Mean-field methods: what can go wrong?

with some applications to bike-sharing systems and caching

Nicolas Gast (Inria)

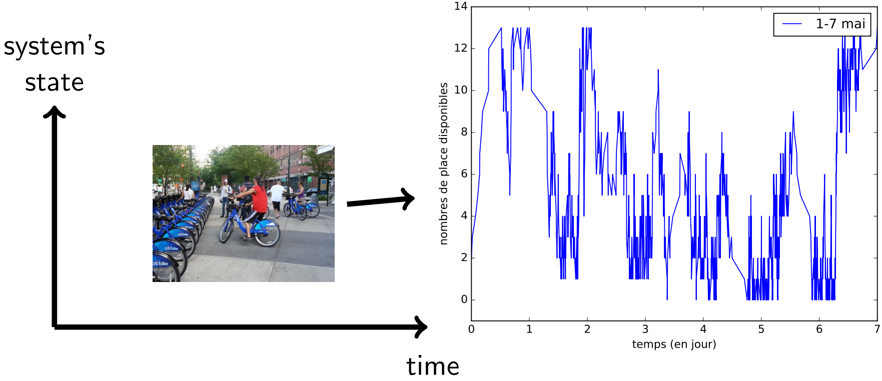
Inria, Grenoble, France

Grenoble's RO Summer School, July 2016

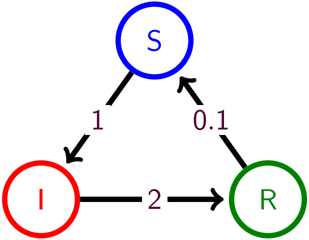
In this talk, we will study dynamical systems

From Wikipedia: *In mathematics, a dynamical system is a system in which a function describes the time dependence of a point in a geometrical space.*

From Wikipedia: *In mathematics, a dynamical system is a system in which a function describes the time dependence of a point in a geometrical space.*



Continuous time Markov chains: Three possible definitions



Transition graph

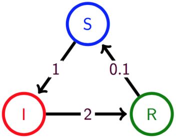
$$Q = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -2 & 2 \\ 0.1 & 0 & -0.1 \end{pmatrix}$$

Infinitesimal generator

$$\begin{aligned} \mathbb{P}(Z(t+dt)=j \mid Z(t)=i \wedge \text{ the past}) &= \mathbb{P}(Z(t+dt)=j \mid Z(t)=i) \\ &= Q_{ij}dt + o(dt) \qquad \text{if } i \neq j \end{aligned}$$

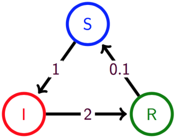
Markov property

Transient and steady-state analysis



$$Q = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -2 & 2 \\ 0.1 & 0 & -0.1 \end{pmatrix}$$

Transient and steady-state analysis



$$Q = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -2 & 2 \\ 0.1 & 0 & -0.1 \end{pmatrix}$$

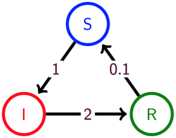
Transient analysis: the master equation

If X is a CTMC (continuous time Markov chain) with generator Q:

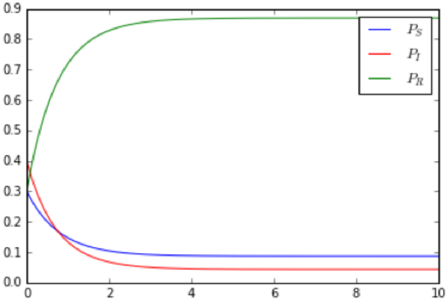
$$\frac{d}{dt}P_i(t) = \sum_{j \in S} P_j(t) Q_{ji},$$

where $P_i(t) = \mathbb{P}(X(t) = i)$.

Transient and steady-state analysis



$$Q = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -2 & 2 \\ 0.1 & 0 & -0.1 \end{pmatrix}$$



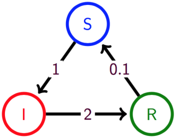
Transient analysis: the master equation

If X is a CTMC (continuous time Markov chain) with generator Q:

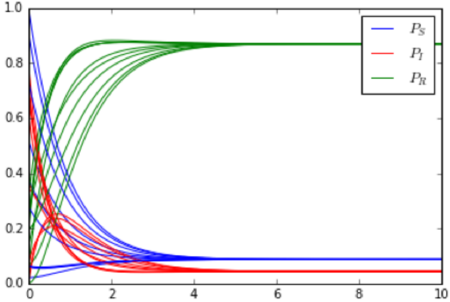
$$\frac{d}{dt}P(t) = P(t)Q,$$

where $P_i(t) = \mathbb{P}(X(t) = i)$.

Transient and steady-state analysis

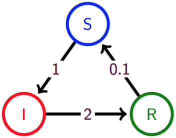


$$Q = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -2 & 2 \\ 0.1 & 0 & -0.1 \end{pmatrix}$$

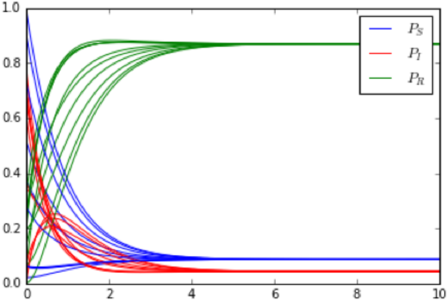


Steady-state analysis

Transient and steady-state analysis



$$Q = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -2 & 2 \\ 0.1 & 0 & -0.1 \end{pmatrix}$$

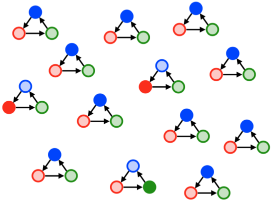


Steady-state analysis

If the chain is irreducible,

- The equation $\pi Q = 0$ has a unique solution such that $\sum_i \pi_i = 1$.
- $\lim_{i \to \infty} P_i(t) = \pi_i$

State space explosion and decoupling method



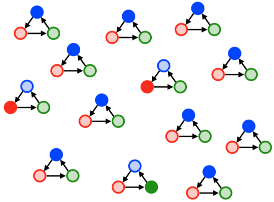
$3^{13} \approx 10^6$ states.

We need to keep track of S^N states

$$\mathbb{P}(Z_1(t) = i_1, \dots, Z_n(t) = i_n)$$

The generator Q has S^N entries.

State space explosion and decoupling method



$3^{13} \approx 10^6$ states.

We need to keep track of S^N states

$$\mathbb{P}(Z_1(t) = i_1, \dots, Z_n(t) = i_n)$$

The generator Q has S^N entries.

The decoupling assumption is

$$\underbrace{\mathbb{P}(Z_1(t) = i_1, \dots, Z_n(t) = i_n)}_{S^N \text{ variables}} \approx \underbrace{\mathbb{P}(Z_1(t) = i_1) \dots \mathbb{P}(Z_n(t) = i_n)}_{N \times S \text{ variables}}$$

Question: when is this (not) valid?

Outline

1. Population models and mean-field
2. The decoupling method: finite and infinite time horizon
   - Finite time horizon: some theory
   - Steady-state regime
   - Rate of convergence
3. Case-studies
   - Bike-sharing systems
   - Cache replacement policy
4. Conclusion and recap

Outline

1. Population models and mean-field

2. The decoupling method: finite and infinite time horizon
   - Finite time horizon: some theory
   - Steady-state regime
   - Rate of convergence

3. Case-studies
   - Bike-sharing systems
   - Cache replacement policy

4. Conclusion and recap

Mean field methods have been used in a multiple contexts

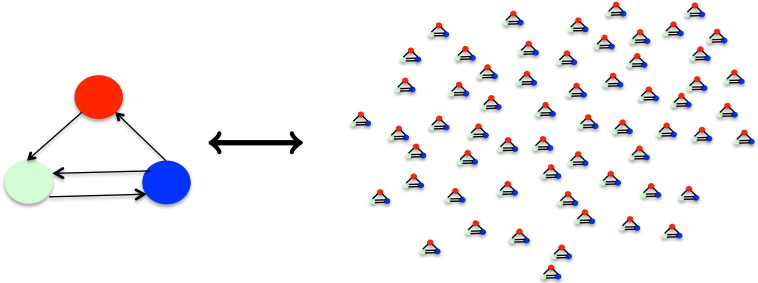
ex: model-checking, performance of SSD, load balancing, MAC protocol,...

