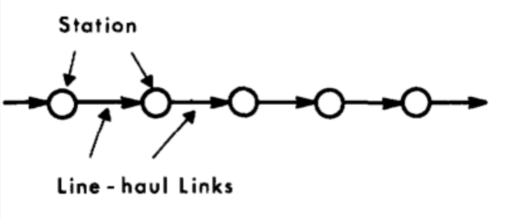
Cornell University



- Vehicular traffic through streets and intersections is not the only flow in the urban area. Passenger transit-lines capture the movement of people
- A **transit line** can be represented by a simple linear network in which the transit stations (or bus stops) are nodes and the line-haul portions are links
- The link impedance can include not just the travel time but also the transit fare, weight time etc.

Transit networks

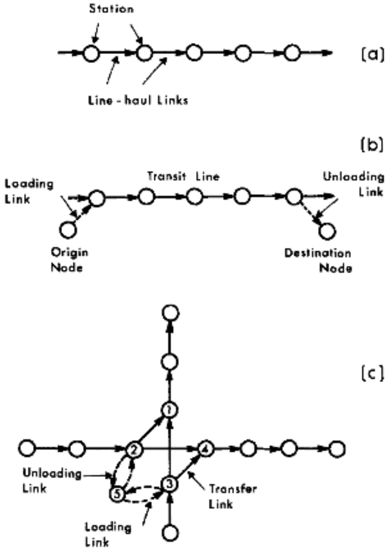
Cornell University



- Vehicular traffic through streets and intersections is not the only flow in the urban area. Passenger transit-lines capture the movement of people
- A **transit line** can be represented by a simple linear network in which the transit stations (or bus stops) are nodes and the line-haul portions are links
- The link impedance can include not just the travel time but also the transit fare, weight time etc.

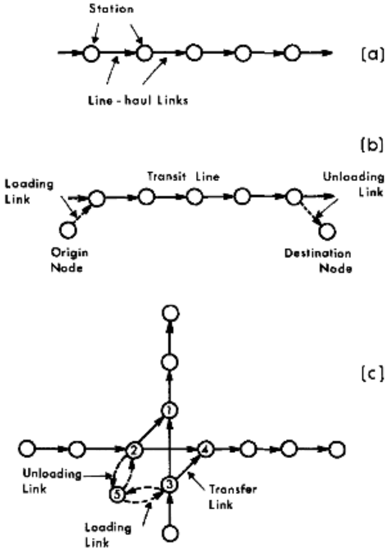
Transit lines

Cornell University



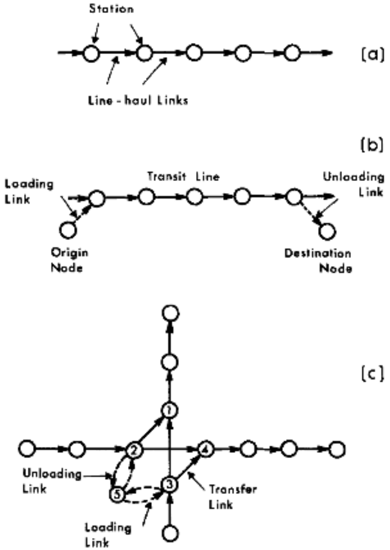
- Reaching a point in the transit line or reaching a specific destination may require walking. This can be captured adding links that connect to loading and unloading points (figure (b)).
- A more detailed representation of a transit station would include transfer and boarding links (figure (c)).

Transit lines



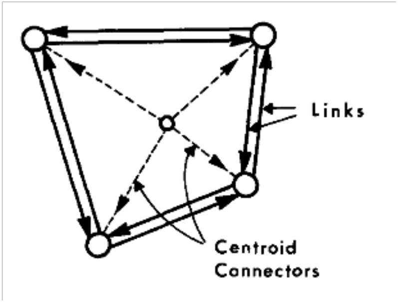
- Reaching a point in the transit line or reaching a specific destination may require walking. This can be captured adding links that connect to loading and unloading points (figure (b)).
- A more detailed representation of a transit station would include transfer and boarding links (figure (c)).

Transit lines



- Reaching a point in the transit line or reaching a specific destination may require walking. This can be captured adding links that connect to loading and unloading points (figure (b)).
- A more detailed representation of a transit station would include transfer and boarding links (figure (c)).

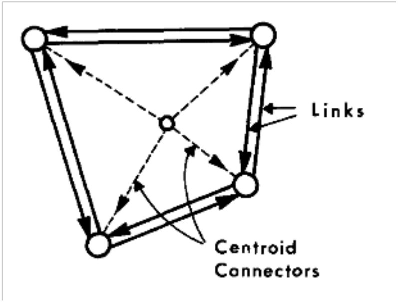
Centroids and traffic-zones



- In transportation planning urban areas are typically partitioned into traffic zones.
- The size of each traffic zone can vary from a city block to a whole neighborhood or a town within the urban area.
- Each traffic zone is represented by a node known as **centroid** that are the aggregate "source" and "sink" nodes where traffic originates and to which traffic is destined.

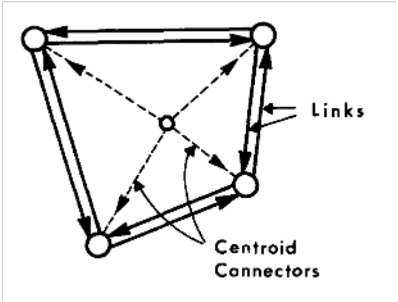
Centroids and traffic-zones

Cornell University



- In transportation planning urban areas are typically partitioned into traffic zones.
- The size of each traffic zone can vary from a city block to a whole neighborhood or a town within the urban area.
- Each traffic zone is represented by a node known as **centroid** that are the aggregate "source" and "sink" nodes where traffic originates and to which traffic is destined.

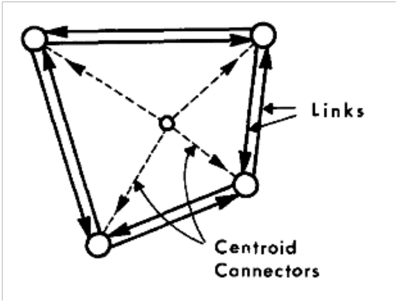
Centroids and traffic-zones



- In transportation planning urban areas are typically partitioned into traffic zones.
- The size of each traffic zone can vary from a city block to a whole neighborhood or a town within the urban area.
- Each traffic zone is represented by a node known as **centroid** that are the aggregate "source" and "sink" nodes where traffic originates and to which traffic is destined.

Centroids and traffic-zones

Cornell University



- In transportation planning urban areas are typically partitioned into traffic zones.
- The size of each traffic zone can vary from a city block to a whole neighborhood or a town within the urban area.
- Each traffic zone is represented by a node known as **centroid** that are the aggregate "source" and "sink" nodes where traffic originates and to which traffic is destined.

Link weights and travel impedance

Cornell University

- Links representing an urban network can have many attributes, reflecting travel time, safety, cost of travel, stability of flow, and others.
- The primary component of the link impedance, however, is travel time, which is often used as the sole measure of link impedance.
- If included in the analysis all impedance measures should be expressed in travel-time units.

Link weights and travel impedance

Cornell University

- Links representing an urban network can have many attributes, reflecting travel time, safety, cost of travel, stability of flow, and others.
- The primary component of the link impedance, however, is travel time, which is often used as the sole measure of link impedance.
- If included in the analysis all impedance measures should be expressed in travel-time units.

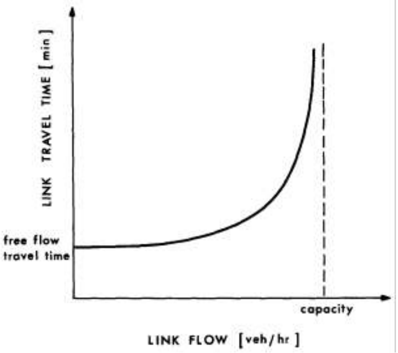
Link weights and travel impedance

Cornell University

- Links representing an urban network can have many attributes, reflecting travel time, safety, cost of travel, stability of flow, and others.
- The primary component of the link impedance, however, is travel time, which is often used as the sole measure of link impedance.
- If included in the analysis all impedance measures should be expressed in travel-time units.

