Social and Communication Networks:

probabilistic models and algorithms

Laurent Massoulié

laurent.massoulie@inria.fr

Scope

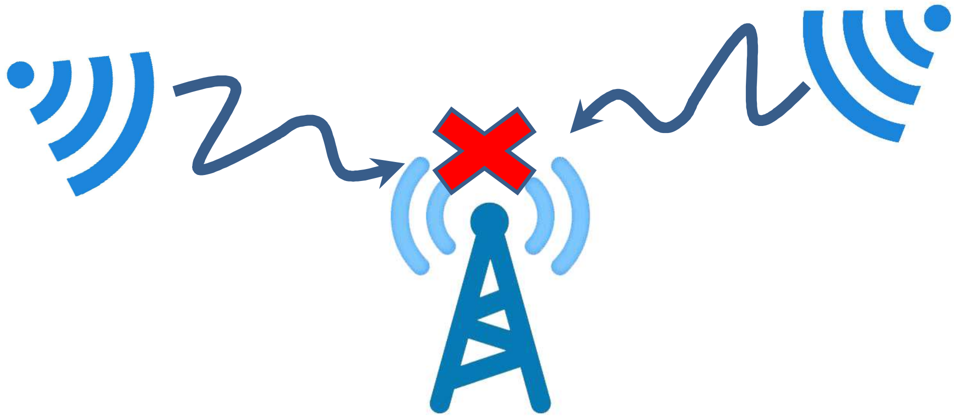
Communication Networks, Online Social Networks:

Algorithms for their control and optimization

→Modeling (probability, graphs)
→Analysis (Markov chains, Markov processes, optimization)

Tools applicable beyond chosen application domain

Communication Networks



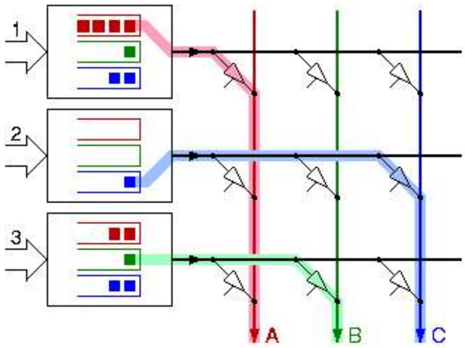
How to manage **collisions** (i.e. lost transmissions because of interference) between wireless transmitters

- ❑Aloha and Ethernet protocols
- →Markov chains and their long-term properties

Communication Networks

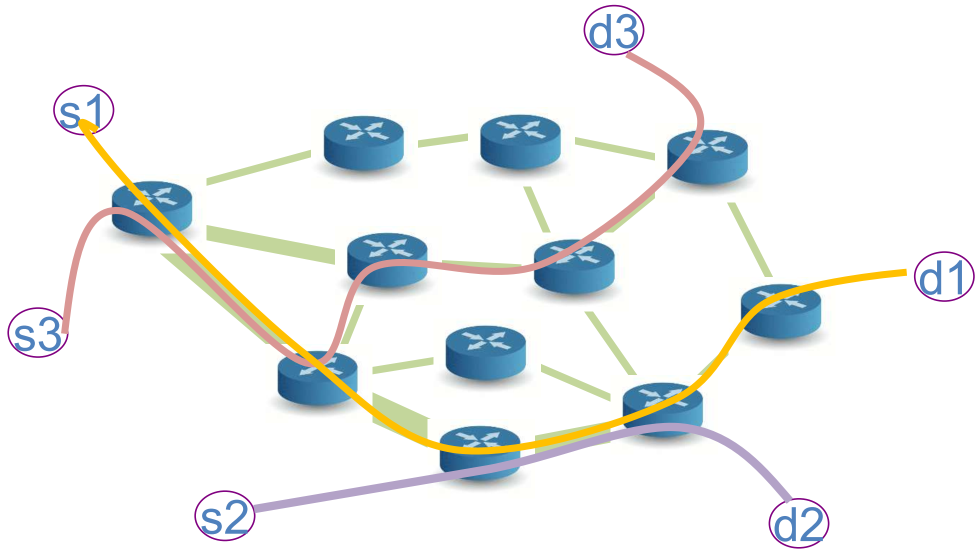
- ❑ How to schedule transmissions in switches, routers, and multi-hop wireless networks

Crossbar switch
with input queues:



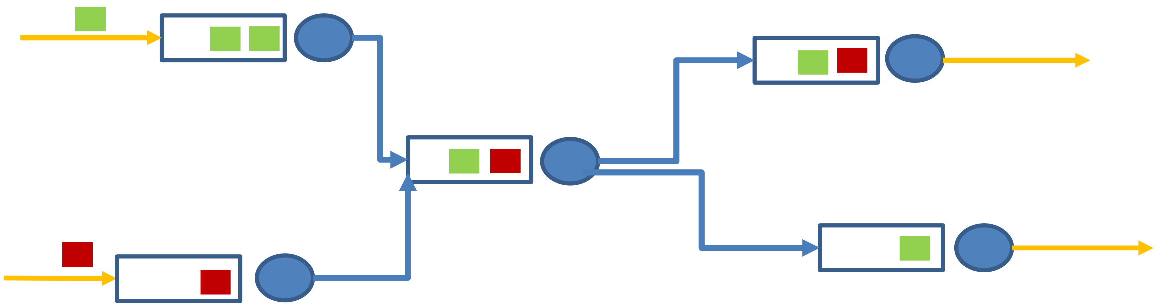
- ❑ Markov chains, Max-weight & backpressure algorithms

Communication Networks



- How to assign bandwidth in networks
  - Understanding TCP, the protocol regulating most Internet traffic
  - →Convex optimization & dynamical systems

Communication Networks

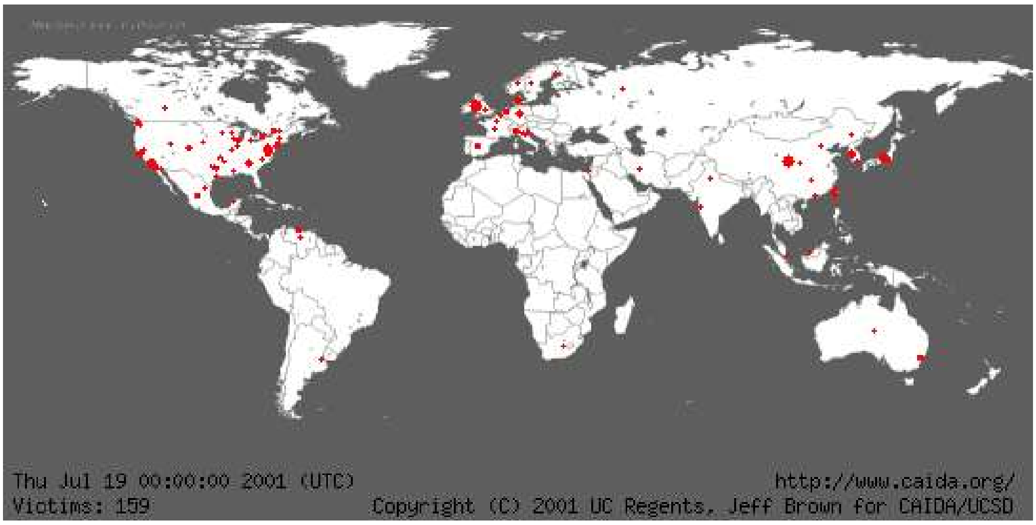


- Dynamics of queueing networks
  - Dimensioning of service systems (data centers, Velib,…); Justification of TCP; …
  - →Continuous time Markov processes, Poisson processes

Social Networks

epidemic-like propagation along a network

Examples: viruses, news, rumours, bank defaults,…



Spread of “CodeRed” Internet worm, 2001

Social Networks

epidemic-like propagation along a network

Examples: viruses, news, rumours, bank defaults,…

Spread of a picture on facebook
https://stamen.com/work/facebook-flowers/

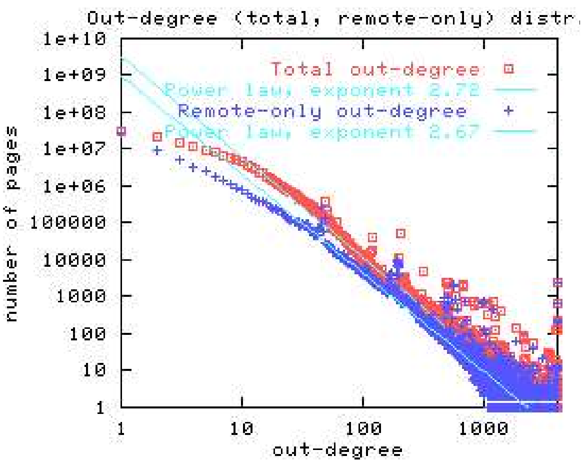
Social Networks

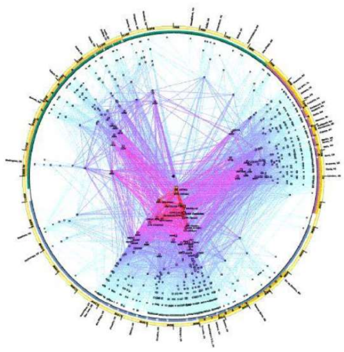
- ❑ What makes an epidemic potent or weak
  →random graphs, branching processes and phase transitions