- JAP 90 — On an index policy for restless bandits by Weber and Weiss
- SPAA 98 — Analyses of Load Stealing Models Based on Differential Equations by Mitzenmacher
- JSAC 2000 — Performance Analysis of the IEEE 802.11 Distributed Coordination Function by Bianchi
- FOCS 2002 — Load balancing with memory by Mitzenmacher et al.
- Ramaiyan et al — Fixed point analys is of single cell IEEE 802.11e WLANs: Uniqueness, multistability by ToN 2008
- SIGMETRICS 2013 — A mean field model for a class of garbage collection algorithms in flash-based solid state drives by Van Houdt
- EJTL 2014 — Incentives and redistribution in homogeneous bike-sharing systems with stations of finite capacities by Fricker and G.
- SIGMETRICS 2015 — Transient and Steady-state Regime of a Family of List-based Cache Replacement Algorithms by G. and Van Houdt

⋮ ⋮

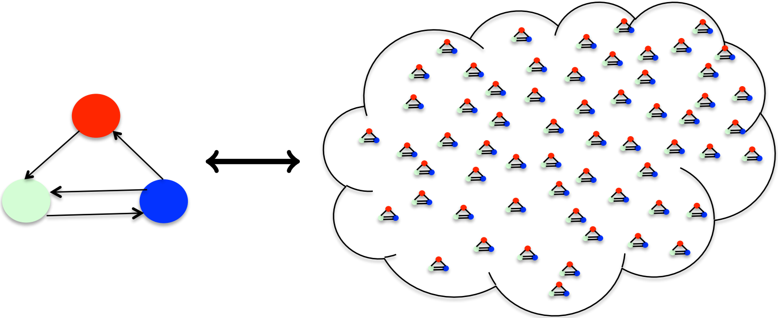
These models correspond to distributed systems

Each object interacts with the mass



These models correspond to distributed systems

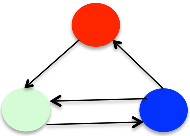
Each object interacts with the mass



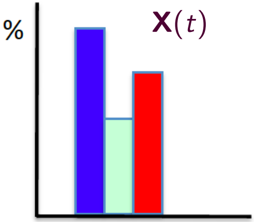
We view the population of objects more abstractly, assuming that individuals are indistinguishable.

These models correspond to distributed systems

Each object interacts with the mass



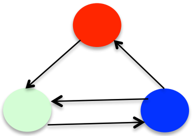
$\longleftrightarrow$

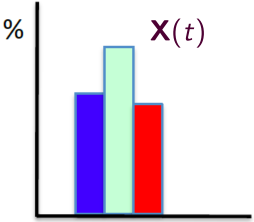


We view the population of objects more abstractly, assuming that individuals are indistinguishable.
An occupancy measure records the proportion of agents that are currently exhibiting each possible state.

These models correspond to distributed systems

Each object interacts with the mass

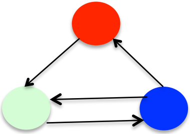




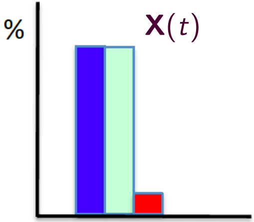
We view the population of objects more abstractly, assuming that individuals are indistinguishable.
An occupancy measure records the proportion of agents that are currently exhibiting each possible state.

These models correspond to distributed systems

Each object interacts with the mass



$\longleftrightarrow$



We view the population of objects more abstractly, assuming that individuals are indistinguishable.
An occupancy measure records the proportion of agents that are currently exhibiting each possible state.

Population CTMC

A population process is a sequence of CTMC $\mathbf{X}^N$, indexed by the population size N, with state spaces $\mathbf{E}^N \subset E \subset \mathbb{R}^d$ such that the transitions are (for $\ell \in \mathcal{L}$):

$$X \mapsto X + \frac{\ell}{N} \qquad \text{at rate } N\beta_\ell(X).$$

The drift is $f(x) = \sum_\ell \ell\beta_\ell(x)$.

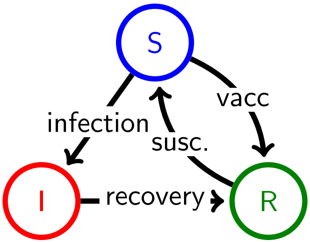
Population CTMC

A population process is a sequence of CTMC $\mathbf{X}^N$, indexed by the population size N, with state spaces $\mathbf{E}^N \subset E \subset \mathbb{R}^d$ such that the transitions are (for $\ell \in \mathcal{L}$):

$$X \mapsto X + \frac{\ell}{N} \qquad \text{at rate } N\beta_\ell(X).$$

The drift is $f(x) = \sum_\ell \ell\beta_\ell(x)$.

Example : SIRS model

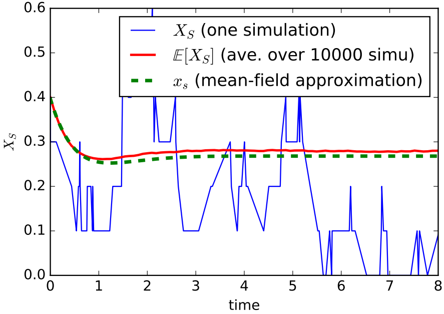


The state is (x_S, x_I, x_R). The transitions are

	ℓ	$\beta_\ell(x)$
Infection	$(-1,+1,0)$	$x_S + x_S x_I$
Recovery	$(0,-1,+1)$	x_I
Susceptible	$(+1,0,-1)$	x_R
Vaccination	$(-1,0,+1)$	x_S

Kurtz' convergence theorem

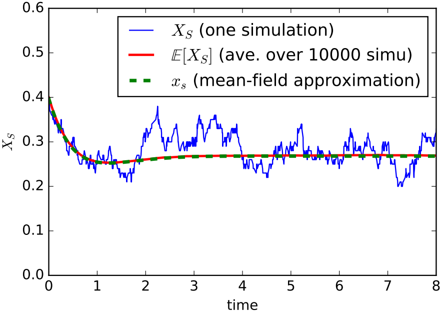
Theorem: Let $\mathbf{X}$ be a population process and assume that its drift f is Lipschitz-continuous and that $\sup_{\ell\in\mathcal{L}} |\ell| < \infty$. If $X^N(0)$ converges (in probability) to a point x, then the stochastic process $\mathbf{X}^N$ converges (in probability) to the solutions of the differential equation $\dot{x} = f(x)$, where f is the drift.



$N = 10$

Kurtz' convergence theorem

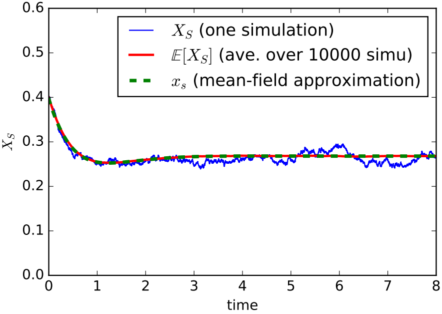
Theorem: Let $\mathbf{X}$ be a population process and assume that its drift f is Lipschitz-continuous and that $\sup_{\ell \in \mathcal{L}} |\ell| < \infty$. If $X^N(0)$ converges (in probability) to a point x, then the stochastic process $\mathbf{X}^N$ converges (in probability) to the solutions of the differential equation $\dot{x} = f(x)$, where f is the drift.