Link Performance function

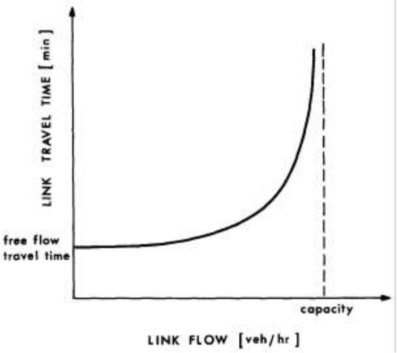
Cornell University



- The typical approach in transportation studies it to associate to each link a **performance function** rather than a constant travel-time measure
- The performance function maps the travel time (which, as mentioned above, may stand for generalized impedance) on each link to the flow traversing this link.
- Travel time at zero flow is known as the free-flow travel time.

Link Performance function

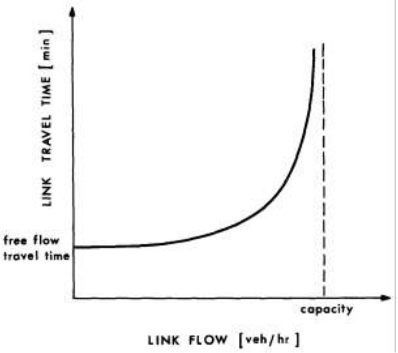
Cornell University



- The typical approach in transportation studies it to associate to each link a **performance function** rather than a constant travel-time measure
- The performance function maps the travel time (which, as mentioned above, may stand for generalized impedance) on each link to the flow traversing this link.
- Travel time at zero flow is known as the free-flow travel time.

Link Performance function

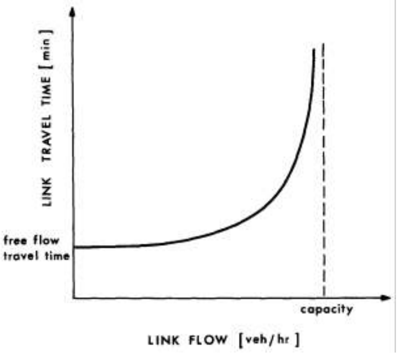
Cornell University



- The typical approach in transportation studies it to associate to each link a **performance function** rather than a constant travel-time measure
- The performance function maps the travel time (which, as mentioned above, may stand for generalized impedance) on each link to the flow traversing this link.
- Travel time at zero flow is known as the free-flow travel time.

Link Performance function

Cornell University



- The typical approach in transportation studies it to associate to each link a **performance function** rather than a constant travel-time measure
- The performance function maps the travel time (which, as mentioned above, may stand for generalized impedance) on each link to the flow traversing this link.
- Travel time at zero flow is known as the free-flow travel time.

Definitions of Equilibria

Cornell University

- There are several multimodal transportation systems analyses
- Let us focus on the case of vehicular traffic. Given:
  1. A graph representation of the urban transportation network
  2. The associated link performance functions
  3. An origin-destination matrix

  find the flow (and travel time) on each of the network links.
- This is known as the **traffic assignment problem** since the issue is how to assign the O-D matrix onto the network.
- The resulting link flows are then used to calculate an array of measures, which can be used, in turn, to evaluate the network

Definitions of Equilibria

Cornell University

- There are several multimodal transportation systems analyses
- Let us focus on the case of vehicular traffic. Given:
  1. A graph representation of the urban transportation network
  2. The associated link performance functions
  3. An origin-destination matrix

  find the flow (and travel time) on each of the network links.
- This is known as the **traffic assignment problem** since the issue is how to assign the O-D matrix onto the network.
- The resulting link flows are then used to calculate an array of measures, which can be used, in turn, to evaluate the network

Definitions of Equilibria

Cornell University

- There are several multimodal transportation systems analyses
- Let us focus on the case of vehicular traffic. Given:
  1. A graph representation of the urban transportation network
  2. The associated link performance functions
  3. An origin-destination matrix

  find the flow (and travel time) on each of the network links.
- This is known as the **traffic assignment problem** since the issue is how to assign the O-D matrix onto the network.
- The resulting link flows are then used to calculate an array of measures, which can be used, in turn, to evaluate the network

Definitions of Equilibria

Cornell University

- There are several multimodal transportation systems analyses
- Let us focus on the case of vehicular traffic. Given:
  1. A graph representation of the urban transportation network
  2. The associated link performance functions
  3. An origin-destination matrix

  find the flow (and travel time) on each of the network links.
- This is known as the **traffic assignment problem** since the issue is how to assign the O-D matrix onto the network.
- The resulting link flows are then used to calculate an array of measures, which can be used, in turn, to evaluate the network

Traffic assignment problem

Cornell University

- To solve the traffic assignment problem, one needs to know the **rule by which motorists choose a route**.
- This rule depends on the demand for travel over the different paths.
- The assignment problem is solved seeking an *equilibrium*.
- The equilibrium flows and corresponding travel times throughout the network are determined by the routes chosen between all O-D pairs, on the one hand, and the performance functions on all the network links, on the other.

Traffic assignment problem

Cornell University

- To solve the traffic assignment problem, one needs to know the **rule by which motorists choose a route**.
- This rule depends on the demand for travel over the different paths.
- The assignment problem is solved seeking an *equilibrium*.
- The equilibrium flows and corresponding travel times throughout the network are determined by the routes chosen between all O-D pairs, on the one hand, and the performance functions on all the network links, on the other.

Traffic assignment problem

Cornell University

- To solve the traffic assignment problem, one needs to know the **rule by which motorists choose a route**.
- This rule depends on the demand for travel over the different paths.
- The assignment problem is solved seeking an *equilibrium*.
- The equilibrium flows and corresponding travel times throughout the network are determined by the routes chosen between all O-D pairs, on the one hand, and the performance functions on all the network links, on the other.

Traffic assignment problem

Cornell University

- To solve the traffic assignment problem, one needs to know the **rule by which motorists choose a route**.
- This rule depends on the demand for travel over the different paths.
- The assignment problem is solved seeking an *equilibrium*.
- The equilibrium flows and corresponding travel times throughout the network are determined by the routes chosen between all O-D pairs, on the one hand, and the performance functions on all the network links, on the other.

Modeling route choice

Cornell University

- It is reasonable to assume that every motorist will try to minimize his or her own travel time from origin to destination.
- The travel time on each link changes with the flow and therefore, the travel time on several of the network paths changes as the link flows change.

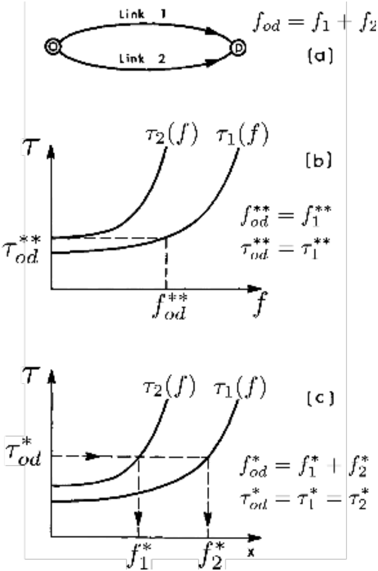
User-equilibrium (UE) condition

A **stable condition** is reached only when no traveler can improve his travel time by unilaterally changing routes.

- **Modeling assumptions**: The UE definition implies that all motorists have identical behavior, have full information (i.e., they know the travel time on every possible route) and that they consistently make the correct route decisions. All this is clearly untrue.