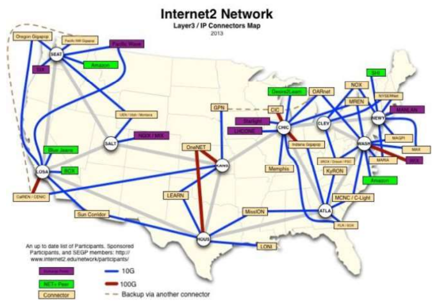
- ❑ What features of network topology affect epidemic outbreak
  →graph topology descriptors, comparison of Markov chains by “coupling”

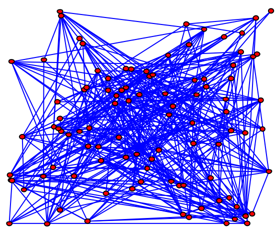
- ❑ How to maximize size of outbreak (e.g. for viral marketing)
  →NP-completeness, submodular functions and greedy maximization

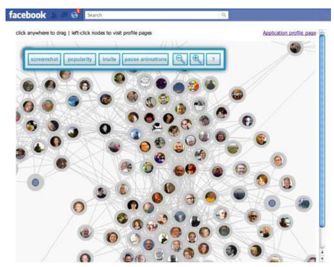
Social Networks





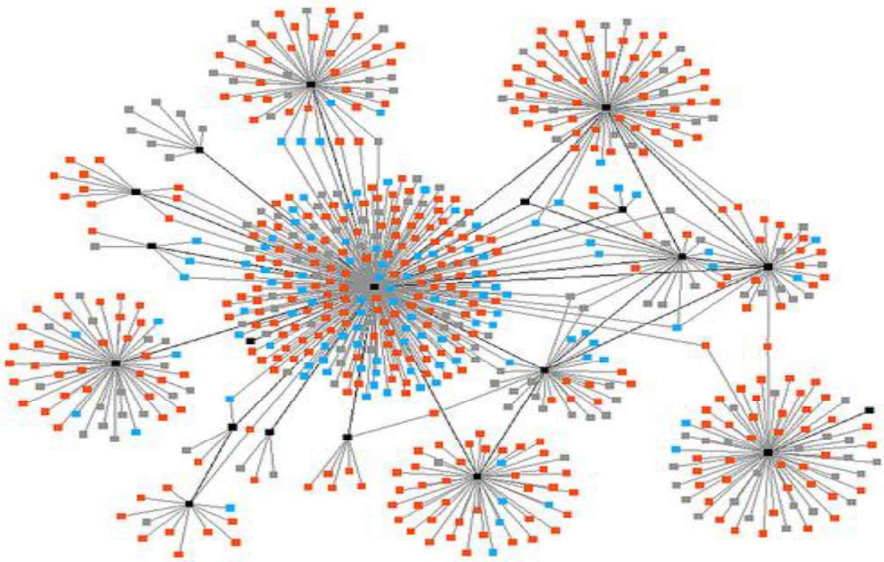






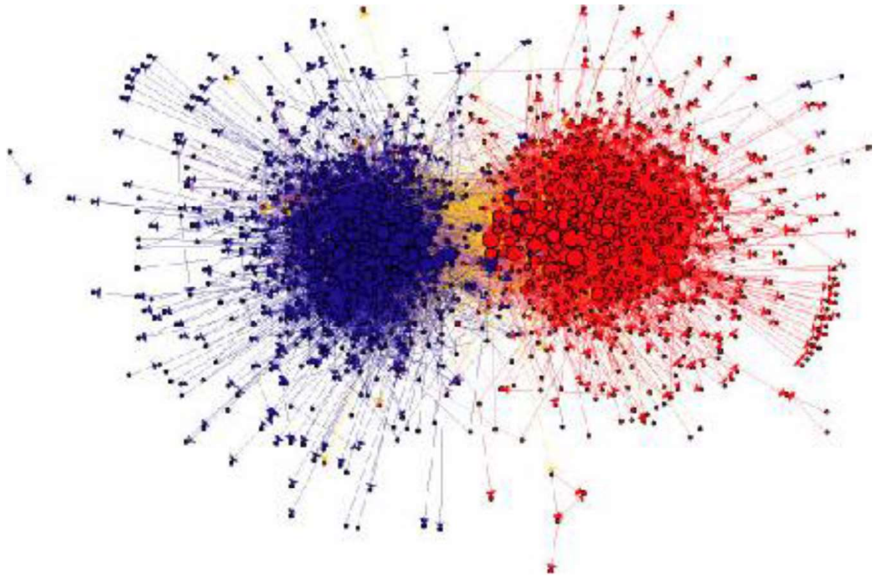
- ❑ Why are most networks “scale-free” (a.k.a. power-law)
  →Coupling and concentration inequalities

Social Networks



- What is a “small world” network
  - And how to search for information in it

Social Networks



Political blogs:
Republican vs Democrats

- How to find community structure and recommend contacts in a social network
  - →spectra of random graphs and spectral methods

Today:

- ❑Random access protocols for channel access
- ❑ Markov chains and their long-term behavior

Aloha: the first random access protocol for channel access [Abramson, Hawaii 70]

