Performance Evaluation

A not so Short Introduction
Stochastic Modeling of Computer Systems (2)

Jean-Marc Vincent[1,2]

[1] Laboratoire LIG, projet Inria-Mescal
UniversitéJoseph Fourier
Jean-Marc.Vincent@imag.fr
[2] LICIA
Laboratoire International de Calcul Intensif
et d'Informatique Ambiante

CAPES/COFECUB Project
Gestion de Ressources pour le Calcul Parallèle sur Grilles

2011 June 1

Outline

Traffic model

Continuous time modeling : occurence of events

- request on a database
- hit on web servers
- messages on a link
- phone calls
- ...

Randomness due to complexity of the environment
Superposition of many behaviors

$$\{N_t\}_{t\in\mathbb{R}}$$

N_t = number of arrivals of events in $[0, t[$

Traffic performance

Communication model : 2 counting processes
- emission process
- reception process associated to the emission process

Throughput

$$\lambda = \lim_{t \to +\infty} \frac{1}{t} N_t.$$

Data transfer
Streaming
Link capacity

Jitter

Variability of inter-arrivals
Voice transfer

Latency

Response time
Round trip time
Real-time applications

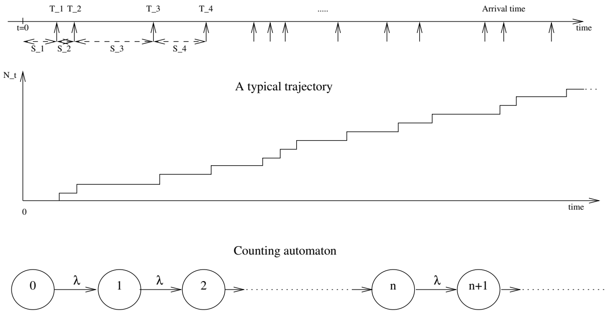
Loss rates

Communication reliability

$$\lambda_{emission} - \lambda_{reception}$$

Outline

Counting process



Macroscopic modeling

Definition

Macroscopic definition A continuous time stochastic process $\{N_t\}_{t \in \mathbb{R}^+}$ is a counting Poisson process with intensity λ iff
i) $N_0 = 0$
ii) $\{N_t\}_{t \in \mathbb{R}^+}$ have independent increments
iii) The number of events occurring in a time interval $]a, b]$ is Poisson distributed with parameter $\lambda(b-a)$;

$$\mathbb{P}(N_b - N_a = k) = e^{-\lambda(b-a)} \frac{(\lambda(b-a))^k}{k!}.$$

Properties

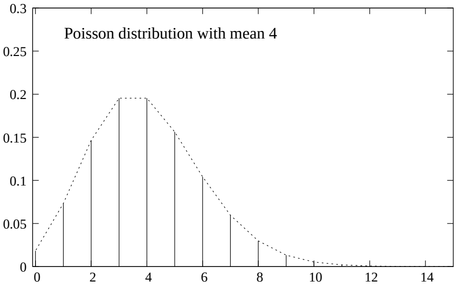
- increments are stationary : homogeneous in time
- $\mathbb{E}(N_b - N_a) = \lambda(b-a)$
- $\lambda =$ intensity or throughput of the process
  number of events per unit of time

Poisson distribution $\mathcal{P}(\lambda)$

X random variable Poisson distributed with parameter λ

$$\mathbb{P}(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}.$$

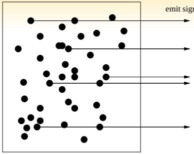
$$\mathbb{E}X = \lambda; \quad VarX = \lambda.$$



If X and Y are independent random variable Poisson distributed with mean λ and μ then

$$X + Y \sim \mathcal{P}(\lambda + \mu).$$

Interpretation



N elements, each of them $p =$ probability of signal emission
X total number of emissions: binomial distribution $\mathcal{B}(N, p)$.
$\mathbb{E}X = Np \stackrel{def}{=} \lambda$ mean number of emissions.

