The mathematics of file dissemination

Philippe Nain (INRIA)

YEQT-VI, Eindhoven, Nov. 1-3, 2012

INSTITUT NATIONAL
DE RECHERCHE
EN INFORMATIQUE
ET EN AUTOMATIQUE

INRIA

centre de recherche
SOPHIA ANTIPOLIS - MÉDITERRANÉE

Introduction

General problem: time to transfer content (file) between one user (or peer) and another one

- Users (peers) may be stationary or mobile

- Two different paradigms:

  1. Client-server model

  2. Peer-to-peer model (full cooperation, limited cooperation)

Client-server model with stationary users

- N+1 users including 1 source (publisher)
- All users present at t=0
- **T** = file transmission time between source and user (**deterministic**)
- D_{cs}= dissemination time (time for all users to get file)

$$\mathbf{D_{cs}} = N\,\mathbf{T}$$

P2P model with stationary users : full cooperation

- N+1 users (peers) including 1publisher, all present at t=0
- Upon receiving file peers become **permanent** publishers
- **T** = file transmission time (**deterministic**)
- D_{p2p} = dissemination time

$$
\begin{aligned}
\mathbf{D}_{\mathbf{p2p}} &= \mathbf{T} \quad \text{if } N = 1 \\
&= 2\mathbf{T} \text{ if } N = 2,3 \\
&= \ \ldots \\
&= K\mathbf{T} \text{ if } 2^{K-1} \le N < 2^K
\end{aligned}
$$

$$\mathbf{D}_{\mathbf{p2p}} = \lceil \log_2 N \rceil \, \mathbf{T}$$

Numerical examples (T deterministic)

- N = 1000

$\mathbf{D_{cs}} = 1000\ \mathbf{T}$; $\mathbf{D_{p2p}} = 9.965\ \mathbf{T}$

- N = 10000

$\mathbf{D_{cs}} = 10000\ \mathbf{T}$; $\mathbf{D_{p2p}} = 13.287\ \mathbf{T}$

$\mathbf{T}$ = **exponentially** distributed (mean $1/\tau$)

$$E[\mathbf{D}_{cs}] = N / \tau$$

$$E[\mathbf{D}_{p2p}] = ?$$

P2P model with stationary users : full cooperation, exponential transmission time

P2P model with stationary users : full cooperation, exponential transmission time (cont.)

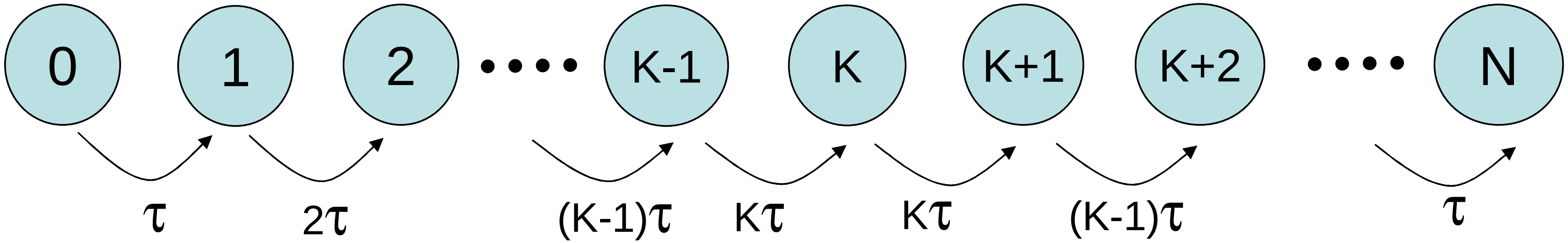
X(t) = # peers with file at time t excluding initial publisher

$\{\mathbf{X(t)}\}_t$: Markov (birth) process

P2P model with stationary users : full cooperation, exponential transmission time (cont.)