- ❑ Goal: allow machines on remote islands to transmit by radio to « master machine » without heavy coordination between them

Aloha: the first random access protocol for channel access [Abramson, Hawaii 70]

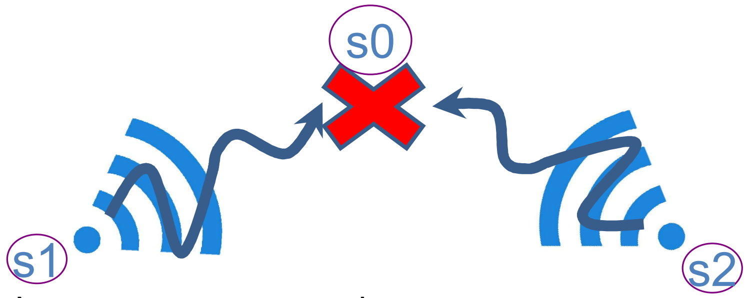
- ❑ Goal: allow machines on remote islands to transmit by radio to « master machine » without heavy coordination between them
- ❑ Key idea: use randomization for scheduling transmissions to avoid collisions between transmitters

Aloha: the first random access protocol for channel access [Abramson, Hawaii 70]

- ❑ Goal: allow machines on remote islands to transmit by radio to « master machine » without heavy coordination between them
- ❑ Key idea: use randomization for scheduling transmissions to avoid collisions between transmitters

→ A randomized, distributed algorithm

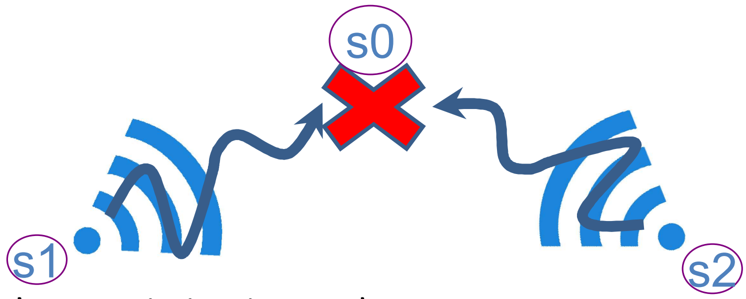
Aloha’s principle



Slotted time: fixed transmission intervals

Station with message to send: emits it with probability p

Aloha’s principle



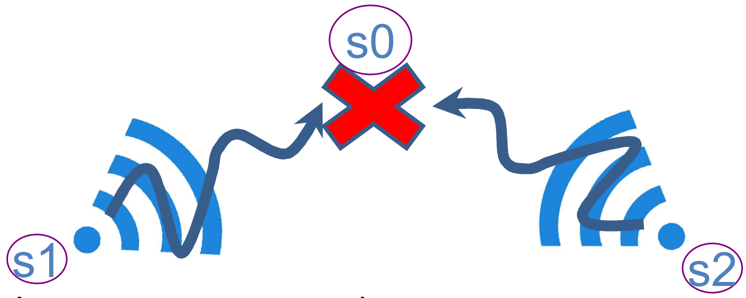
Slotted time: fixed transmission intervals