$$\begin{aligned}\mathbb{P}(X = k) &= \binom{N}{k} p^k (1-p)^{N-k};\\ &= \underbrace{\frac{N(N-1)\cdots(N-k+1)}{N.N\cdots N}}_{\to 1} \underbrace{\frac{1}{(1-\frac{\lambda}{N})^k}}_{\to 1} \frac{\lambda^k}{k!} \underbrace{(1-\frac{\lambda}{N})^N}_{\to e^{-\lambda}};\\ &\simeq e^{-\lambda}\frac{\lambda^k}{k!}.\end{aligned}$$

for very large N, X is asymptotically Poisson distributed

Traffic analysis

Traffic generated by a huge amount of users $\Rightarrow$ Poisson process

- requests arrival on a web server
- arrivals of phone calls
- routed packets in a network
- carrs on a road network
- ...

How to detect non-Poisson traffic

- Time dependence or correlation (burstyness, periodicity,. . .)
- Mean $<$ Variance : too much variability
- smoothers of the traffic (peack avoidance strategies)
- ...

Microscopic modelling

Definition

Microscopic definition A continuous time stochastic process $\{N_t\}_{t\in\mathbb{R}^+}$ *is a counting Poisson process with intensity* λ *iff i)* $N_0 = 0$
ii) $\{N_t\}_{t\in\mathbb{R}^+}$ *have independent and stationary increments*
iii) On a very small intervall $]t, t+dt]$ *we have :*

$$\begin{aligned}
\mathbb{P}(N_{t+dt} - N_t = 1) &= \lambda dt + o(dt) \\
\mathbb{P}(N_{t+dt} - N_t = 0) &= 1 - \lambda dt + o(dt) \\
\mathbb{P}(N_{t+dt} - N_t \geqslant 2) &= o(dt)
\end{aligned}$$

Properties

- increments are stationary : homogeneous in time
- $\mathbb{E}(N_b - N_a) = \lambda(b - a)$
- $\lambda =$ intensity or throughput of the process
  number of events per unit of time

Differential system

$$p_n(t) = \mathbb{P}(N_t = n)$$

$$\begin{aligned}
p_n(t+dt) &= \mathbb{P}(N_{t+dt} = n) \\
&= \mathbb{P}(N_{t+dt} = n | N_t = n)\mathbb{P}(N_t = n) \text{ nothing happens} \\
&\quad + \mathbb{P}(N_{t+dt} = n | N_t = n-1)\mathbb{P}(N_t = n-1) \text{ one arrival} \\
&\quad + \mathbb{P}(N_{t+dt} = n | N_t < n-1)\mathbb{P}(N_t < n-1) \text{ more than one arrival} \\
&\quad \text{ independent increments} \\
&= \mathbb{P}(N_{t+dt} - N_t = 0) p_n(t) \text{ nothing happens} \\
&\quad + \mathbb{P}(N_{t+dt} - N_t = 1) p_{n-1}(t) \text{ one arrival} \\
&\quad + \mathbb{P}(N_{t+dt} - N_t \geqslant 2)\mathbb{P}(N_t < n-1) \text{ more than one arrival} \\
&= (1 - \lambda dt + o(dt)) p_n(t) + (\lambda dt + o(dt)) p_{n-1}(t) + o(dt) \\
&= p_n(t) + \lambda(p_{n-1}(t) - p_n(t)) dt + o(dt)
\end{aligned}$$

recurrent differential equations

$$p_n'(t) = \lambda(p_{n-1}(t) - p_n(t)), \quad p_0(t) = \lambda p_0(t)$$

which is solved by recurrence (put $q_n(t) = e^{\lambda t} p_n(t)$)