$N = 100$

Kurtz' convergence theorem

Theorem: Let $\mathbf{X}$ be a population process and assume that its drift f is Lipschitz-continuous and that $\sup_{\ell \in \mathcal{L}} |\ell| < \infty$. If $X^N(0)$ converges (in probability) to a point x, then the stochastic process $\mathbf{X}^N$ converges (in probability) to the solutions of the differential equation $\dot{x} = f(x)$, where f is the drift.



$N = 1000$

Outline

1. Population models and mean-field
2. The decoupling method: finite and infinite time horizon
   - Finite time horizon: some theory
   - Steady-state regime
   - Rate of convergence
3. Case-studies
   - Bike-sharing systems
   - Cache replacement policy
4. Conclusion and recap

Outline

1 Population models and mean-field

2 The decoupling method: finite and infinite time horizon
- Finite time horizon: some theory
- Steady-state regime
- Rate of convergence

3 Case-studies
- Bike-sharing systems
- Cache replacement policy

4 Conclusion and recap

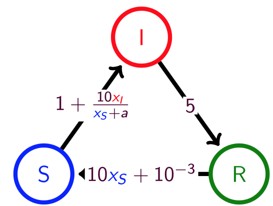
Decoupling and $\dot{x} = xQ(x)$

$$\mathbb{P}(Z_1(t) = i_1, \ldots, Z_n(t) = i_n) \approx \underbrace{\mathbb{P}(Z_1(t) = i_1)}_{=x_{1,i_1}(t)} \ldots \underbrace{\mathbb{P}(Z_n(t) = i_n)}_{=x_{n,i_n}(t)}$$

When we zoom on one object

$$\mathbb{P}(Z_1(t+dt) = j | Z_1(t) = i) \approx Q^{(1)}_{i,j}(\mathbf{x}(t))$$
$$:= \sum_{i_2 \ldots i_n, j_2 \ldots j_n} K_{(i,i_2\ldots i_n)\to(j,j_2\ldots j_n)} x_{2,i_2} \ldots x_{n,i_n}$$

We then get: $\frac{d}{dt} x_{1,j}(t) \approx \sum_i x_{1,i} Q^{(1)}_{i,j}(\mathbf{x}(t))$



Transient regime

For fixed t, the decoupling assumption is equivalent to the mean-field convergence.

Theorem (Snitzman (99), Kurtz (70'), Benaim, Le Boudec (08),...)

Let $\mathbf{X}^N$ be a population process such that the drift is Lipschitz-continuous. Then for any finite k:

$$\lim_{N\to\infty} \mathbf{P}\left[Z_1(t) = i_1 \dots Z_k(t) = i_k\right] = x_{i_1}(t) \dots x_{i_k}(t).$$

1 Population models and mean-field

2 The decoupling method: finite and infinite time horizon
- Finite time horizon: some theory
- Steady-state regime
- Rate of convergence

3 Case-studies
- Bike-sharing systems
- Cache replacement policy

4 Conclusion and recap

The fixed point method

Markov chain

Transient regime $\dot{p} = pK$

$t \to \infty$

Stationary $\pi K = 0$

The fixed point method

	Markov chain		Mean-field
Transient regime	$\dot{p} = pK$	$\xrightarrow{\quad N \to \infty \quad}$	$\dot{x} = xQ(x)$
	$\downarrow t \to \infty$		
Stationary	$\pi K = 0$	$\xrightarrow{\quad ? \quad}$	$\boxed{x^* Q(x^*) = 0}$ fixed points

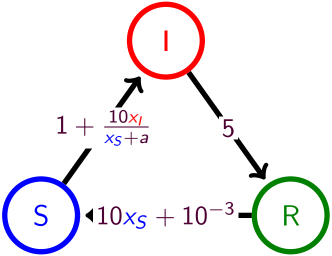
Method was used in many papers:

- Bianchi 00, Performance analysis of the IEEE 802.11 distributed coordination function.
- Ramaiyan et al. 08, Fixed point analys is of single cell IEEE 802.11e WLANs: Uniqueness, multistability.
- Kwak et al. 05, Performance analysis of exponenetial backoff.
- Kumar et al 08, New insights from a fixed-point analysis of single cell IEEE 802.11 WLANs.

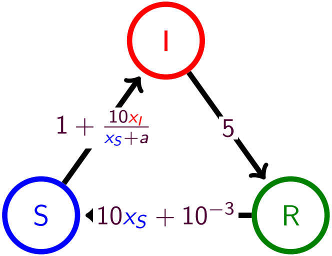
Does it always work?

SIRS model:

- A node S becomes I at rate 1 (external infection)
- When a S meets an I, it becomes infected at rate $1/(S+a)$
- An I recovers at rate 5.
- A node R becomes S by:
  - meeting a node S (rate $10S$)
  - alone (at rate 10^{-3}).

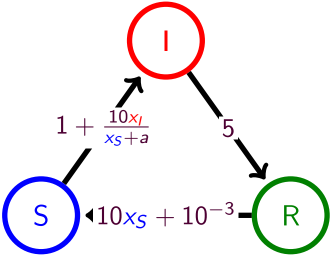


Does it always work?



- Markov chain is irreducible.
- Unique fixed point $x^*Q(x^*)=0$.

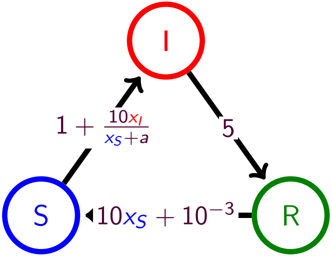
Does it always work?



- Markov chain is irreducible.
- Unique fixed point $x^* Q(x^*) = 0$.

	Fixed point $xQ(x) = 0$		Stat. measure $N = 10^3, 10^4 \dots$	
	x_S	x_I	π_S	π_I
$a = .3$	0.209	0.234	0.209	0.234

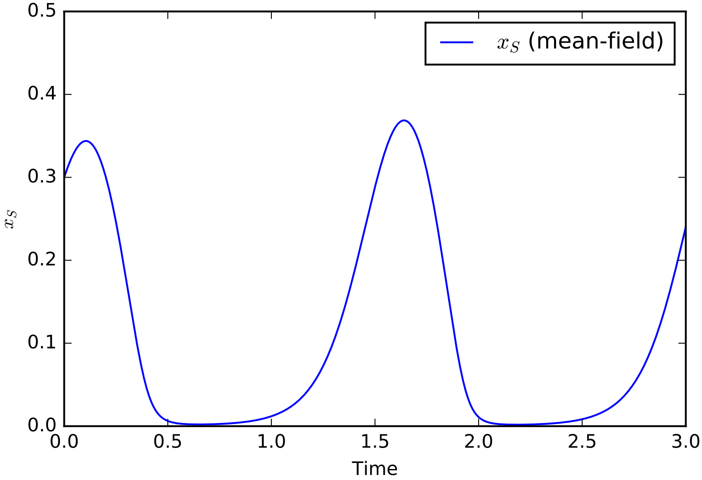
Does it always work?