Station with message to send: emits it with probability p

By end of interval: learns whether msg successfully received, or not (due to collision or other interference)

Repeat until no message left to be sent

Aloha's principle



Slotted time: fixed transmission intervals

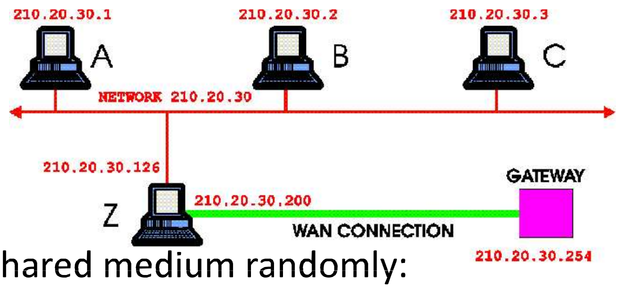
Station with message to send: emits it with probability p

By end of interval: learns whether msg successfully received, or not (due to collision or other interference)

Repeat until no message left to be sent

→ Minimal feedback (only listen for ack after having emitted)
→implicit coordination by receiver's acknowledgement

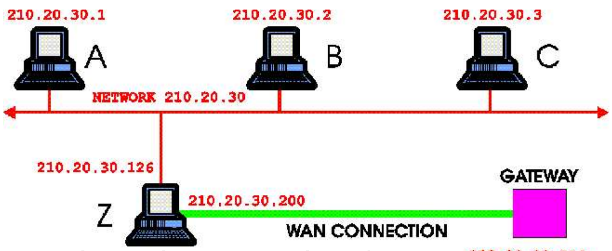
Ethernet principles [Metcalfe, Xerox PARC 73]



Machine emits on shared medium randomly:

After k failed attempts, waits before retransmitting for random number of slots picked uniformy from $\{1,2,\ldots,2^k\}$ (so-called contention window)

Ethernet principles [Metcalfe, Xerox PARC 73]

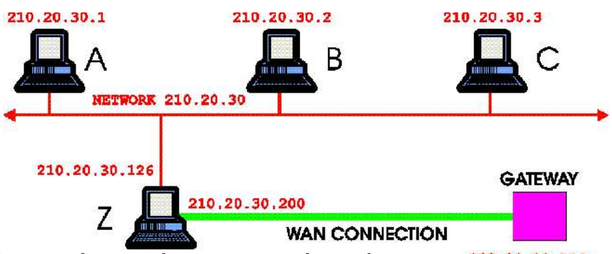


Machine emits on shared medium randomly:

After k failed attempts, waits before retransmitting for random number of slots picked uniformy from {1,2,…,2^k} (so-called contention window)

→The exponential backoff method, a refinement over Aloha

Ethernet principles [Metcalfe, Xerox PARC 73]



Machine emits on shared medium randomly:

After k failed attempts, waits before retransmitting for random number of slots picked uniformy from {1,2,…,2^k} (so-called contention window)

→The exponential backoff method, a refinement over Aloha

Other refinement: sense channel before transmitting (allows to compete by random access only during small fraction of total time)

Principles underly 802.11x (Wi-Fi) protocols

Goals

Understand performance of random access protocols

→for given ***traffic***, or ***workload*** offered to system (=process of message request arrivals),

Does system transmit them all?

Does it reach some steady state behaviour?

How long do transmissions take?

Outline

- ❑Introduction to Markov chain theory
  - ❑Fundamental notions (recurrence, irreducibility, ergodicity, transience)
  - ❑Criteria for ergodicity or transience
- ❑Performance of Random Access Protocols
  - ❑Aloha with finitely many stations
  - ❑Aloha with an infinite number of stations
  - ❑Results for Ethernet and other variants

Markov chains

- E a countable set (e.g., $\mathbb{N}$ or $[n] = \{1, \dots, n\}$)
- Definition: $\{X_n\}_{n\in\mathbb{N}}$ Markov chain with transition matrix P iff $\forall n > 0, \forall x_0^n = \{x_0, \dots, x_n\} \in E^{n+1}$,
  $\mathbb{P}(X_n = x_n | X_0^{n-1} = x_0^{n-1}) = \mathbb{P}(X_n = x_n | X_{n-1} = x_{n-1}) = p_{x_{n-1}x_n}$
  where $\forall x, y \in E,\ p_{xy} \geq 0$ and $\sum_{z\in E} p_{xz} = 1$
  (i.e. P is a *stochastic matrix*)
- Canonical example
  X_0 independent of $\{Y_n\}_{n\geq 0}$ an i.i.d. sequence, $Y_n \in E'$
  For some function $f : E \times E' \to E$,
  $$\forall n \geq 0,\ X_{n+1} = f(X_n, Y_n)$$
- Illustration: reflected Random Walk on $\mathbb{N}$:
  $X_{n+1} = \max(0, X_n + Y_n)$