$$p_n(t) = e^{-\lambda t}\frac{(\lambda t)^n}{n!}$$

Differential system

$$p_n(t) = \mathbb{P}(N_t = n)$$

$$\begin{aligned}
p_n(t+dt) &= \mathbb{P}(N_{t+dt} = n) \\
&= \mathbb{P}(N_{t+dt} = n | N_t = n)\mathbb{P}(N_t = n) \text{ nothing happens} \\
&\quad + \mathbb{P}(N_{t+dt} = n | N_t = n-1)\mathbb{P}(N_t = n-1) \text{ one arrival} \\
&\quad + \mathbb{P}(N_{t+dt} = n | N_t < n-1)\mathbb{P}(N_t < n-1) \text{ more than one arrival} \\
&\quad \text{ independent increments} \\
&= \mathbb{P}(N_{t+dt} - N_t = 0)p_n(t) \text{ nothing happens} \\
&\quad + \mathbb{P}(N_{t+dt} - N_t = 1)p_{n-1}(t) \text{ one arrival} \\
&\quad + \mathbb{P}(N_{t+dt} - N_t \geqslant 2)\mathbb{P}(N_t < n-1) \text{ more than one arrival} \\
&= (1 - \lambda dt + o(dt))p_n(t) + (\lambda dt + o(dt))p_{n-1}(t) + o(dt) \\
&= p_n(t) + \lambda(p_{n-1}(t) - p_n(t))dt + o(dt)
\end{aligned}$$

recurrent differential equations

$$p_n'(t) = \lambda(p_{n-1}(t) - p_n(t)), \quad p_0(t) = \lambda p_0(t)$$

which is solved by recurrence (put $q_n(t) = e^{\lambda t}p_n(t)$)

$$p_n(t) = e^{-\lambda t}\frac{(\lambda t)^n}{n!}$$

Interarrivals

Let t be a fixed time and let T_t be the time to the next arrival after time t.

$$\mathbb{P}(T_t \geqslant s) = \mathbb{P}(N_{t+s} - N_t = 0) = e^{-\lambda s}.$$

T_t is exponentially distributed with rate λ

The inter-arrival process $\{A_n\}_{n\in\mathbb{N}}$ is a sequence of independent exponentially distributed random variable with rate λ

Exponential distribution

Density, rate λ :

$$f(x) = \lambda e^{-\lambda x}$$

Cumulative distribution function

$$F(x) = 1 - e^{-\lambda x}$$

Mean, Variance

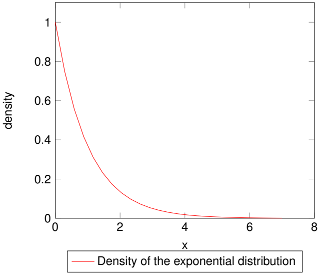
$$\mathbb{E}X = \frac{1}{\lambda}, \quad \mathrm{Var}X = \frac{1}{\lambda^2}$$

Hazard rate

$$h(x) = \lambda$$

Laplace transform

$$\mathcal{L}(t) = \mathbb{E}e^{-tX} = \frac{\lambda}{t+\lambda}$$



Memoryless property

$$\mathbb{P}(X > t + s | X > t) = \mathbb{P}(X > s)$$

Equivalence of definitions

Theorem

Macroscopic, ***microscopic*** *and* ***independent exponentially distributed inter-arrivals*** *are equivalent definitions of a Poisson process*

Proof : classical books

Spread of points

Let $[a, b]$ an interval, knowing $N_b - N_a = n$ the n points are distributed as the rearrangement of n points independents and uniformly distributed points on $[a, b]$

The **Poisson process** is the model of process with a fixed intensity and **minimal *a priori* information**

Equivalence of definitions

Theorem

Macroscopic, *microscopic* and *independent exponentially distributed inter-arrivals* are equivalent definitions of a Poisson process

Proof : classical books

Spread of points

Let $[a, b]$ an interval, knowing $N_b - N_a = n$ the n points are distributed as the rearrangement of n points independents and uniformly distributed points on $[a, b]$

The **Poisson process** is the model of process with a fixed intensity and **minimal *a priori* information**

Equivalence of definitions

Theorem

Macroscopic, ***microscopic*** *and* ***independent exponentially distributed inter-arrivals*** *are equivalent definitions of a Poisson process*

Proof : classical books

Spread of points

Let $[a, b]$ an interval, knowing $N_b - N_a = n$ the n points are distributed as the rearrangement of n points independents and uniformly distributed points on $[a, b]$

The **Poisson process** is the model of process with a fixed intensity and **minimal *a priori* information**

Outline

Continuous time Markov chain

Definition 1: Poisson System

Consider a discrete time Markov chain $\{X_n\}_{n\in\mathbb{N}}$ and $\{N_t\}_{t\in\mathbb{R}}$ a Poisson process $\mathcal{PP}(\lambda)$
At each "tick" of the Poisson process change the state according the Markov chain.