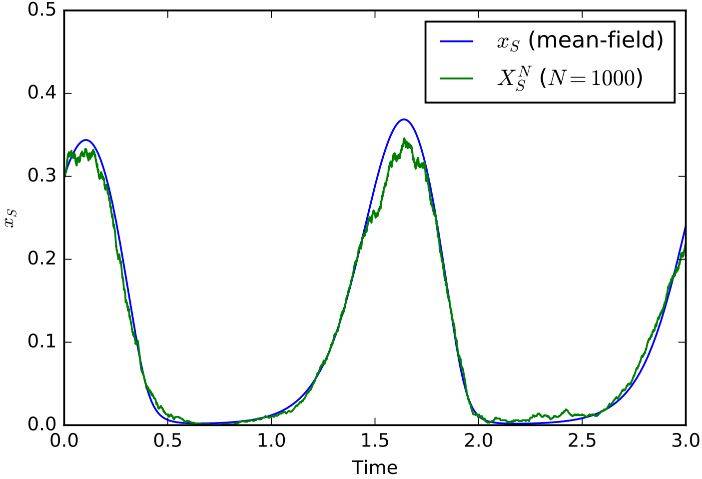
- Markov chain is irreducible.
- Unique fixed point $x^* Q(x^*) = 0$.

	Fixed point $xQ(x)=0$		Stat. measure $N = 10^3, 10^4 \ldots$	
	x_S	x_I	π_S	π_I
$a = .3$	0.209	0.234	0.209	0.234
$a = .1$	0.078	0.126	0.11	0.13

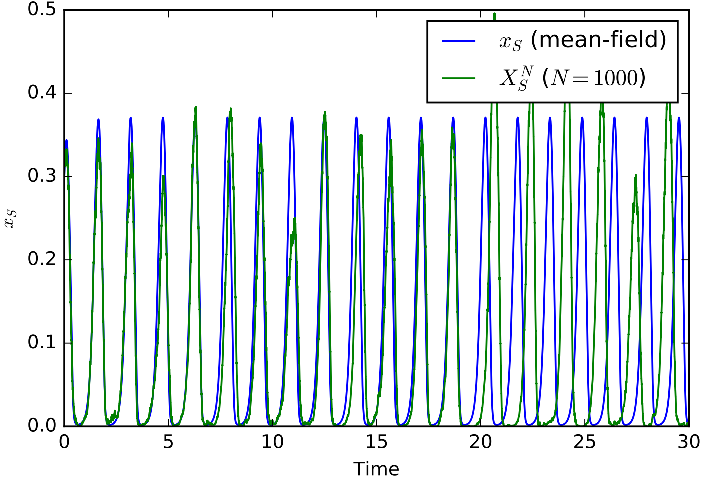
What happened?



What happened?

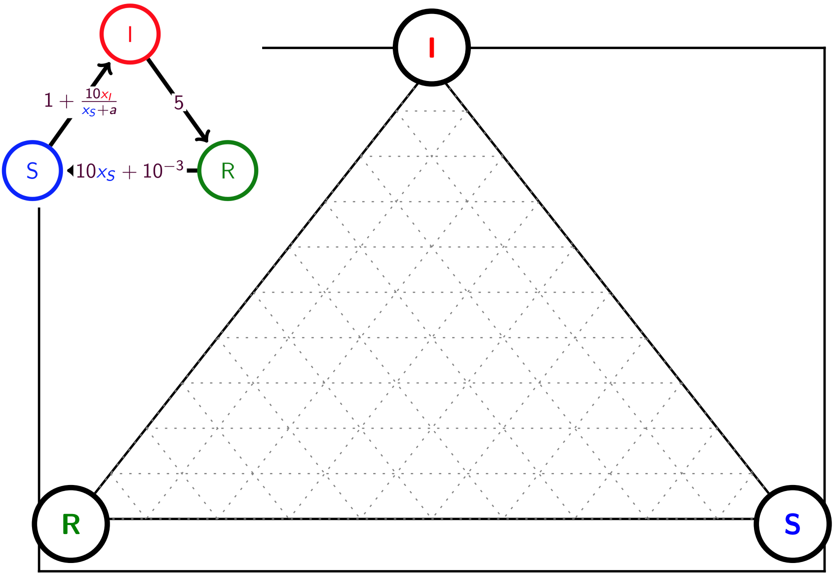


What happened?



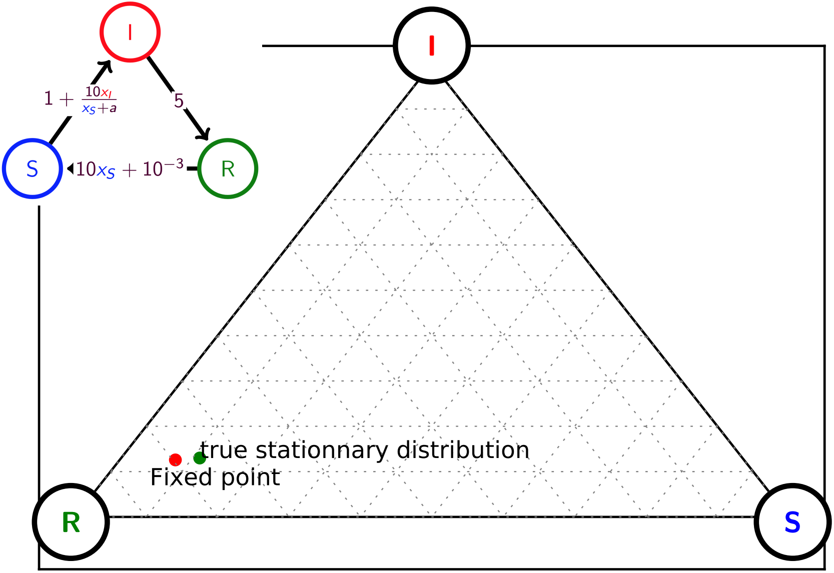
What happened?

$(x_S = 0.078, x_I = 0.126), (\pi_S = 0.11, \pi_I = 0.13)$



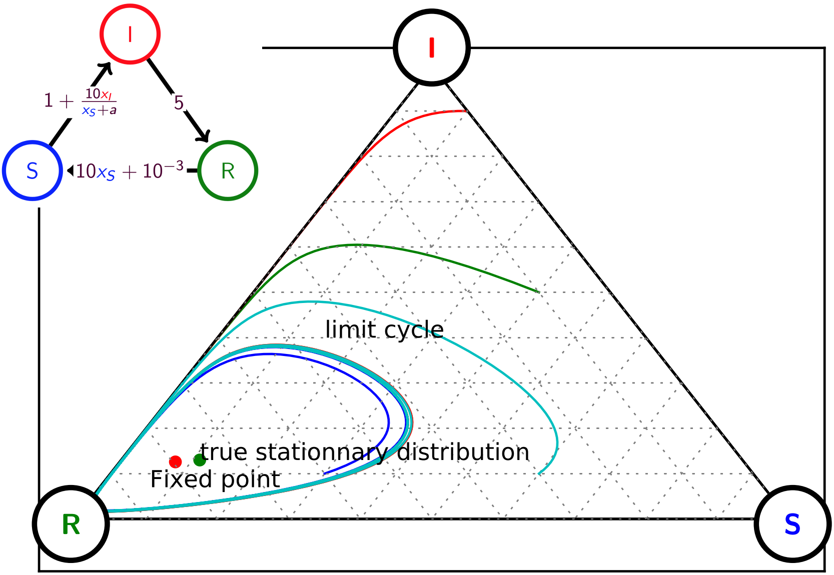
What happened?

$(x_S = 0.078, x_I = 0.126), (\pi_S = 0.11, \pi_I = 0.13)$



What happened?

$(x_S = 0.078, x_I = 0.126), (\pi_S = 0.11, \pi_I = 0.13)$



Fixed points?