Basic properties

- By induction $\mathbb{P}(X_n^{n+m} = x_n^{n+m}) = \mathbb{P}(X_n = x_n) \prod_{i=n+1}^{n+m} p_{x_{i-1}x_i}$
  $$\Rightarrow \mathbb{P}(X_0^{n+m} = x_0^{n+m} | X_n = x_n) \begin{array}{l} = \mathbb{P}(X_0^{n-1} = x_0^{n-1} | X_n = x_n) \times \cdots \\ \cdots \times \mathbb{P}(X_{n+1}^{n+m} = x_{n+1}^{n+m} | X_n = x_n) \end{array}$$
  (past and future independent conditionally on present)
- Noting $p_{x,y}^n = \mathbb{P}(X_n = y | X_0 = x)$, semi-group property:
  $$p_{xy}^{n+m} = \sum_{z\in E} p_{xz}^n p_{zy}^m$$
- Linear algebra interpretation
  For finite E (e.g. $E = [k]$), Matrix p^n = n-th power of P

Further properties

- Denote $\mathbb{P}_x(\cdot) = \mathbb{P}(\cdot|X_0 = x)$ distribution of chain started in state x at time 0
- Def: $T \in \mathbb{N} \cup \{+\infty\}$ **stopping time** iff $\forall n \in \mathbb{N}$, $\{T = n\}$ is $\sigma(X_0^n)$-measurable, i.e. $\exists \phi_n : E^{n+1} \to \{0,1\}$ such that $\mathbb{I}_{T=n} = \phi_n(X_0^n)$
- Key example $T_x := \inf\{n > 0 : X_n = x\}$
- **Strong Markov property**
  Markov chain X_0^∞ with transition matrix P, stopping time T
  Then conditionally on $T < +\infty$ and $X_T = x$,
  X_0^T and X_T^∞ independent with $X_T^\infty \sim \mathbb{P}_x$

Positive recurrence, null recurrence, transience, periodicity

State x is

- **recurrent** if $\mathbb{P}_x(T_x < +\infty) = 1$
- **positive recurrent** if $\mathbb{E}_x(T_x) < +\infty$
- **null recurrent** if $\mathbb{P}_x(T_x < +\infty) = 1$ & $\mathbb{E}_x(T_x) = +\infty$
- **transient** if not recurrent, i.e. $\mathbb{P}_x(T_x < +\infty) < 1$
- d-**periodic** if $d = \text{GCD}(n \geq 0 : p_{xx}^n > 0)$

ILLUSTRATION: reflected random walk on $\mathbb{N}$,
$S_{n+1} = \max(0, S_n + Y_n)$
State 0 is

- **positive recurrent** if $\mathbb{E}(Y_n) < 0$
- **transient** if $\mathbb{E}(Y_n) > 0$
- **null recurrent** if $\mathbb{E}(Y_n) = 0$ & $0 < \text{Var}(Y_n) < +\infty$

Decomposition of recurrent chains in cycles

Fix a state x that is recurrent ($\mathbb{P}_x(T_x < +\infty) = 1$),

Let $T_{x,k}$ = instant of k-th visit to state x

$\Rightarrow$ Trajectory X_1^∞: concatenation of cycles
$C_k := \{X_n\}_{T_{x,k} < n \le T_{x,k+1}}$

Strong Markov property $\Rightarrow$ cycles C_k are i.i.d.