$$X_t = X_{N_t}.$$

Definition 2: Microscopic evolution (rates)

$$\begin{aligned}
\mathbb{P}(X_{t+dt} = j | X_t = i) &= q_{i,j}dt + o(dt) \text{ for } j \neq i \\
\mathbb{P}(X_{t+dt} = i | X_t = i) &= 1 + q_{i,i}dt + o(dt)
\end{aligned}$$

$Q = ((q_{i,j}))$ matrix of rates

Continuous time Markov chain (2)

Definition 3: Sojourn and jump

Consider a discrete time Markov chain $\{\tilde{X}_n\}_{n \in \mathbb{N}}$ and sojourn times for each state iid $\mathcal{E}(-q_{i,i})$ for state i
After each jump the Markov chain stays in the state for an exponentially distributed time.

Formal definition

Let $\{X_t\}_{t\in\mathbb{R}}$ a stochastic process in a discrete state-space $\mathcal{X}$

$\{X_t\}_{t\in\mathbb{R}}$ is a **Markov chain** with initial law $\pi(0)$ iff

- $X_0 \sim \pi(0)$ and
- for all $n \in \mathbb{N}$ for all $s, t > t_{n-1} > \cdots > t_0$ and for all $(j, i, i_{n-1}, \cdots, i_0) \in \mathcal{X}^{n+2}$

$$\mathbb{P}(X_{t+s} = j | X_t = i, X_{t_{n-1}} = i_{n-1}, \cdots, X_{t_0} = i_0) = \mathbb{P}(X_{t+s} = j | X_t = i).$$

$\{X_t\}_{t\in\mathbb{N}}$ is a **homogeneous** Markov chain iff

- for all $t \in \mathbb{R}$ and for all $(j, i) \in \mathcal{X}^2$

$$\mathbb{P}(X_{t+s} = j | X_t = i) = \mathbb{P}(X_s = j | X_0 = i) \stackrel{def}{=} p_{i,j}(s).$$

(invariance during time of probability transition)

Formal definition

Let $\{X_t\}_{t\in\mathbb{R}}$ a stochastic process in a discrete state-space $\mathcal{X}$

$\{X_t\}_{t\in\mathbb{R}}$ is a **Markov chain** with initial law $\pi(0)$ iff

- $X_0 \sim \pi(0)$ and
- for all $n \in \mathbb{N}$ for all $s, t > t_{n-1} > \cdots > t_0$ and for all $(j, i, i_{n-1}, \cdots, i_0) \in \mathcal{X}^{n+2}$

$$\mathbb{P}(X_{t+s} = j | X_t = i, X_{t_{n-1}} = i_{n-1}, \cdots, X_{t_0} = i_0) = \mathbb{P}(X_{t+s} = j | X_t = i).$$

$\{X_t\}_{t\in\mathbb{N}}$ is a **homogeneous** Markov chain iff

- for all $t \in \mathbb{R}$ and for all $(j, i) \in \mathcal{X}^2$

$$\mathbb{P}(X_{t+s} = j | X_t = i) = \mathbb{P}(X_s = j | X_0 = i) \stackrel{def}{=} p_{i,j}(s).$$

(invariance during time of probability transition)

Algebraic representation

$P(s)$ is the **transition matrix** of the chain on an interval of length s

- P is a **stochastic matrix**
  Linear recurrence equation, let $\pi_i(n) = \mathbb{P}(X_n = i)$ then $\pi_t = \pi_0 P(t)$.
- Equation of **Chapman-Kolmogorov** (homogeneous):

  $$P(s+t) = P(s).P(t)$$

  (semi-group structure)

  $$\begin{aligned}\mathbb{P}(X_{s+t} = j | X_0 = i) &= \sum_k \mathbb{P}(X_{s+t} = j | X_t = k)\mathbb{P}(X_t = k | X_0 = i);\\ &= \sum_k \mathbb{P}(X_s = j | X_0 = k)\mathbb{P}(X_t = k | X_0 = i). \text{ homogeneity}\end{aligned}$$