X(t) = i if i peers have file at t excluding initial publisher

- N **even** (K = N/2)

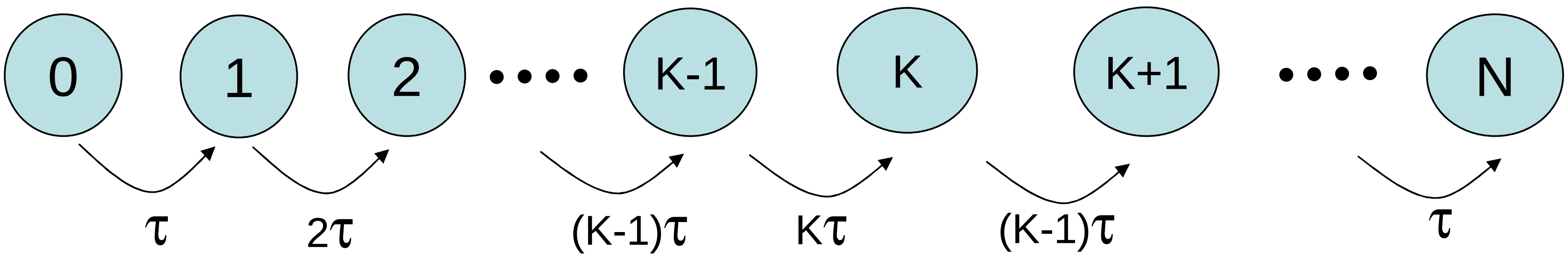


$$E[\mathbf{D}_{\mathbf{p2p}}] = \frac{2}{\tau}\sum_{i=1}^{K}\frac{1}{i} = \frac{2}{\tau}(\ln(\frac{N}{2}) + \gamma + O(\frac{1}{N}))$$

γ = 0.57721… Euler's constant

P2P model with stationary users : full cooperation, exponential transmission time (cont.)

- N **odd** (K = (N+1)/2)



$$E[\mathbf{D_{p2p}}] = \frac{2}{\tau}\sum_{i=1}^{K-1}\frac{1}{i} + \frac{1}{\tau K}$$

$$= \frac{2}{\tau}\left(\ln\left(\frac{N+1}{2}\right) - \frac{1}{N+1} + \gamma + O\left(\frac{1}{N+1}\right)\right)$$

Numerical examples ($T \sim exp(\tau)$)

$E[T] = \tau^{-1}$

- N = 1000

$E[\mathbf{D}_{cs}] = 1000\tau^{-1}$

$E[\mathbf{D}_{p2p}] = 13.586\tau^{-1}$

$E[\mathbf{D}_{cs}]$ linear in N

$E[\mathbf{D}_{p2p}]$ ~ logarithmic in N

- N = 10000

$E[\mathbf{D}_{cs}] = 10000\tau^{-1}$

$E[\mathbf{D}_{p2p}] = 18.189\tau^{-1}$

Conclusion

Huge benefit in sharing content!

Note

Analytic untractability quickly around the corner …

Examples (**limited** cooperation):

- after exponential duration peer with file leaves network
- after k uploads a peer leaves

P2P: mobile and cooperative users
(Epidemic-like dissemination)

- N+1 **mobile** peers including 1 publisher.
- 1 file in system at t=0
- Transmission occurs each time peer with file meets peer without file (**instantaneous** transmissions)
- Successive contact times between any pair of peers form **Poisson** process (rate λ) [Groenevelt, P.N., Koole, PEVA 2005]

X(t) = # files at time t (**X(0-)**=1)

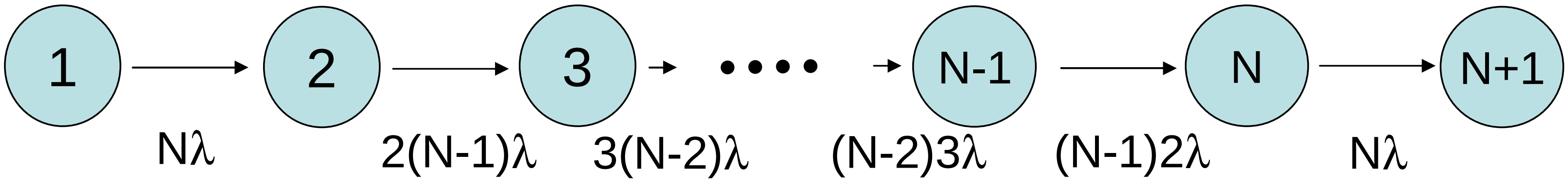
$\{\mathbf{X(t)}\}_t$ Markov

rate from state i to state i+1 : $i(N+1-i)\lambda$

P2P: mobile and cooperative users

(Epidemic-like dissemination) (cont.)