UE computation

Cornell University

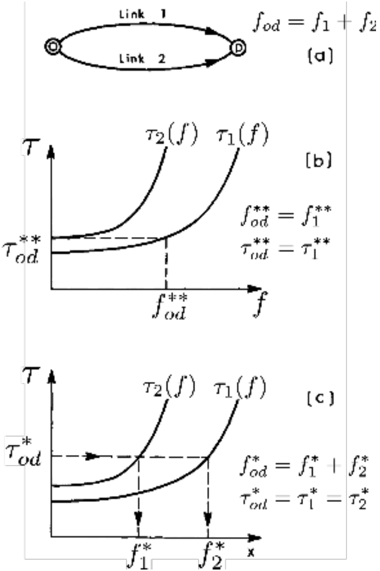


In this simple two routes example UE condition is as follows

- If the total flow is less than f_{od}^{**}, then all traffic flows in the first link
- If the total flow exceeds f_{od}^{**}, then the solution to accomodate f_{od}^{*} is always that the traffic splits so that the performance function gives equal time for the two routes
- The graphical method used to solve this small two-links network cannot be used to solve large networks because there are too many routes to consider.

UE computation

Cornell University

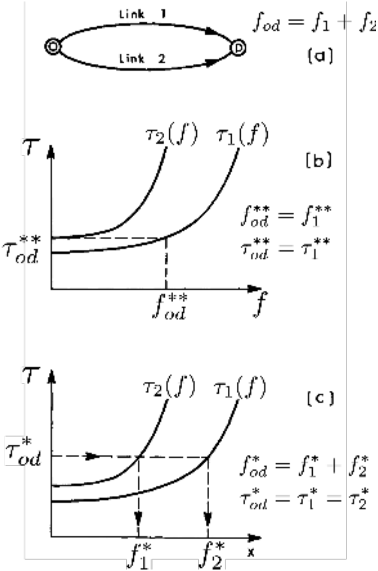


In this simple two routes example UE condition is as follows

- If the total flow is less than f^{**}_{od}, then all traffic flows in the first link
- If the total flow exceeds f^{**}_{od}, then the solution to accomodate f^{*}_{od} is always that the traffic splits so that the performance function gives equal time for the two routes
- The graphical method used to solve this small two-links network cannot be used to solve large networks because there are too many routes to consider.

UE computation

Cornell University

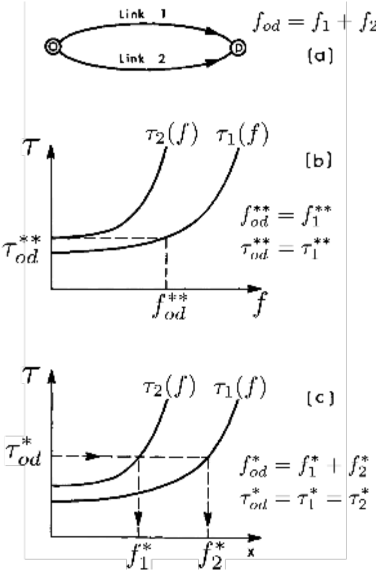


In this simple two routes example UE condition is as follows

- If the total flow is less than f_{od}^{**}, then all traffic flows in the first link
- If the total flow exceeds f_{od}^{**}, then the solution to accomodate f_{od}^{*} is always that the traffic splits so that the performance function gives equal time for the two routes
- The graphical method used to solve this small two-links network cannot be used to solve large networks because there are too many routes to consider.

UE computation

Cornell University



In this simple two routes example UE condition is as follows

- If the total flow is less than f_{od}^{**}, then all traffic flows in the first link
- If the total flow exceeds f_{od}^{**}, then the solution to accomodate f_{od}^* is always that the traffic splits so that the performance function gives equal time for the two routes
- The graphical method used to solve this small two-links network cannot be used to solve large networks because there are too many routes to consider.

General problem formulation

Cornell University

- The sets of centroids representing sources and destinations are $\mathcal{O}$ and $\mathcal{D}$ respectively, such that $\mathcal{O} \cap \mathcal{D} = \emptyset$
- $\boldsymbol{F}$ is the $|\mathcal{O}| \times |\mathcal{D}|$ matrix of all origin destination flows $[\boldsymbol{F}]_{od} = f_{od} \geq 0$.
- In the network flow model $\sum_{d\in\mathcal{D}} f_{od} = b_o$ and $\sum_{o\in\mathcal{O}} f_{od} = f_d^{\text{out}}$.
- Let F_e be the flow on link e with performance function $\tau_e(f)$.
- The origin destination flows are split through a set $\mathcal{K}^{od}$ of possible routes $\mathcal{P}_k^{od}$, whose flows are denoted by f_k^{od} and travel times by τ_k^{od}
- We need to use indicator variables $\delta_{e,k}^{od} \in \{0,1\}$:

$$\delta_{e,k}^{od} = \begin{cases} 1 & \text{if edge } e \in \mathcal{P}_k^{od} \\ 0 & \text{else} \end{cases}$$

General problem formulation

Cornell University

- The sets of centroids representing sources and destinations are $\mathcal{O}$ and $\mathcal{D}$ respectively, such that $\mathcal{O} \cap \mathcal{D} = \emptyset$
- $\boldsymbol{F}$ is the $|\mathcal{O}| \times |\mathcal{D}|$ matrix of all origin destination flows $[\boldsymbol{F}]_{od} = f_{od} \geq 0$.
- In the network flow model $\sum_{d \in \mathcal{D}} f_{od} = b_o$ and $\sum_{o \in \mathcal{O}} f_{od} = f_d^{\text{out}}$.
- Let F_e be the flow on link e with performance function $\tau_e(f)$.
- The origin destination flows are split through a set $\mathcal{K}^{od}$ of possible routes $\mathcal{P}_k^{od}$, whose flows are denoted by f_k^{od} and travel times by τ_k^{od}
- We need to use indicator variables $\delta_{e,k}^{od} \in \{0, 1\}$:

$$\delta_{e,k}^{od} = \begin{cases} 1 & \text{if edge } e \in \mathcal{P}_k^{od} \\ 0 & \text{else} \end{cases}$$

Identities

Cornell University

- The following identities hold for each route $k \in \mathcal{K}_{od}$ and each origin and destination $(o, d) \in \mathcal{O} \times \mathcal{D}$:

$$\tau_k^{od} = \sum_{e \in \mathcal{E}} \tau_e(F_e)\delta_{e,k}^{od},$$

$$F_e = \sum_{o \in \mathcal{O}} \sum_{d \in \mathcal{D}} \sum_{k} f_k^{od}\delta_{e,k}^{od}, \quad \forall e \in \mathcal{E}$$

$$f_{od} = \sum_{k \in \mathcal{K}^{od}} f_k^{od},$$

$$f_k^{od} \geq 0$$

- It is important to note that this formulation assumes that the travel time on a given link is a function of the flow on that link only and not of the flow on any other link in the network. This somewhat restrictive

Identities

Cornell University

- The following identities hold for each route $k \in \mathcal{K}_{od}$ and each origin and destination $(o, d) \in \mathcal{O} \times \mathcal{D}$:

$$\tau_k^{od} = \sum_{e \in \mathcal{E}} \tau_e(F_e)\delta_{e,k}^{od},$$

$$F_e = \sum_{o \in \mathcal{O}} \sum_{d \in \mathcal{D}} \sum_{k} f_k^{od}\delta_{e,k}^{od}, \quad \forall e \in \mathcal{E}$$

$$f_{od} = \sum_{k \in \mathcal{K}^{od}} f_k^{od},$$

$$f_k^{od} \geq 0$$

- It is important to note that this formulation assumes that the travel time on a given link is a function of the flow on that link only and not of the flow on any other link in the network. This somewhat restrictive

Beckmann transformation

Cornell University

- In the 50s Beckmann proved that the UE can be found solving the following mimization:

$$\min_{F_e, e\in\mathcal{E}} \left(\sum_{e\in\mathcal{E}} \int_0^{F_e} \tau_e(v)dv\right)$$

$$\text{subject to } F_e = \sum_{o\in\mathcal{O}} \sum_{d\in\mathcal{D}} \sum_{k} f_k^{od} \delta_{e,k}^{od}$$

$$f_{od} = \sum_{k\in\mathcal{K}^{od}} f_k^{od},$$

$$f_k^{od} \geq 0$$

- For a linear $\tau_e(F)$ function this is a mixed integer linear program
- Basically Beckmann proved that the solution gives a flow pattern satisfying the UE conditions: the flows over the routes are either zero or take equal time. Also, he proved the solution is unique.