  Interpretation: decomposition of the set of paths with length $n+m$ from i to j.
- Differential equation

  $$\frac{d}{dt}P(t) = P'(0)P(t) = P(t)P'(0)$$
- $Q = P'(0)$ is the **generator** of the chain and $P(t) = exp(tQ)$
  for $i \neq j$, $q_{i,j} \geqslant 0$ and $\sum_j q_{i,j} = 0$ (sum on rows)

Problems

Finite horizon

- Estimation of $\pi(t)$
- Estimation of stopping times

$$\tau_A = \inf\{t \geqslant 0;\ X_n \in A\}$$

- $\dots$

Infinite horizon

- Convergence properties
- Estimation of the asymptotics
- Estimation speed of convergence
- $\dots$

Problems

Finite horizon

- Estimation of $\pi(t)$
- Estimation of stopping times

$$\tau_A = \inf\{t \geqslant 0;\ X_n \in A\}$$

- ...

Infinite horizon

- Convergence properties
- Estimation of the asymptotics
- Estimation speed of convergence
- ...

Convergence In Law

Let $\{X_t\}_{t\in\mathbb{N}}$ a homogeneous, irreducible and ~~aperiodic~~ Markov chain taking values in a discrete state $\mathcal{X}$ then

- The following limits exist (and do not depend on i)

$$\lim_{n\to+\infty} \mathbb{P}(X_t = j | X_0 = i) = \pi_j;$$

- π is the unique probability vector invariant by Q

$$\pi Q = 0;$$

- The convergence is rapid (geometric); there is $C > 0$ and $0 < \alpha$ such that

$$||\mathbb{P}(X_t = j | X_0 = i) - \pi_j|| \leqslant C.e^{-\alpha t}.$$

Denote

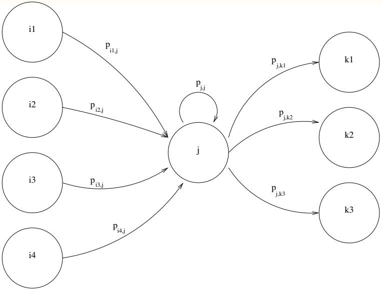
$$X_t \xrightarrow{\mathcal{L}} X_\infty;$$

with X_∞ with law π

π is the **steady-state probability** associated to the chain

Interpretation

Equilibrium equation



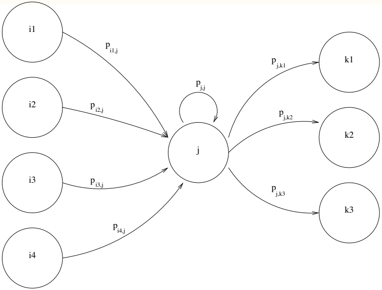
Probability to enter j =probability to exit j
balance equation

$$\sum_{i \neq j} \pi_i q_{i,j} = \sum_{k \neq j} \pi_j q_{j,k} = \pi_j \sum_{k \neq j} p_{j,k} = \pi_j - q_{j,j}$$

$\pi \stackrel{def}{=}$ steady-state.
If $\pi_0 = \pi$ the process is stationary ($\pi_t = \pi$)

Interpretation

Equilibrium equation



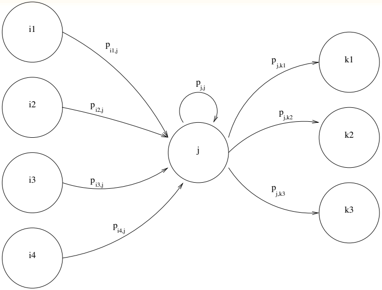
Probability to enter j =probability to exit j
balance equation

$$\sum_{i \neq j} \pi_i q_{i,j} = \sum_{k \neq j} \pi_j q_{j,k} = \pi_j \sum_{k \neq j} p_{j,k} = \pi_j - q_{j,j}$$

$\pi \stackrel{def}{=}$ steady-state.
If $\pi_0 = \pi$ the process is stationary ($\pi_t = \pi$)