	Markov chain
Transient regime	$\dot{p} = pK$
	$\Big\downarrow\, t \to \infty$
Stationary	$\pi K = 0$

Fixed points?

$$\begin{tikzcd}
 & \text{Markov chain} & & \text{Mean-field} \\
\text{Transient regime} & \dot{p} = pK \arrow[rr, "N \to \infty"] \arrow[d, "t \to \infty"] & & \dot{x} = xQ(x) \\
\text{Stationary} & \pi K = 0 \arrow[rr, "?"] & & \underbrace{\boxed{x^* Q(x^*) = 0}}_{\text{fixed points}}
\end{tikzcd}$$

Fixed points?

$$\begin{tikzcd}[column sep=huge]
& \text{Markov chain} & \text{Mean-field} \\
\text{Transient regime} & \dot{p} = pK \arrow[r, "N \to \infty"] \arrow[d, "t \to \infty"] & \dot{x} = xQ(x) \arrow[d, "\cancel{t \to \infty}"] \\
\text{Stationary} & \pi K = 0 \arrow[r, "\cancel{N \to \infty}"] & \underset{\text{fixed points}}{\boxed{\cancel{x^* Q(x^*) = 0}}}
\end{tikzcd}$$

Fixed points?

$$
\begin{tikzcd}[column sep=large]
 & \text{Markov chain} & & \text{Mean-field} \\
\text{Transient regime} & \dot{p} = pK \arrow[rr, "N \to \infty"] \arrow[d, "t \to \infty"] & & \dot{x} = xQ(x) \arrow[d, "\cancel{t \to \infty}"'] \arrow[d, bend left, "\text{if yes}"] \\
\text{Stationary} & \pi K = 0 \arrow[rr, "\cancel{N \to \infty}"] \arrow[rr, bend right, "\text{then yes}"'] & & \boxed{x^* Q(x^*) = 0} \\
 & & & \text{fixed points}
\end{tikzcd}
$$

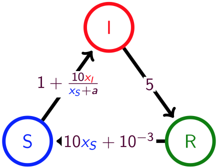
Theorem (Benaim Le Boudec 08)

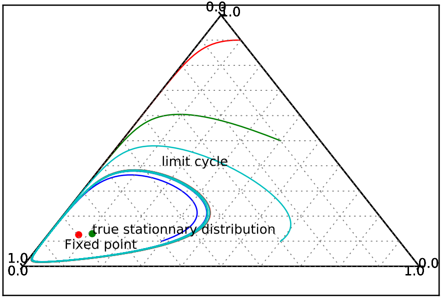
If all trajectories of the ODE converges to the fixed points, the stationary distribution π^N concentrates on the fixed points

In that case, we also have:

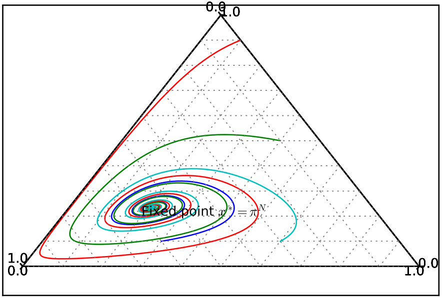
$$\lim_{N\to\infty} \mathbf{P}\left[Z_1 = i_1 \dots Z_k = i_k\right] = x_1^* \dots x_k^*.$$

Steady-state: illustration





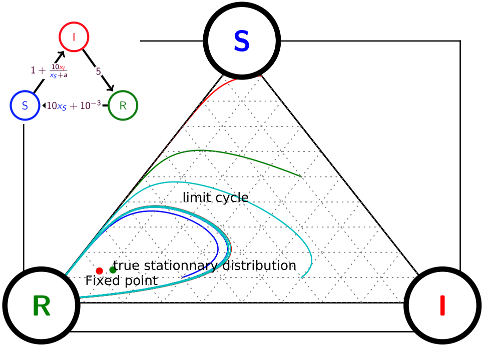
$a = .1$



$a = .3$

Quiz

Consider the SIRS model:



Under the stationary distribution π^N:

(A) As the trajectory converge to a fixed point, there is no such stationary distribution.

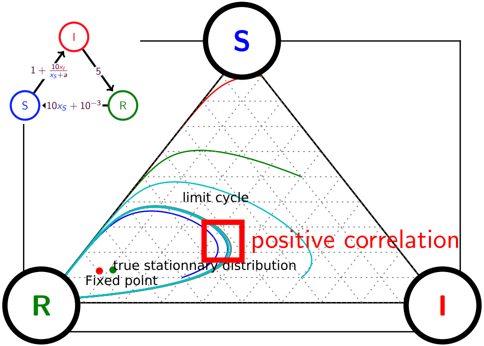
(B) $P(Z_1 = S, Z_2 = S) \approx P(Z_1 = S)P(Z_2 = S)$

(C) $P(Z_1 = S, Z_2 = S) > P(Z_1 = S)P(Z_2 = S)$

(D) $P(Z_1 = S, Z_2 = S) < P(Z_1 = S)P(Z_2 = S)$

Quiz

Consider the SIRS model:



Under the stationary distribution π^N:

(A) As the trajectory converge to a fixed point, there is no such stationary distribution.

(B) $P(Z_1 = S, Z_2 = S) \approx P(Z_1 = S)P(Z_2 = S)$

(C) $P(Z_1 = S, Z_2 = S) > P(Z_1 = S)P(Z_2 = S)$

(D) $P(Z_1 = S, Z_2 = S) < P(Z_1 = S)P(Z_2 = S)$

Answer: C

$P(Z_1(t) = S, Z_2(t) = S) = x_1(t)^2$. Thus: positively correlated.

How to show that trajectories converge to a fixed point?

Possible solution: find Lyapunov function [G. 2016]

How to show that trajectories converge to a fixed point?

Possible solution: find Lyapunov function [G. 2016]

A Lyapunov function if a function f such that

- Lower bounded: $\inf_x f(x) > +\infty$
- Decreasing along trajectories:

$$\frac{d}{dt} f(x(t)) < 0,$$

  whenever $x(t)Q(x(t)) \neq 0$.

If there exists a Lyapunov function, then $\dot{x} = xQ(x)$ converges to a fixed point $x^* Q(x^*) = 0$.

How to find a Lyapunov function

- Energy? Entropy? (or often: Luck)

Outline

1. Population models and mean-field
2. The decoupling method: finite and infinite time horizon
   - Finite time horizon: some theory
   - Steady-state regime
   - Rate of convergence
3. Case-studies
   - Bike-sharing systems
   - Cache replacement policy
4. Conclusion and recap

The rate of convergence is $O(1/\sqrt{N})$

Theorem

Let $\mathbf{X}$ be a population process such that its drift is L-Lipschitz-continuous. Then: if $X^N(0) = x_0$:

$$\mathbb{E}\left[\sup_{t\le T}\left\|X^N(t) - x(t)\right\|\right] \le O\Big(\frac{1}{\sqrt{N}}\Big)e^{LT}.$$

Note: we also have

$$|\mathbf{P}\left[Z(t) = i\right] - x_i(t)| = O(1/N).$$

Can be extended to:

- Steady-state
- Non-homogeneous objects.
- Non-smooth dynamics

A martingale argument

Recall that the transitions are $x \mapsto x + \ell/N$ at rate $N\beta_\ell(x)$. Then, $f(x) = \sum_\ell \ell\beta_\ell(x)$ satisfies:

$$\lim_{dt\to 0} \frac{1}{dt}\mathbb{E}\left[X(t+dt) - X(t) | X(t) = x\right] = f(x)$$

$$\lim_{dt\to 0} \frac{1}{dt}\mathrm{var}\left[X(t+dt) - X(t) - f(X(t)) | X(t) = x\right] \leq C/N$$

This means that:

$$M(t) = X(t) - (x_0 - \int_0^t f(X(s))ds)$$

is such that:

$$\underbrace{\mathbb{E}\left[M(t) \mid \mathcal{F}_s\right] = M(s)}_{M(t) \text{ is a martingale}} \qquad \wedge \qquad \underbrace{\mathrm{var}\left[M(t)\right] \leq Ct/N}_{\text{Small variance}}.$$

By Doob's inequality:

$$\mathbf{P}\left[\sup_{t\leq T} \|M(t)\| \geq \epsilon\right] \leq \frac{C}{N\epsilon^2}.$$

Mean-field convergence

We then have

$$X(t) = x_0 + \int_0^t f(X(s))ds + \underbrace{M(t)}_{\text{small by previous slide}}$$

Let $x(t)$ be the solution of the ODE $\dot{x} = f(x)$ such that $x(0) = x_0$.