Beckmann transformation

Cornell University

- In the 50s Beckmann proved that the UE can be found solving the following mimization:

$$\min_{F_e, e\in\mathcal{E}} \left(\sum_{e\in\mathcal{E}} \int_0^{F_e} \tau_e(v)dv \right)$$

$$\text{subject to } F_e = \sum_{o\in\mathcal{O}} \sum_{d\in\mathcal{D}} \sum_{k} f_k^{od} \delta_{e,k}^{od}$$

$$f_{od} = \sum_{k\in\mathcal{K}^{od}} f_k^{od},$$

$$f_k^{od} \geq 0$$

- For a linear $\tau_e(F)$ function this is a mixed integer linear program
- Basically Beckmann proved that the solution gives a flow pattern satisfying the UE conditions: the flows over the routes are either zero or take equal time. Also, he proved the solution is unique.

Beckmann transformation

- In the 50s Beckmann proved that the UE can be found solving the following mimization:

$$\min_{F_e, e\in\mathcal{E}} \left(\sum_{e\in\mathcal{E}} \int_0^{F_e} \tau_e(v)dv \right)$$

$$\text{subject to } F_e = \sum_{o\in\mathcal{O}} \sum_{d\in\mathcal{D}} \sum_{k} f_k^{od}\delta_{e,k}^{od}$$

$$f_{od} = \sum_{k\in\mathcal{K}^{od}} f_k^{od},$$

$$f_k^{od} \geq 0$$

- For a linear $\tau_e(F)$ function this is a mixed integer linear program
- Basically Beckmann proved that the solution gives a flow pattern satisfying the UE conditions: the flows over the routes are either zero or take equal time. Also, he proved the solution is unique.

Generalization: Stochastic User Equilibrium

- One can introduced a random variable modeling the **perceived travel time**, based on which the motorists make the decisions.

Stochastic user-equilibrium (UE) condition

A **stable condition** is reached only when no traveler **believes** he/she can improve his travel time by unilaterally changing routes.

Consideration on equilibrium studies

Cornell University

- Steady-state equilibrium analyses are valid only if the flows are constant.
- Consequently, the analysis is done for certain time periods such as the “morning peak”, “evening peak”, or “midday” considering origin-destination flows within each of these respective periods approximately constant.
- The period of analysis cannot be too long but also not very small, since it has longer than the typical duration of trips at this time.

Consideration on equilibrium studies

Cornell University

- Steady-state equilibrium analyses are valid only if the flows are constant.
- Consequently, the analysis is done for certain time periods such as the "morning peak", "evening peak", or "midday" considering origin-destination flows within each of these respective periods approximately constant.
- The period of analysis cannot be too long but also not very small, since it has longer than the typical duration of trips at this time.

System optimum formulation

Cornell University

- In this case the total travel time in the network is minimized (this is not in general an UE condition)
- The formulation is:

$$\min_{F_e, e\in\mathcal{E}} \left(\sum_{e\in\mathcal{E}} F_e \tau_e(F_e) \right)$$

$$\text{subject to } F_e = \sum_{o\in\mathcal{O}} \sum_{d\in\mathcal{D}} \sum_{k} f_k^{od} \delta_{e,k}^{od}$$

$$f_{od} = \sum_{k\in\mathcal{K}^{od}} f_k^{od},$$

$$f_k^{od} \geq 0$$

- Note that here we are not considering capacity constraints, basically because we are assuming that we can route in finite time (else we would exceed the capacity of the road)

System optimum formulation

Cornell University

- In this case the total travel time in the network is minimized (this is not in general an UE condition)
- The formulation is:

$$\min_{F_e, e\in\mathcal{E}} \left(\sum_{e\in\mathcal{E}} F_e \tau_e(F_e)\right)$$

$$\text{subject to } F_e = \sum_{o\in\mathcal{O}} \sum_{d\in\mathcal{D}} \sum_{k} f_k^{od} \delta_{e,k}^{od}$$

$$f_{od} = \sum_{k\in\mathcal{K}^{od}} f_k^{od},$$

$$f_k^{od} \geq 0$$

- Note that here we are not considering capacity constraints, basically because we are assuming that we can route in finite time (else we would exceed the capacity of the road)

Data-driven cyber-physical transportation networks

- The roadway network is not only a graph of roads and intersections: it is also a **measured** and **controlled** system.
- Typical data sources include:
  - loop detectors, cameras, tolling data,
  - GPS traces and mobile-phone location data,
  - transit smart-card and schedule data.
- These data support:
  1. **demand estimation** (origin-destination inference),
  2. **travel-time prediction**,
  3. **congestion detection**,
  4. **control and optimization** through tolls, signal timing, routing guidance, and transit scheduling.
- User equilibrium and system optimum provide the mathematical basis for these analytics and control decisions.

Data-driven cyber-physical transportation networks

Cornell University

- The roadway network is not only a graph of roads and intersections: it is also a **measured** and **controlled** system.
- Typical data sources include:
  - loop detectors, cameras, tolling data,
  - GPS traces and mobile-phone location data,
  - transit smart-card and schedule data.
- These data support:
  1. **demand estimation** (origin-destination inference),
  2. **travel-time prediction**,
  3. **congestion detection**,
  4. **control and optimization** through tolls, signal timing, routing guidance, and transit scheduling.
- User equilibrium and system optimum provide the mathematical basis for these analytics and control decisions.

Data-driven cyber-physical transportation networks

- The roadway network is not only a graph of roads and intersections: it is also a **measured** and **controlled** system.
- Typical data sources include:
  - loop detectors, cameras, tolling data,
  - GPS traces and mobile-phone location data,
  - transit smart-card and schedule data.
- These data support:
  1. **demand estimation** (origin-destination inference),
  2. **travel-time prediction**,
  3. **congestion detection**,
  4. **control and optimization** through tolls, signal timing, routing guidance, and transit scheduling.
- User equilibrium and system optimum provide the mathematical basis for these analytics and control decisions.

Data-driven cyber-physical transportation networks

Cornell University

- The roadway network is not only a graph of roads and intersections: it is also a **measured** and **controlled** system.
- Typical data sources include:
  - loop detectors, cameras, tolling data,
  - GPS traces and mobile-phone location data,
  - transit smart-card and schedule data.
- These data support:
  1. **demand estimation** (origin-destination inference),
  2. **travel-time prediction**,
  3. **congestion detection**,
  4. **control and optimization** through tolls, signal timing, routing guidance, and transit scheduling.
- User equilibrium and system optimum provide the mathematical basis for these analytics and control decisions.

Outline

Cornell University

The Internet as a cyber-physical infrastructure

Cornell University

- The Internet is the infrastructure that moves **information flows**.
- Its physical components are links, routers, switches, servers, wireless access points, and data centers.
- Its dynamic state includes:
  - traffic demand,
  - link utilization,
  - queue occupancies,
  - delay, throughput, and packet loss.
- The operational questions are therefore not only architectural:
  1. How do we **route** traffic?
  2. How do we **control congestion**?
  3. How do we **monitor** and **detect anomalies**?
  4. How do we **allocate capacity** efficiently and robustly?

The Internet as a cyber-physical infrastructure

Cornell University

- The Internet is the infrastructure that moves **information flows**.
- Its physical components are links, routers, switches, servers, wireless access points, and data centers.
- Its dynamic state includes:
  - traffic demand,
  - link utilization,
  - queue occupancies,
  - delay, throughput, and packet loss.
- The operational questions are therefore not only architectural:
  1. How do we **route** traffic?
  2. How do we **control congestion**?
  3. How do we **monitor** and **detect anomalies**?
  4. How do we **allocate capacity** efficiently and robustly?