D_e = time before full dissemination



$$E[D_e] = \frac{1}{\lambda}\sum_{i=1}^{N}\frac{1}{i(N+1-i)} = \frac{2}{\lambda(N+1)}\sum_{i=1}^{N}\frac{1}{i}$$

$$= \frac{2}{\lambda(N+1)}(\ln(N)+\gamma+O(\frac{1}{N}))$$

P2P: mobile and cooperative users
(Epidemic-like dissemination) (cont.)

- Transient behavior of X(t) for N large

$dE[\mathbf{X(t)}]/dt = \lambda\, E[\,\mathbf{X(t)}(N+1-\mathbf{X(t)})]$ **(Kolmorogov eq.)**

$$E[\,\mathbf{X(t)}(N+1-\mathbf{X(t)}\,)] \leq E[\mathbf{X(t)}]\,(N+1-E[\mathbf{X(t)}])$$

so that $$dE[\mathbf{X(t)}]/dt \leq \lambda\, E[\mathbf{X(t)}]\,(N+1-E[\mathbf{X(t)}])$$

Subsequently, $E[\mathbf{X(t)}] \leq z(t) = \dfrac{N+1}{1+Ne^{-\lambda(N+1)t}}$

with z solution of $dz/dt = \lambda z(N+1-z), \quad z(0)=1$

P2P: mobile and cooperative users
(Epidemic-like dissemination) (cont.)

Assume $\lambda = \beta/(N+1)$

Take $y(t) = z(t)/(N+1)$

$$dz/dt = \lambda z(N+1-z)$$

yields $$dy/dt = \beta y(1-y)$$

so that $$y(t) = y(0)/(y(0)+(1-y(0))e^{-\beta t})$$

[Kurtz, 70] If $\lim_{N\to\infty}(N+1)\mathbf{X(0)} = y(0)>0$ then

$$\lim_{N\uparrow\infty} P\left(\sup_{s\le t} \left|\frac{X(s)}{N+1} - y(s)\right| > \varepsilon \right) = 0 \quad \forall \varepsilon > 0$$

P2P: mobile and cooperative users
(Epidemic-like dissemination) (cont.)

In summary:

$$E[\mathbf{X(t)}] \sim (N+1)/(1+Ne^{-\beta t})$$

as number peers (N) large and $\lambda = \beta/(N+1)$

« Proof » : take $y(0)=1/(N+1)$ in [Kurtz]

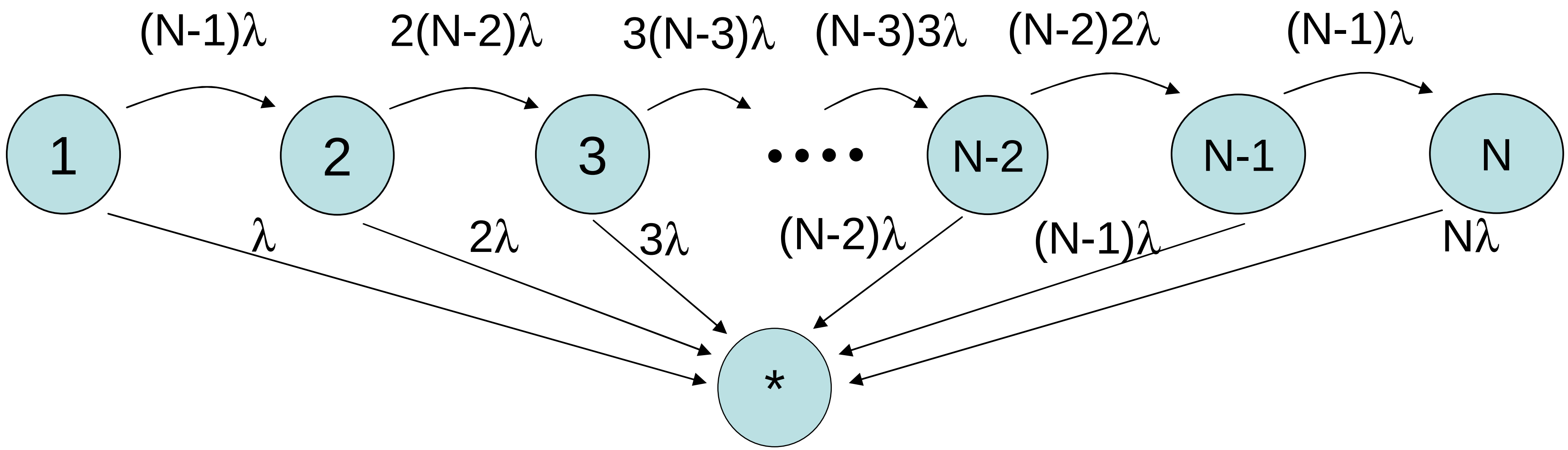
P2P: mobile and cooperative users
(Epidemic-like dissemination) (cont.)