Gronwall's Lemma

If f is Lipschitz-continuous, then

$$\sup_{t \le T} \|X(t) - x(t)\| \le \sup_{t \le T} \|M(t)\| \, e^{LT}.$$

Recap

- Decoupling $\approx$ mean-field convergence
- If the rates are continuous, convergence always holds for the transient regime
- The stationary regime should be handle with care
  - The uniqueness of the fixed point is not enough.
  - Lyapunov functions can help but are not easy to find.

Outline

1 Population models and mean-field

2 The decoupling method: finite and infinite time horizon
- Finite time horizon: some theory
- Steady-state regime
- Rate of convergence

3 Case-studies
- Bike-sharing systems
- Cache replacement policy

4 Conclusion and recap

Outline

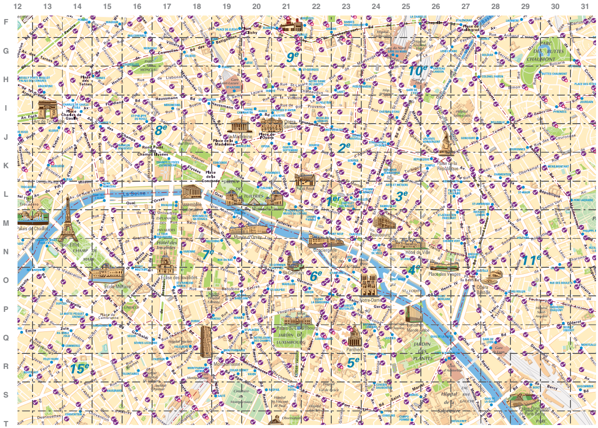
1. Population models and mean-field

2. The decoupling method: finite and infinite time horizon
   - Finite time horizon: some theory
   - Steady-state regime
   - Rate of convergence

3. Case-studies
   - Bike-sharing systems
   - Cache replacement policy

4. Conclusion and recap

Bike-sharing systems

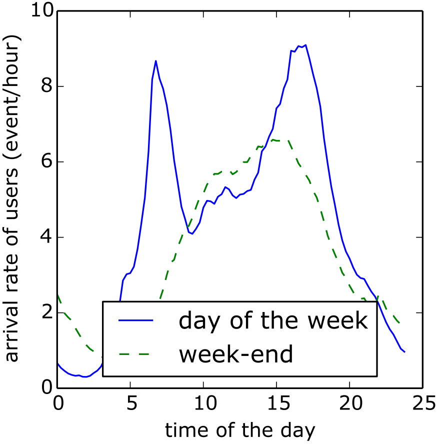


Empty station

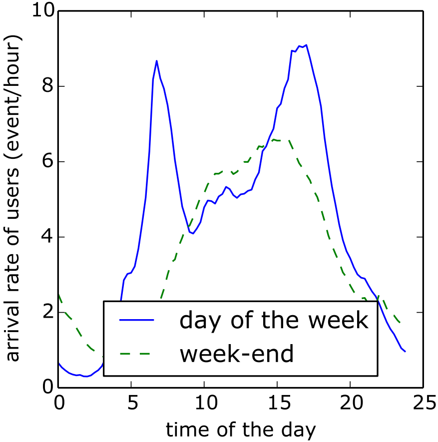
Full station

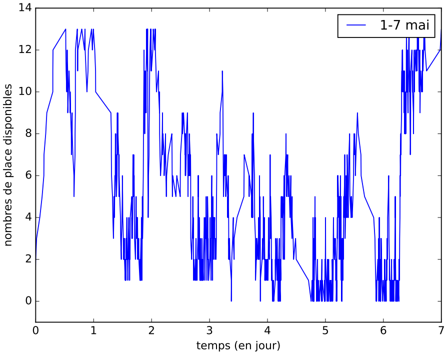
- Each station has a given number of parking slots.
- Users enter the system by picking up a bike at a station and making a trip to another station, where they drop the bike on an available parking spot.