The Internet as a cyber-physical infrastructure

- The Internet is the infrastructure that moves **information flows**.
- Its physical components are links, routers, switches, servers, wireless access points, and data centers.
- Its dynamic state includes:
  - traffic demand,
  - link utilization,
  - queue occupancies,
  - delay, throughput, and packet loss.
- The operational questions are therefore not only architectural:
  1. How do we **route** traffic?
  2. How do we **control congestion**?
  3. How do we **monitor** and **detect anomalies**?
  4. How do we **allocate capacity** efficiently and robustly?

The Internet as a cyber-physical infrastructure

- The Internet is the infrastructure that moves **information flows**.
- Its physical components are links, routers, switches, servers, wireless access points, and data centers.
- Its dynamic state includes:
  - traffic demand,
  - link utilization,
  - queue occupancies,
  - delay, throughput, and packet loss.
- The operational questions are therefore not only architectural:
  1. How do we **route** traffic?
  2. How do we **control congestion**?
  3. How do we **monitor** and **detect anomalies**?
  4. How do we **allocate capacity** efficiently and robustly?

Outline

Cornell University

5 The Internet and Packet Switched Networks

- The architecture of communication networks
- Congestion control
- Internet routing
- Graph data analysis: traffic data

Modern communication systems

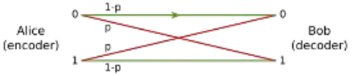
Cornell University

Two key advances

- A universal mathematical notion of information source and channel (Claude Shannon 1945)
- A (software) decomposition of networks architectures into layers (1970)
  The concept of a seven-layer architecture, evolved from early network systems like ARPANET: the Open Systems Interconnection (OSI) Reference Model (1983)

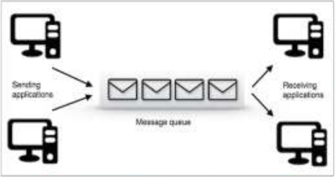
Packet switching

Cornell University

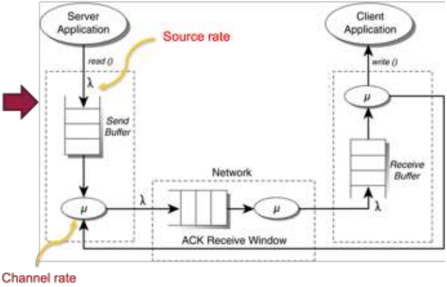


Link abstraction ⇒

- Digital data can be buffered and re-routed



If the channel is very fast, there is very little latency. The delay grows with the queue in front.



Why packet switching matters

Cornell University

- Digital information can be divided into packets, buffered, and routed over different paths.
- This makes communication networks:
  - scalable,
  - reconfigurable,
  - robust to local failures,
  - amenable to optimization and adaptive control.
- The price is that packets compete for limited resources, creating congestion, delay, and queueing.

Why packet switching matters

Cornell University

- Digital information can be divided into packets, buffered, and routed over different paths.
- This makes communication networks:
  - scalable,
  - reconfigurable,
  - robust to local failures,
  - amenable to optimization and adaptive control.
- The price is that packets compete for limited resources, creating congestion, delay, and queueing.

Why packet switching matters

- Digital information can be divided into packets, buffered, and routed over different paths.
- This makes communication networks:
  - scalable,
  - reconfigurable,
  - robust to local failures,
  - amenable to optimization and adaptive control.
- The price is that packets compete for limited resources, creating congestion, delay, and queueing.

Outline

Cornell University

5 The Internet and Packet Switched Networks

- The architecture of communication networks
- Congestion control
- Internet routing
- Graph data analysis: traffic data

Transmission Control Protocol

Cornell University

- The Transmission Control Protocol (TCP) ensures the end to end delivery (through acknowledgements and retransmissions) while the internal routing decisions are made at the IP layer.
- The User Datagram Protocol (UDP) is a sloppier/but faster, version where there is no acknowledgement
- TCP is connection-oriented, that is connection between client and server is established before data can be sent.
- TCP carries also information about a *window* that is used for **congestion control** controls the input flow to the system, i.e. the u_i from a certain source node i accessing the Internet.
- The mechanism determines the input and output flows of the network from a host $f_i(t)$.

Transmission Control Protocol

Cornell University

- The Transmission Control Protocol (TCP) ensures the end to end delivery (through acknowledgements and retransmissions) while the internal routing decisions are made at the IP layer.
- The User Datagram Protocol (UDP) is a sloppier/but faster, version where there is no acknowledgement
- TCP is connection-oriented, that is connection between client and server is established before data can be sent.
- TCP carries also information about a *window* that is used for **congestion control** controls the input flow to the system, i.e. the u_i from a certain source node i accessing the Internet.
- The mechanism determines the input and output flows of the network from a host $f_i(t)$.

Transmission Control Protocol

Cornell University

- The Transmission Control Protocol (TCP) ensures the end to end delivery (through acknowledgements and retransmissions) while the internal routing decisions are made at the IP layer.
- The User Datagram Protocol (UDP) is a sloppier/but faster, version where there is no acknowledgement
- TCP is connection-oriented, that is connection between client and server is established before data can be sent.
- TCP carries also information about a *window* that is used for **congestion control** controls the input flow to the system, i.e. the u_i from a certain source node i accessing the Internet.
- The mechanism determines the input and output flows of the network from a host $f_i(t)$.

Transmission Control Protocol

Cornell University

- The Transmission Control Protocol (TCP) ensures the end to end delivery (through acknowledgements and retransmissions) while the internal routing decisions are made at the IP layer.
- The User Datagram Protocol (UDP) is a sloppier/but faster, version where there is no acknowledgement
- TCP is connection-oriented, that is connection between client and server is established before data can be sent.
- TCP carries also information about a *window* that is used for **congestion control** controls the input flow to the system, i.e. the u_i from a certain source node i accessing the Internet.
- The mechanism determines the input and output flows of the network from a host $f_i(t)$.

Transmission Control Protocol

Cornell University

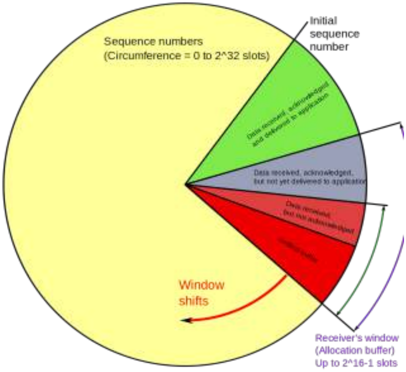
- The Transmission Control Protocol (TCP) ensures the end to end delivery (through acknowledgements and retransmissions) while the internal routing decisions are made at the IP layer.
- The User Datagram Protocol (UDP) is a sloppier/but faster, version where there is no acknowledgement
- TCP is connection-oriented, that is connection between client and server is established before data can be sent.
- TCP carries also information about a *window* that is used for **congestion control** controls the input flow to the system, i.e. the u_i from a certain source node i accessing the Internet.
- The mechanism determines the input and output flows of the network from a host $f_i(t)$.

Congestion control

Cornell University

TCP sequence numbers and receive windows behave very much like a clock.

The receive window shifts each time the receiver receives and acknowledges a new segment of data. Once it runs out of sequence numbers, the sequence number loops back to 0.



Congestion control strategies used in modern TCP are:

1. slow-start
2. congestion avoidance
3. fast retransmit
4. fast recovery

TCP native protocol is a distributed heuristic.

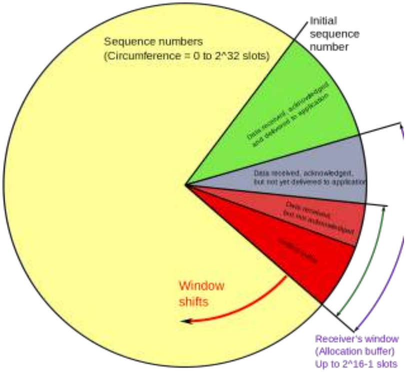
Next: the Internet analogue of transportation assignment: users compete for shared link capacity.