Interpretation

Equilibrium equation



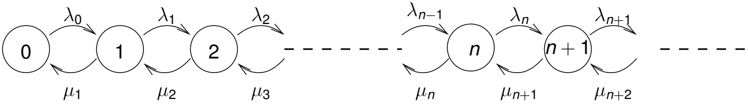
Probability to enter j =probability to exit j
balance equation

$$\sum_{i \neq j} \pi_i q_{i,j} = \sum_{k \neq j} \pi_j q_{j,k} = \pi_j \sum_{k \neq j} p_{j,k} = \pi_j - q_{j,j}$$

$\pi \overset{def}{=}$ steady-state.
If $\pi_0 = \pi$ the process is stationary ($\pi_t = \pi$)

Outline

Birth and Death models



- Stationary distribution

$$\pi_n = \pi_0 . \frac{\lambda_0 \lambda_1 \cdots \lambda_{n-1}}{\mu_1 \mu_2 \cdots \mu_n}$$

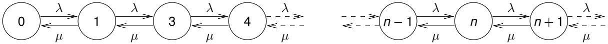
- Stability condition

$$\sum_n \frac{\lambda_0 \lambda_1 \cdots \lambda_{n-1}}{\mu_1 \mu_2 \cdots \mu_n} < \infty$$

- Model of many systems
  queueing systems, reliability models, epidemic models, systems of particles, ...

The classic M/M/1 queue

$\{X_t\}_{t\in\mathbb{R}}$, number of clients in the queue at time t is a Birth-and-Death process

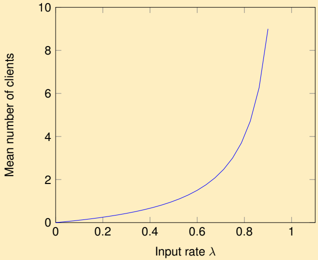


Formula

- **Load** of the queue $\rho = \frac{\lambda}{\mu}$ (utilization of the server)
- **Stability** of the queue $\rho < 1$
- **Steady-state** of the queue $\pi_n = (1-\rho)\rho^n$ (geometric distribution)
- **Mean number of clients** in the queue $\overline{X} = \frac{\rho}{1-\rho}$
- **Mean response time** of clients $\overline{W} = \frac{1}{\mu-\lambda}$
- **Overflow probability** $\mathbb{P}(X \geqslant n) = \rho^n$

Performances

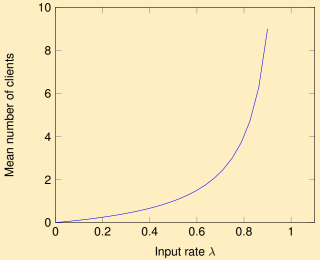
Mean number of clients $\mu = 1$



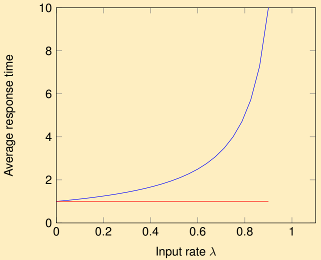
Average response time $\mu = 1$

Performances

Mean number of clients $\mu = 1$



Average response time $\mu = 1$



Outline

Synthesis

Markov chains

- plays the role of linear models for deterministic dynamical systems
  $\Rightarrow$ First order approximation
- algebraic methods to solve the model
  $\Rightarrow$ Numerically tractable
- systems with "steady-state" behavior
  $\Rightarrow$ Not highly variable

Markovian queues

- brick for modeling distributed/parallel resources
  $\Rightarrow$ basics for networking, operating systems
- robustness of the single queue model
  $\Rightarrow$ variation of input or service process
  $\Rightarrow$ composition of queues (networks)
- analytical form
  $\Rightarrow$ computable product form, asymptotic independence

Synthesis

Markov chains

- plays the role of linear models for deterministic dynamical systems
  $\Rightarrow$ First order approximation
- algebraic methods to solve the model
  $\Rightarrow$ Numerically tractable
- systems with "steady-state" behavior
  $\Rightarrow$ Not highly variable

Markovian queues

- brick for modeling distributed/parallel resources
  $\Rightarrow$ basics for networking, operating systems
- robustness of the single queue model
  $\Rightarrow$ variation of input or service process
  $\Rightarrow$ composition of queues (networks)
- analytical form
  $\Rightarrow$ computable product form, asymptotic independence