A time-varying system



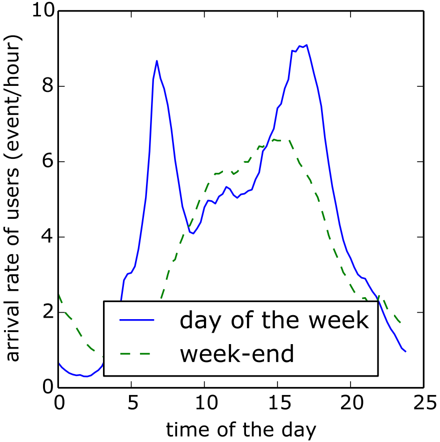
A time-varying and stochastic system

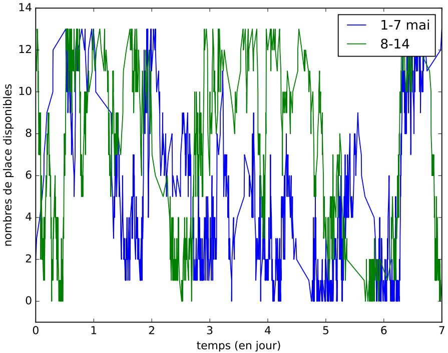




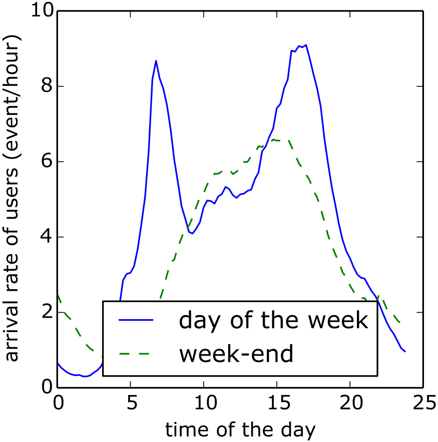
Gare de l'Est

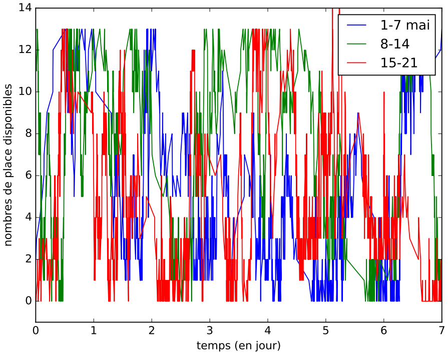
A time-varying and stochastic system



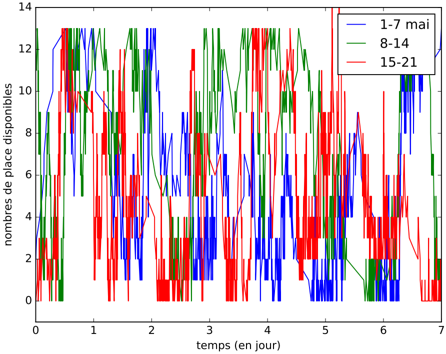


A time-varying and stochastic system





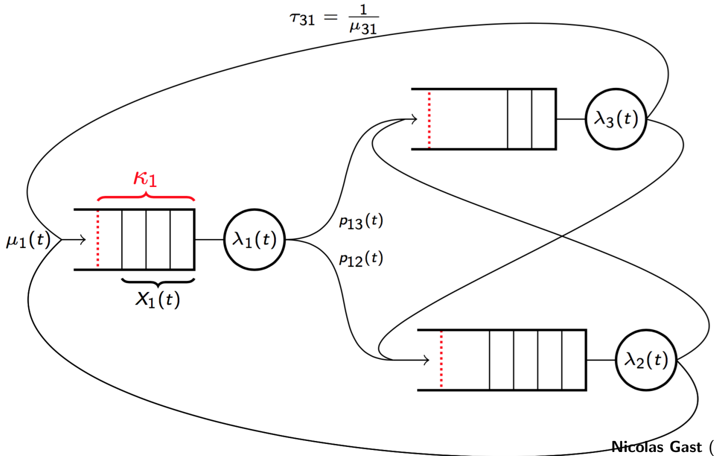
We need stochastic forecasts



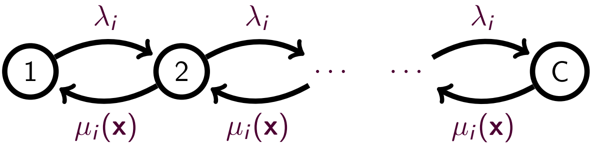
Exercise

Assuming independence, write down an approximation for

$$\mathbf{P}(k \text{ bikes are parked at a given station})$$



Solution: a time-inhomogeneous CTMC per station



$$\begin{aligned}\mu_i(x) &= p_i\mu\#\{\text{bike circulating}\}\\ &= p_i\mu\left(\#\{\text{total bikes}\} - \sum_{i\in\text{stations}}\sum_{k=1}^{C_i} kx_{i,k}\right)\end{aligned}$$

Two types of results

BSS model

theoretical

Sizing, incentives

[Fricker G. 2014, Fricker et al. 2013]

data-analysis

Prediction

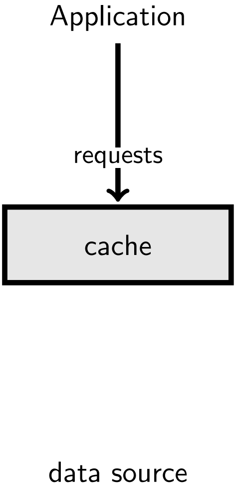
[G. et al 2015]

Outline

1. Population models and mean-field
2. The decoupling method: finite and infinite time horizon
   - Finite time horizon: some theory
   - Steady-state regime
   - Rate of convergence
3. Case-studies
   - Bike-sharing systems
   - Cache replacement policy
4. Conclusion and recap

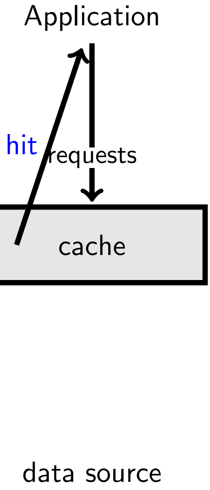
A cache-replacement policy

G. Van Houdt, 2015



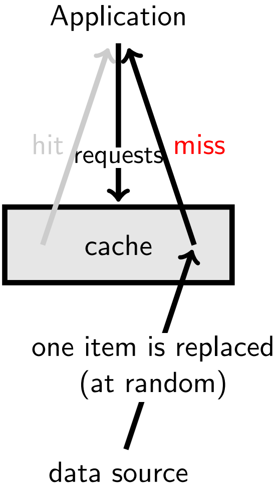
A cache-replacement policy

G. Van Houdt, 2015



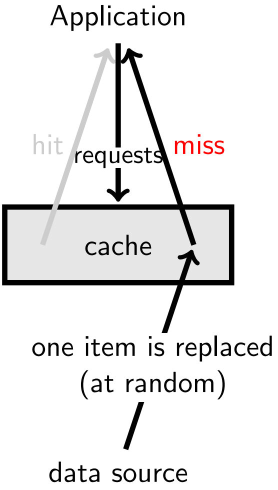
A cache-replacement policy

G. Van Houdt, 2015



A cache-replacement policy

G. Van Houdt, 2015



Model:

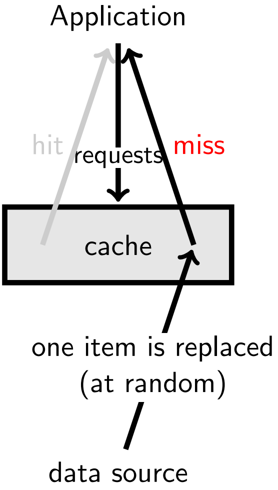
- Items have the same size.
- Cache can store m items.
- There are n items. Item i is requested with probability p_i.

Exercise

- By using the independence assumption, find an approximation for $\mathbb{P}(\text{item i is in cache at time } t)$.

A cache-replacement policy

G. Van Houdt, 2015



Markov model

State space : set of m distinct items.

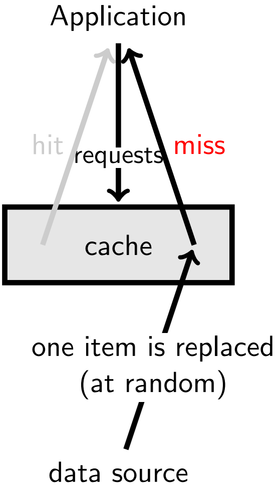
Transitions:

$$\{i_1 \dots i_m\} \mapsto \{i_1 \dots i_{k-1}, j, i_{k+1} \dots i_n\}$$

with probability p_j/m.

A cache-replacement policy

G. Van Houdt, 2015



Markov model

State space : set of m distinct items.

Transitions:

$$\{i_1 \dots i_m\} \mapsto \{i_1 \dots i_{k-1}, j, i_{k+1} \dots i_n\}$$

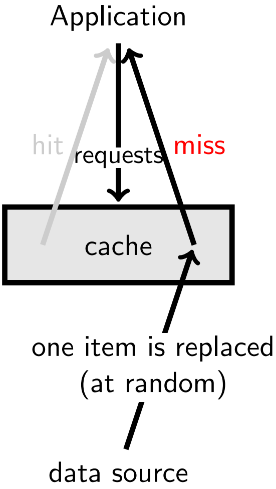
with probability p_j/m.

Decoupling assumption

$$\mathbb{P}(i_1 \dots i_m) \approx \underbrace{\mathbb{P}(i_1)}_{=:x_{i_1}} \dots \mathbb{P}(i_m)$$

A cache-replacement policy

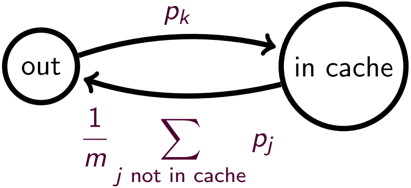
G. Van Houdt, 2015



Decoupling assumption

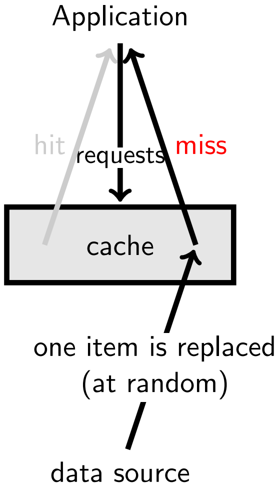
$$\mathbb{P}(i_1 \dots i_m) \approx \underbrace{\mathbb{P}(i_1)}_{=:x_{i_1}} \dots \mathbb{P}(i_m)$$

If we zoom on object k:



A cache-replacement policy

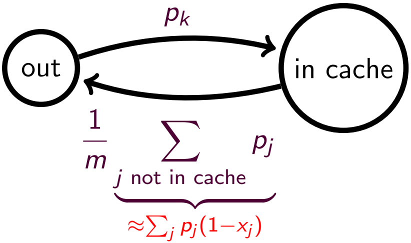
G. Van Houdt, 2015



Decoupling assumption

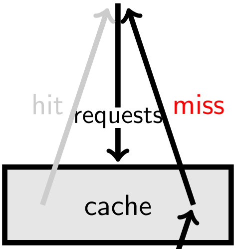
$$\mathbb{P}(i_1 \dots i_m) \approx \underbrace{\mathbb{P}(i_1)}_{=:x_{i_1}} \dots \mathbb{P}(i_m)$$

If we zoom on object k:

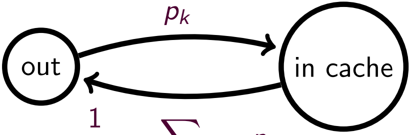


A cache-replacement policy

G. Van Houdt, 2015



If we zoom on object k:



$$\frac{1}{m} \underbrace{\sum_{j \text{ not in cache}} p_j}_{\approx \sum_j p_j(1-x_j)}$$

Mean-field model

Let $x_k := \mathbb{P}(\text{item } k \text{ is in the cache})$.

$$\dot{x}_k = p_k(1 - x_k) - \frac{\sum_\ell (p_\ell(1-x_\ell))}{m} x_k.$$

g

A cache-replacement policy: simulation

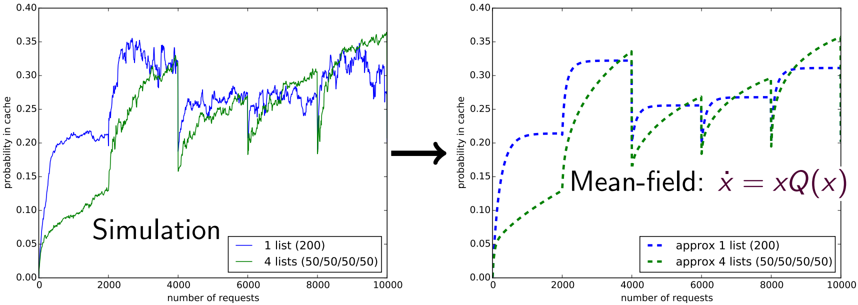


Figure: Popularities of objects change every 2000 steps.

A cache-replacement policy: simulation

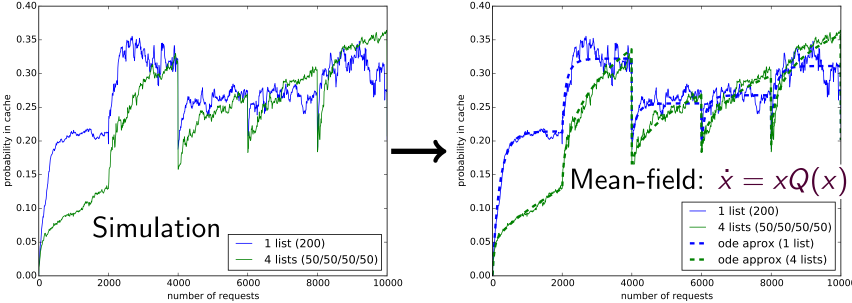


Figure: Popularities of objects change every 2000 steps.

Stationary distribution

Fixed point equation

- $0 = \dot{x}_k = p_k(1 - x_k) - \frac{\sum_\ell (p_\ell(1-x_\ell))}{m} x_k.$
- $\sum_k x_k = m.$

Stationary distribution

Fixed point equation

- $0 = \dot{x}_k = p_k(1 - x_k) - \frac{\sum_\ell (p_\ell(1-x_\ell))}{m} x_k$.
- $\sum_k x_k = m$.

Algorithm: easy to solve:

1. Define $x_k(T)$ the solution of $p_k(1 - x_k) - Tx_k$.
   - $x_k(T) = p_k/(1 + T)$
2. Find T such that $\sum_k (1 - x_k(T)) = m$.

Outline

1 Population models and mean-field

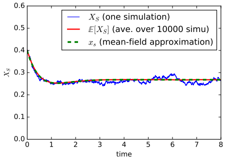
2 The decoupling method: finite and infinite time horizon
- Finite time horizon: some theory
- Steady-state regime
- Rate of convergence

3 Case-studies
- Bike-sharing systems
- Cache replacement policy

4 Conclusion and recap

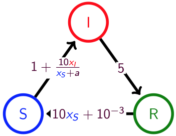
Recap

Mean field methods are useful to study large stochastic systems.



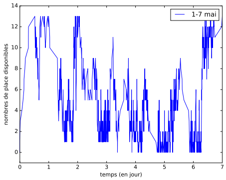
Mean-field ≈ decoupling assumption

- Valid for finite time.
- Infinite horizon should be handle with care



Applications:

- Give ideas on how to construct models
- Provide good approximations



Extensions: centralized optimization (OK), mean-field game (not that OK, see tomorrow)

Thank you!

http://mescal.imag.fr/membres/nicolas.gast

nicolas.gast@inria.fr

Mean-field and decoupling

Benaïm, Le Boudec 08 — *A class of mean field interaction models for computer and communication systems*, M.Benaïm and J.Y. Le Boudec., Performance evaluation, 2008.

Le Boudec 10 — *The stationary behaviour of fluid limits of reversible processes is concentrated on stationary points.*, J.-Y. L. Boudec. , Arxiv:1009.5021, 2010

Darling Norris 08 — *R. W. R. Darling and J. R. Norris*, Differential equation approximations for Markov chains, Probability Surveys 2008

G. 16 — *Construction of Lyapunov functions via relative entropy with application to caching*, Gast, N., ACM MAMA 2016

Budhiraja et al. 15 — *Limits of relative entropies associated with weakly interacting particle systems.*, A. S. Budhiraja, P. Dupuis, M. Fischer, and K. Ramanan. , Electronic journal of probability, 20, 2015.

References (continued)

Optimal control and mean-field games:

G.,Gaujal Le Boudec 12 — *Mean field for Markov decision processes: from discrete to continuous optimization*, N.Gast,B.Gaujal,J.Y.Le Boudec, IEEE TAC, 2012

G. Gaujal 12 — *Markov chains with discontinuous drifts have differential inclusion limits.*, Gast N. and Gaujal B., Performance Evaluation, 2012

Puterman — *Markov decision processes: discrete stochastic dynamic programming*, M.L. Puterman, John Wiley & Sons, 2014.

Lasry Lions — *Mean field games*, J.-M. Lasry and P.-L. Lions, Japanese Journal of Mathematics, 2007.

Tembine at al 09 — *Mean field asymptotics of markov decision evolutionary games and teams*, H. Tembine, J.-Y. L. Boudec, R. El-Azouzi, and E. Altman., GameNets 00

Applications: caches, bikes

Don and Towsley — *An approximate analysis of the LRU and FIFO buffer replacement schemes*, A. Dan and D. Towsley., SIGMETRICS 1990

G. Van Houdt 15 — *Transient and Steady-state Regime of a Family of List-based Cache Replacement Algorithms.*, Gast, Van Houdt., ACM Sigmetrics 2015

Fricker-Gast 14 — *Incentives and redistribution in homogeneous bike-sharing systems with stations of finite capacity.*, C. Fricker and N. Gast. , EJTL, 2014.

Fricket et al. 13 — *Mean field analysis for inhomogeneous bike sharing systems*, Fricker, Gast, Mohamed, Discrete Mathematics and Theoretical Computer Science DMTCS

G. et al 15 — *Probabilistic forecasts of bike-sharing systems for journey planning*, N. Gast, G. Massonnet, D. Reijsbergen, and M. Tribastone, CIKM 2015