Irreducibility

Markov chain is **irreducible** iff $\forall x, y \in E$,
$\exists n \in \mathbb{N}, x_0^n \in E^{n+1} \mid x_0 = x,\ x_n = y\ \&\ \prod_{i=1}^n p_{x_{i-1}x_i} > 0$

i.e., graph on E with directed edge (x, y) iff $p_{xy} > 0$ **strongly connected**

EXAMPLE
Standard random walk on graph G irreducible iff G connected

Proposition

For irreducible chain, if one state x is transient (resp. null recurrent, positive recurrent, d-periodic) then all are

Stationary measures

Non-negative measure π on E is **stationary** for P iff
$\forall x \in E, \pi_x = \sum_{y \in E} \pi_y p_{yx}$

Notation: $\mathbb{P}_\nu := \sum_{x \in E} \nu_x \mathbb{P}_x$ chain's distribution when $X_0 \sim \nu$

$\Rightarrow$ For stationary probability distribution π,
$\forall n > 0, \mathbb{P}_\pi(X_n^\infty \in \cdot) = \mathbb{P}_\pi(X_0^\infty \in \cdot)$

Limit theorems 1

Recurrence and stationary measures

Irreducible recurrent chain admits a stationary measure, unique up to multiplicative factor $\forall y \in E,\ \pi_y = \mathbb{E}_x \sum_{n=1}^{T_x} \mathbb{I}_{X_n = y}$

Irreducible chain admits a stationary probability distribution iff it is positive recurrent

Ergodic theorem

Irreducible, positive recurrent chain satisfies almost sure convergence

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} f(X_n) = \sum_{x \in E} \pi_x f(x)$$

for all π-integrable f, where $\pi =$ unique stationary distribution

Such chains are called ergodic

Limit theorems 2

Convergence in distribution

Ergodic, aperiodic chain satisfies $\forall x \in E, \lim_{n\to\infty} \mathbb{P}(X_n = x) = \pi_x$ where π: unique stationary distribution

"Converse"

Irreducible, non-ergodic chain satisfies $\forall x \in E, \lim_{n\to\infty} \mathbb{P}(X_n = x) = 0$

Foster-Lyapunov criterion for ergodicity

Theorem

An irreducible chain such that there exist $V : E \to \mathbb{R}_+$, a finite set $K \subset E$ and $\epsilon, b > 0$ satisfying

$$\mathbb{E}(V(X_{n+1}) - V(X_n) | X_n = x) \leq \begin{cases} -\epsilon, & x \notin K, \\ b - \epsilon, & x \in K, \end{cases}$$

is then ergodic.

Aloha with finitely many stations

Stations $s \in \mathcal{S}$, $|\mathcal{S}| < \infty$

- New arrivals at station s in slot n: $A_{n,s} \in \mathbb{N}$, $\{A_{n,s}\}_{n\geq 0}$ i.i.d.
- Probability of transmission by s if message in queue: p_s
- Source of randomness: $\{B_{n,s}\}_{n\geq 0}$ i.i.d., Bernoulli(p_s)
- Transmits iff $B'_{n,s} = 1$ where $B'_{n,s} = B_{n,s}\mathbb{I}_{L_{n,s}>0}$

Queue dynamics

$$L_{n+1,s} = L_{n,s} + A_{n,s} - B'_{n,s} \prod_{s'\neq s}(1 - B'_{n,s'})$$

Aloha with finitely many stations

Assume $\forall s,\ 0 < \mathbb{P}(A_{n,s} = 0) < 1$
Then chain $\{(L_{n,s})_{s\in\mathcal{S}}\}_{n\geq 0}$ is irreducible and aperiodic

Sufficient condition for ergodicity

$$\forall s, \mathbb{E}(A_{n,s}^2) < +\infty \text{ and } \lambda_s := \mathbb{E}(A_{n,s}) < p_s \prod_{s'\neq s}(1 - p_{s'})$$

Sufficient condition for transience

$$\forall s, \lambda_s > p_s \prod_{s'\neq s}(1 - p_{s'})$$

Aloha with finitely many stations

Symmetric case $\lambda_s = \lambda/|\mathcal{S}|$, $p_s \equiv p$:

Recurrence if $\lambda < |\mathcal{S}|p(1-p)^{|\mathcal{S}|-1}$

Transience if $\lambda > |\mathcal{S}|p(1-p)^{|\mathcal{S}|-1}$

$\Rightarrow$ To achieve stability (ergodicity) for fixed λ, need $p \to 0$ as $|\mathcal{S}| \to \infty$

Impractical! (Collisions take forever to be resolved)

Takeaway messages

- Markov chain theory: framework for system and algorithm performance analysis
- Ergodicity (stability) analysis:
  $\rightarrow$ Determines for what demands system stabilizes into steady state
  $\rightarrow$ A "first order" performance index (know when delays remain stable, not their magnitude)
- Foster-Lyapunov criteria to prove ergodicity