Congestion control

Cornell University

TCP sequence numbers and receive windows behave very much like a clock.

The receive window shifts each time the receiver receives and acknowledges a new segment of data. Once it runs out of sequence numbers, the sequence number loops back to 0.



Congestion control strategies used in modern TCP are:

1. slow-start
2. congestion avoidance
3. fast retransmit
4. fast recovery

TCP native protocol is a distributed heuristic.

Next: the Internet analogue of transportation assignment: users compete for shared link capacity.

Optimum congestion control

Cornell University

- The optimum cogestion control mechanism can be formulated as a Network Utility Maximization (NUM)
- Assuming that the utility of a flow for a certain source is $U_s(f)$, the system optimum is the maximization of the total amount of utility $U_s(f_s)$ for every s subject to the flow constraints.

$$\max \sum_s U_s(f_s)$$

$$\text{subject to } [\boldsymbol{R}\boldsymbol{f}]_i = 0, i \in \mathcal{V}/\{s,t\},$$
$$[\boldsymbol{R}\boldsymbol{f}]_s = -[\boldsymbol{R}\boldsymbol{f}]_t = f_s,$$
$$0 \leq \boldsymbol{f} \leq \overline{\boldsymbol{f}}, \quad , f_s \geq 0$$

where $\boldsymbol{R}$ is the routing matrix, i.e.

$$[\boldsymbol{R}]_{e,i} = \begin{cases} 1 & \text{if flow } i \text{ uses link } e \\ 0 & \text{else} \end{cases}$$

Optimum congestion control

Cornell University

- The optimum cogestion control mechanism can be formulated as a Network Utility Maximization (NUM)
- Assuming that the utility of a flow for a certain source is $U_s(f)$, the system optimum is the maximization of the total amount of utility $U_s(f_s)$ for every s subject to the flow constraints.

$$\max \sum_s U_s(f_s)$$

$$\text{subject to } [\boldsymbol{Rf}]_i = 0, i \in \mathcal{V}/\{s,t\},$$
$$[\boldsymbol{Rf}]_s = -[\boldsymbol{Rf}]_t = f_s,$$
$$0 \leq \boldsymbol{f} \leq \overline{\boldsymbol{f}}, \quad , f_s \geq 0$$

where $\boldsymbol{R}$ is the routing matrix, i.e.

$$[\boldsymbol{R}]_{e,i} = \begin{cases} 1 & \text{if flow } i \text{ uses link } e \\ 0 & \text{else} \end{cases}$$

Outline

Cornell University

5 The Internet and Packet Switched Networks
- The architecture of communication networks
- Congestion control
- Internet routing
- Graph data analysis: traffic data

Internet routing

Cornell University

- The process of determining where to switch packets with an origin destination pair of IP addresses uses static configuration rules or dynamically obtained from **routing protocols**
- There are, of course many routing protocols options and optimization strategies.
- Routing protocols are standardized through the so called "Request for Comments" (RFC)
- Today is possible to change on the fly the routing rules using Software Defined Networks.
- There are efforts undergoing to orchestrate flows from the access network, particularly for wireless access, called Open Radio Access Networks.

Internet routing

Cornell University

- The process of determining where to switch packets with an origin destination pair of IP addresses uses static configuration rules or dynamically obtained from **routing protocols**
- There are, of course many routing protocols options and optimization strategies.
- Routing protocols are standardized through the so called “Request for Comments” (RFC)
- Today is possible to change on the fly the routing rules using Software Defined Networks.
- There are efforts undergoing to orchestrate flows from the access network, particularly for wireless access, called Open Radio Access Networks.

Internet routing

Cornell University

- The process of determining where to switch packets with an origin destination pair of IP addresses uses static configuration rules or dynamically obtained from **routing protocols**
- There are, of course many routing protocols options and optimization strategies.
- Routing protocols are standardized through the so called "Request for Comments" (RFC)
- Today is possible to change on the fly the routing rules using Software Defined Networks.
- There are efforts undergoing to orchestrate flows from the access network, particularly for wireless access, called Open Radio Access Networks.

Internet routing

Cornell University

- The process of determining where to switch packets with an origin destination pair of IP addresses uses static configuration rules or dynamically obtained from **routing protocols**
- There are, of course many routing protocols options and optimization strategies.
- Routing protocols are standardized through the so called "Request for Comments" (RFC)
- Today is possible to change on the fly the routing rules using Software Defined Networks.
- There are efforts undergoing to orchestrate flows from the access network, particularly for wireless access, called Open Radio Access Networks.

Link flow assignment

Cornell University

- Internally each router has a buffer, which instantaneously has a certain state $x_i(t)$. The buffer state at time $t+1$ evolves on average like that of a flow network:

$$x_i(t+1) = \max\left(0, x_i(t) + \sum_{j \in N_i^+} \overline{f}_{j \to i} - \sum_{j \in N_i^-} \overline{f}_{i \to j}\right)$$

  where each link is used at capacity.
- The flow for each IP pairs is chosen depends on the **queueing policy**, e.g.
  1. First in First Out (FIFO), perhaps most common
  2. Last in First Out
  3. Processor Sharing (IP pairs share equally among the link).
  4. Priority or NUM
- This affects the communication delay delay. For instance for FIFO, the delay a packet encounters is a function of the state at the time of its arrival.

Link flow assignment

Cornell University

- Internally each router has a buffer, which instantaneously has a certain state $x_i(t)$. The buffer state at time $t+1$ evolves on average like that of a flow network:

$$x_i(t+1) = \max\left(0, x_i(t) + \sum_{j \in N_i^+} \overline{f}_{j \to i} - \sum_{j \in N_i^-} \overline{f}_{i \to j}\right)$$

where each link is used at capacity.

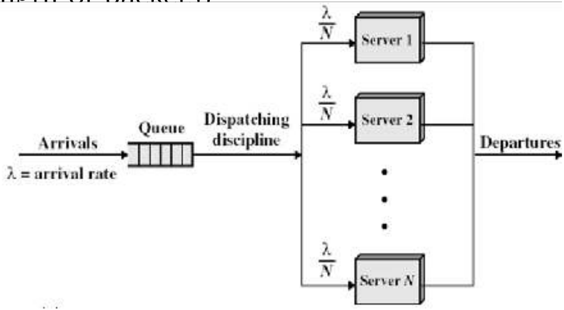
- The flow for each IP pairs is chosen depends on the **queueing policy**, e.g.
  1. First in First Out (FIFO), perhaps most common
  2. Last in First Out
  3. Processor Sharing (IP pairs share equally among the link).
  4. Priority or NUM

- This affects the communication delay delay. For instance for FIFO, the delay a packet encounters is a function of the state at the time of its arrival.

Queuing theory

Cornell University

- For accurate analysis the queues are analyzed as Markov processes, specifying an arrival and service model that can be deterministic or random (the latter is required to capture the random length of packets)



- It allows to understand the performance of the network and therefore design it not for the worse case scenario.

Outline

Cornell University

5 The Internet and Packet Switched Networks

- The architecture of communication networks
- Congestion control
- Internet routing
- Graph data analysis: traffic data

Internet traffic analysis

Cornell University

- Internet traffic analysis utilizes traffic data that, in principle, should not personally identifiable to characterize aggregate as well as subscriber patterns of Internet traffic.
- There are different types of traffic one wants to study. For web services is the number of people visiting, for the management of the network, is the packet flows.
- To better manage the system ISPs' level user traffic profile information and analysis. The types of measurements are
  1. **Passive**: Passive measurement is defined as measurement of existing traffic, observing network packets and connections at different levels (packet or flow). They are useful for traffic engineering, capacity planning and security.
  2. **Active**: Active measurements are obtained through injected traffic with several probe packets are sent continuously.