Peer's point of view: Expected time for peer to get file?

Let $E[\mathbf{T_d}]$ be this time

P2P: mobile and cooperative users

(Epidemic-like dissemination) (cont.)



$$E[\mathbf{T}_d] = \frac{1}{\lambda N}\sum_{i=1}^{N}\frac{1}{i} = \frac{1}{\lambda N}\left(\ln(N)+\gamma+O\left(\frac{1}{N}\right)\right)$$

P2P: mobile and cooperative users
(Epidemic-like dissemination) (cont.)

Mean-field approach:

$dP_N(\mathbf{T_d}>t)/dt = \lambda\,(1- P_N(\mathbf{T_d}>t))\,E_N[\mathbf{X(t)}]$ (Kolmogorov)

$= \beta\,(1- P_N(\mathbf{T_d}>t))\,E_N[\mathbf{X(t)/}(N+1)]$

$N\to\infty$: $E_N[\mathbf{X(t)/}(N+1)] \to y(t)$ and $P_N(\mathbf{T_d}>t) \to P(t)$ given by

$$dP/dt = \beta(1-P)y$$

with $y(t) = 1/(1+Ne^{-\beta t})$ (mean-field for $\{\mathbf{X(t)}/(N+1)\}$)

$$P(t) = 1 - (N+1)/(N+e^{\beta t})$$

P2P: mobile and cooperative users

(Epidemic-like dissemination) (cont.)

$P(t) = 1 - (N+1)/(N+e^{\beta t})$

$\lambda = \beta/(N+1)$

$$E[T_d] \sim \int_0^\infty (1-P(t))dt = \frac{N+1}{\beta N}\ln(N+1) = \frac{1}{\lambda N}\ln(N+1)$$

as N large, matching exact result

P2P: mobile and cooperative users
(Epidemic-like dissemination) (cont.)

- Expected total file dissemination

$$E[D_e] = \frac{2}{\lambda(N+1)} \sum_{i=1}^{N} \frac{1}{i}$$

- Expected peer file acquisition

$$E[T_d] = \frac{1}{\lambda N} \sum_{i=1}^{N} \frac{1}{i}$$

$E[D_e]/E[T_d] = 2N/(N+1)$

P2P: mobile and partially cooperative users
(Epidemic-like dissemination) (cont.)

- After exponential time (rate μ) peer with file leaves network

$X(t)$ = # peers with files at time t

$Y(t)$ = # peers without file at time t

$(X(t),Y(t)) \rightarrow (X(t)+1,Y(t)-1)$ rate $\lambda X(t)Y(t)$

$\rightarrow (X(t)-1,Y(t))$ rate $\mu X(t)$

P2P: mobile and partially cooperative users
(Epidemic-like dissemination) (cont.)

A - Mean-field approach

B - Branching process

with E. Altman, A. Shwartz, Y. Xu (IEEE/ACM ToN 2012)

P2P: mobile and partially cooperative users
(Epidemic-like dissemination) (cont.)

A - Mean-field approach

B - Branching process

P2P: mobile and partially cooperative users
(Epidemic-like dissemination) (cont.)

Mean-field approach (N peers including 1 initial publisher)

- $\lambda = N^{-1}\beta$ $(X(t) := X_N(t),\ Y(t) := Y_N(t))$

[Kurtz] If $\lim_N N^{-1}X_N(0) = x_0 > 0$, $\lim_N N^{-1}Y_N(0) = y_0 > 0$, $x(0)+y(0)=1$ then

$$(N^{-1}X_N(t), N^{-1}Y_N(t)) \rightarrow_{prob.} (x(t), y(t)) \quad (N \rightarrow \infty)$$

with

$$dx/dt = x(\beta y - \mu), \qquad x(0) = x_0$$

$$dy/dt = -\beta y x, \qquad y(0) = y_0$$

Result holds for $t = \infty$ for that model (does not hold in general)

P2P: mobile and partially cooperative users
(Epidemic-like dissemination) (cont.)