Internet traffic analysis

Cornell University

- Internet traffic analysis utilizes traffic data that, in principle, should not personally identifiable to characterize aggregate as well as subscriber patterns of Internet traffic.
- There are different types of traffic one wants to study. For web services is the number of people visiting, for the management of the network, is the packet flows.
- To better manage the system ISPs' level user traffic profile information and analysis. The types of measurements are
  1. **Passive**: Passive measurement is defined as measurement of existing traffic, observing network packets and connections at different levels (packet or flow). They are useful for traffic engineering, capacity planning and security.
  2. **Active**: Active measurements are obtained through injected traffic with several probe packets are sent continuously.

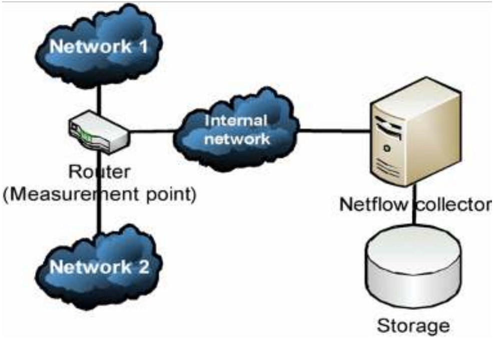
Internet traffic analysis

- Internet traffic analysis utilizes traffic data that, in principle, should not personally identifiable to characterize aggregate as well as subscriber patterns of Internet traffic.
- There are different types of traffic one wants to study. For web services is the number of people visiting, for the management of the network, is the packet flows.
- To better manage the system ISPs' level user traffic profile information and analysis. The types of measurements are
  1. **Passive**: Passive measurement is defined as measurement of existing traffic, observing network packets and connections at different levels (packet or flow). They are useful for traffic engineering, capacity planning and security.
  2. **Active**: Active measurements are obtained through injected traffic with several probe packets are sent continuously.

Passive Internet traffic analysis

Cornell University

- Several tools have a good interface to acquire traffic information e.g.: *Wireshark, TCPdump, SNORT (real time), etc.*

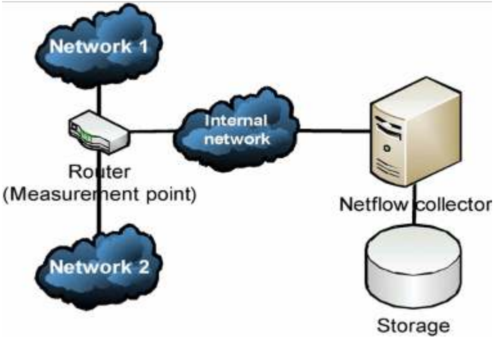


- There are naturally sampling and placement problems to determine.

Passive Internet traffic analysis

Cornell University

- Several tools have a good interface to acquire traffic information e.g.: *Wireshark, TCPdump, SNORT (real time), etc.*



- There are naturally sampling and placement problems to determine.

Outline

Cornell University

5 The Internet and Packet Switched Networks

- The architecture of communication networks
- Congestion control
- Internet routing
- Graph data analysis: traffic data
  - Statistics of network traffic
  - Internet traffic monitoring

Traffic statistics

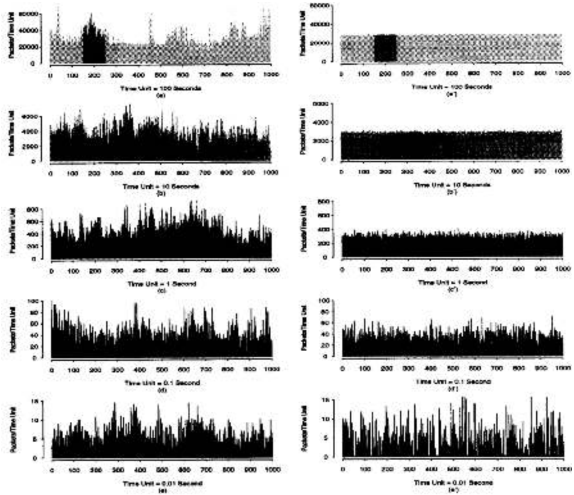
Cornell University

- One of the interesting empirical observations about Internet traffic flows is that they hav [1]:
  - Long Range Dependence: Flows are correlated time series over a long range of time.
  - Heavy tails: When averaged it remains bursty. This property is found in self-similar random processes.
- This is in contrast with the typical models of Poisson random traffic that are used in queueing theory:
  - independent exponential inter-arrival times between packets
  - Let $N(\Delta_t)$ denote the #packets in Δ_t, the Poisson model for the probability is $P(N(\Delta_t) = n) = \frac{(\lambda\Delta_t)^n e^{-\lambda\Delta_t}}{n!}$. The tail of the distribution decays exponentially with n.

[1] W. E. Leland, M. S. Taqqu, W. Willinger and D. V. Wilson, "On the self-similar nature of Ethernet traffic," in IEEE/ACM Transactions on Networking, vol. 2, no. 1, pp. 1-15, Feb. 1994.

Self-similarity of Internet Traffic

Visually what occurs is that if you average the traffic over larger time units you do not smoothen it. Also $P(N > n) = O(n^{-\alpha})$



Self-similar Processes

Cornell University

Let X_k be a weakly stationary (2nd-order stationary) process with mean μ, variance σ^2, and autocorrelation function $\gamma(t)$. Let

$$X_k^{(m)} = \frac{1}{m^H} \sum_{\ell=0}^{m-1} X_{k+\ell}$$

If $\gamma(t) \to t^{-\beta} L(t)$ as $t \to \infty$, where $0 < \beta < 1$ and $\lim_{t\to\infty} \frac{L(tx)}{L(t)} = 1$ for all $x > 0$ (e.g. $L(t) = const$ and $L(t) = \log(t)$) we have:

- **Exactly Self-similar Process:** X_k is exactly self-similar if $\exists H$ such that $X_k^{(m)}$ and X_k have the same distribution.
- **Exactly Second Order Self-similar Process:** X_k is exactly second order self-similar if $\exists H$ such that $X_k^{(m)}$ has the same variance and autocorrelation as X_k.
- **Asymptotic Second Order Self-similar Process:** X_k is an asymptotic second order self-similar process with self-similar parameter H if $\gamma^{(m)}(t) \to \frac{1}{2}[(t+1)^{2H} - 2t^{2H} + (t-1)^{2H}]$ as $m \to \infty$, $\forall t = 1, 2, 3, \ldots$

Self-similar Processes

Let X_k be a weakly stationary (2nd-order stationary) process with mean μ, variance σ^2, and autocorrelation function $\gamma(t)$. Let

$$X_k^{(m)} = \frac{1}{m^H} \sum_{\ell=0}^{m-1} X_{k+\ell}$$

If $\gamma(t) \to t^{-\beta} L(t)$ as $t \to \infty$, where $0 < \beta < 1$ and $\lim_{t\to\infty} \frac{L(tx)}{L(t)} = 1$ for all $x > 0$ (e.g. $L(t) = const$ and $L(t) = \log(t)$) we have:

- **Exactly Self-similar Process:** X_k is exactly self-similar if $\exists H$ such that $X_k^{(m)}$ and X_k have the same distribution.
- **Exactly Second Order Self-similar Process:** X_k is exactly second order self-similar if $\exists H$ such that $X_k^{(m)}$ has the same variance and autocorrelation as X_k.
- **Asymptotic Second Order Self-similar Process:** X_k is an asymptotic second order self-similar process with self-similar parameter H if $\gamma^{(m)}(t) \to \frac{1}{2}[(t+1)^{2H} - 2t^{2H} + (t-1)^{2H}]$ as $m \to \infty$, $\forall t = 1, 2, 3, \ldots$

Outline

Cornell University

5 The Internet and Packet Switched Networks

- The architecture of communication networks
- Congestion control
- Internet routing
- Graph data analysis: traffic data
  - Statistics of network traffic
  - Internet traffic monitoring

Problems in Internet Traffic Analysis

Cornell University

These are some of the topics of interest:[2]

- **Traffic Classification**:
  - Identifying and categorizing different types of traffic (e.g., web browsing, video streaming, VoIP) for network optimization and security purposes.
  - Challenges include encrypted traffic and emerging applications.
- **Anomaly Detection**:
  - Detecting unusual patterns or behaviors in network traffic that may indicate security breaches, network failures, or performance issues.
  - Requires distinguishing between normal and abnormal behavior while handling evolving threats and network dynamics.

[2] A. Callado et al., "A Survey on Internet Traffic Identification," in IEEE Communications Surveys & Tutorials, vol. 11, no. 3, pp. 37-52, 3rd Quarter 2009.

Problems in Internet Traffic Analysis

Cornell University

These are some of the topics of interest:[2]

- **Traffic Classification**:
  - Identifying and categorizing different types of traffic (e.g., web browsing, video streaming, VoIP) for network optimization and security purposes.
  - Challenges include encrypted traffic and emerging applications.
- **Anomaly Detection**:
  - Detecting unusual patterns or behaviors in network traffic that may indicate security breaches, network failures, or performance issues.
  - Requires distinguishing between normal and abnormal behavior while handling evolving threats and network dynamics.

[2] A. Callado et al., "A Survey on Internet Traffic Identification," in IEEE Communications Surveys & Tutorials, vol. 11, no. 3, pp. 37-52, 3rd Quarter 2009.

Problems in Internet Traffic Analysis (cont.)

Cornell University

- **Traffic Prediction**:
  - Forecasting future network traffic volumes and patterns to allocate resources efficiently, plan network upgrades, and mitigate congestion.
  - Incorporates machine learning, time series analysis, and network modeling techniques.
- **Quality of Service (QoS) Management**:
  - Ensuring satisfactory performance for critical applications by prioritizing traffic, managing congestion, and optimizing resource allocation.
  - Balancing conflicting requirements of different services and user demands.

Problems in Internet Traffic Analysis (cont'd)

- **Traffic Prediction**:
  - Forecasting future network traffic volumes and patterns to allocate resources efficiently, plan network upgrades, and mitigate congestion.
  - Incorporates machine learning, time series analysis, and network modeling techniques.
- **Quality of Service (QoS) Management**:
  - Ensuring satisfactory performance for critical applications by prioritizing traffic, managing congestion, and optimizing resource allocation.
  - Balancing conflicting requirements of different services and user demands.

Problems in Internet Traffic Analysis (cont.)

- **Traffic Engineering**:
  - Designing and managing network infrastructure to improve performance, reliability, and efficiency based on traffic characteristics and demands.
  - Involves route optimization, load balancing, and capacity planning.

Traffic monitoring: Classification

- They include: classification (packet or flow level), volume, duration, rate and burstiness.
- Classification at the packet level is useful for
  - traffic shaping (i.e., different applications get different service from the network)
  - traffic filtering (i.e., threats/attacks can be blocked) from *heavy hitters*
  - billing (i.e., per-application charging rates).

Classification

Cornell University

- Classification at the flow level looks into connection patterns in three levels of detail: the social, the functional and the application level. I uses, for instance, the 5-tuple (source and destination addresses, source and destination ports and protocol)
- Classification at the packet level is useful for traffic shaping (i.e., different applications get different service from the network), filtering (i.e., threats/attacks can be blocked) and billing (i.e., per-application charging rates).

Classification

Cornell University

- Classification at the flow level looks into connection patterns in three levels of detail: the social, the functional and the application level. I uses, for instance, the 5-tuple (source and destination addresses, source and destination ports and protocol)
- Classification at the packet level is useful for traffic shaping (i.e., different applications get different service from the network), filtering (i.e., threats/attacks can be blocked) and billing (i.e., per-application charging rates).

Classification

Cornell University

- Classification at the flow level looks into connection patterns in three levels of detail: the social, the functional and the application level. I uses, for instance, the 5-tuple (source and destination addresses, source and destination ports and protocol)
- Classification at the packet level is useful for traffic shaping (i.e., different applications get different service from the network), filtering (i.e., threats/attacks can be blocked) and billing (i.e., per-application charging rates).

Online Identification of Heavy Hitters

Cornell University

- **Heavy Hitters**: In internet traffic analysis, heavy hitters refer to the most significant contributors to network traffic volume, such as top talkers or frequently accessed resources.
- **Online Identification**: In dynamic network environments, online identification methods are essential for real-time detection (e.g. to block DoS attacks).
- **Challenges**:
  - **Real-time Processing**: algorithms must operate in real-time, handling high-speed data streams.
  - **Memory Constraints**: Limited memory resources require algorithms to maintain compact summaries of data streams.
  - **Dynamic Nature**: traffic patterns are non stationary, necessitating adaptive algorithms.

Online Identification of Heavy Hitters

- **Heavy Hitters**: In internet traffic analysis, heavy hitters refer to the most significant contributors to network traffic volume, such as top talkers or frequently accessed resources.
- **Online Identification**: In dynamic network environments, online identification methods are essential for real-time detection (e.g. to block DoS attacks).
- **Challenges**:
  - **Real-time Processing**: algorithms must operate in real-time, handling high-speed data streams.
  - **Memory Constraints**: Limited memory resources require algorithms to maintain compact summaries of data streams.
  - **Dynamic Nature**: traffic patterns are non stationary, necessitating adaptive algorithms.

Online Identification of Heavy Hitters

Cornell University

- **Heavy Hitters**: In internet traffic analysis, heavy hitters refer to the most significant contributors to network traffic volume, such as top talkers or frequently accessed resources.
- **Online Identification**: In dynamic network environments, online identification methods are essential for real-time detection (e.g. to block DoS attacks).
- **Challenges**:
    - **Real-time Processing**: algorithms must operate in real-time, handling high-speed data streams.
    - **Memory Constraints**: Limited memory resources require algorithms to maintain compact summaries of data streams.
    - **Dynamic Nature**: traffic patterns are non stationary, necessitating adaptive algorithms.

Approaches for Identification of Heavy Hitters

- **Sketch-based Methods**: Utilize probabilistic data structures like Count-Min Sketch or Count Sketch to estimate heavy hitters with bounded memory usage.
- **Streaming Algorithms**: Process data streams incrementally, updating statistics and identifying heavy hitters on-the-fly.
- **Sampling Techniques**: Select a representative subset of traffic for analysis, reducing computational overhead while maintaining accuracy.

Because of the challenges in real time-processing only recently these detection algorithms have used more sophisticated, theory based, Bayesian methods. Most are heuristics.

Outline

Cornell University

Take-away messages

Cornell University

- Transportation and communication networks are examples of **cyber-physical infrastructures**.
- They can be described using a common language:
  - graphs,
  - flows,
  - capacities,
  - queues or delays,
  - sensing and feedback.
- Data analytics is central because it enables:
  1. **monitoring and inference**,
  2. **prediction**,
  3. **optimization and control**,
  4. **anomaly detection and resilience**.
- The modern challenge is no longer only how to build

Take-away messages

Cornell University

- Transportation and communication networks are examples of **cyber-physical infrastructures**.
- They can be described using a common language:
  - graphs,
  - flows,
  - capacities,
  - queues or delays,
  - sensing and feedback.
- Data analytics is central because it enables:
  1. **monitoring and inference**,
  2. **prediction**,
  3. **optimization and control**,
  4. **anomaly detection and resilience**.
- The modern challenge is no longer only how to build

Take-away messages

Cornell University

- Transportation and communication networks are examples of **cyber-physical infrastructures**.
- They can be described using a common language:
  - graphs,
  - flows,
  - capacities,
  - queues or delays,
  - sensing and feedback.
- Data analytics is central because it enables:
  1. **monitoring and inference**,
  2. **prediction**,
  3. **optimization and control**,
  4. **anomaly detection and resilience**.
- The modern challenge is no longer only how to build

Take-away messages

Cornell University

- Transportation and communication networks are examples of **cyber-physical infrastructures**.
- They can be described using a common language:
  - graphs,
  - flows,
  - capacities,
  - queues or delays,
  - sensing and feedback.
- Data analytics is central because it enables:
  1. **monitoring and inference**,
  2. **prediction**,
  3. **optimization and control**,
  4. **anomaly detection and resilience**.
- The modern challenge is no longer only how to build