Interpretation: x (resp.) fraction peers with (without) file as N large

$$dx/dt = x(\beta y-\mu)$$

$$dy/dt = - \beta yx$$

Kermack-McKendrick eqns if one adds

$$dz/dt = \mu x$$

z = fraction peers which have left (recovered) by time t

Note here x(t)+y(t)+z(t)=1 for all t

P2P: mobile and partially cooperative users
(Epidemic-like dissemination) (cont.)

$\theta = \mu/\beta$

$$dx/dt = \beta x(y-\theta)$$
$$dy/dt = -\beta yx$$

$dx/dt = dx/dy \, . \, dy/dt = - dx/dy \, . \, \beta yx = \beta x(y-\theta)$

Hence $dx/dy = -1 + \theta/y$

so that $x(y) = -y + \theta \ln(y) + f(\theta)$

with $f(\theta) := x_0 + y_0 - \theta \ln(y_0) = 1 - \theta \ln(y_0)$

P2P: mobile and partially cooperative users

(Epidemic-like dissemination) (cont.)

$$dx/dt = \beta x(y-\theta)$$
$$dy/dt = -\beta yx \qquad (1)$$

$x(y) = -y + \theta \ln(y) + f(\theta)$ (recall $f(\theta) = 1 - \theta \ln(y_0)$)

x maximum when $y = \theta$ from (1)

Therefore $x_{max} = \theta (\ln(\theta) - 1) + f(\theta)$

P2P: mobile and partially cooperative users

(Epidemic-like dissemination) (cont.)

$$x(y) = - y + 1 + \theta \ln(y) - \theta \ln(y_0) \qquad (2)$$

Ratio of peers without file (set x=0 in (2))

$$0 = - y + 1 + \theta \ln(y) - \theta \ln(y_0)$$

Power series expansion at y_0:

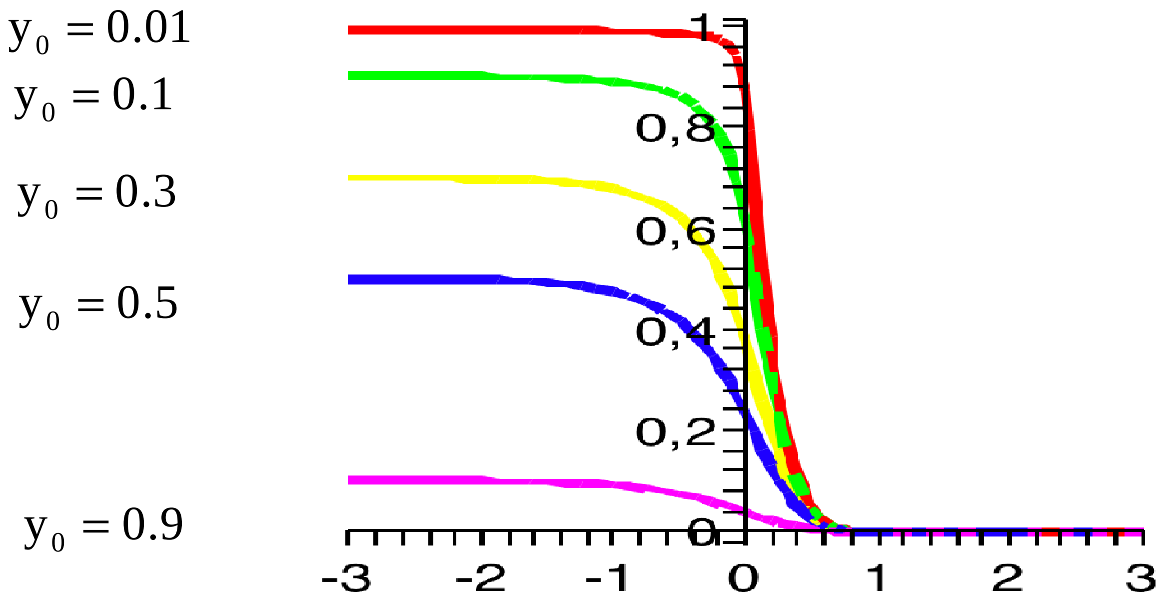
$$0 \sim y (\theta/y_0 - 1) + 1 - \theta - (\theta/2)(y/y_0 - 1)^2 + o((y-y_0)^2)$$

$\theta((y/y_0-1)^2)$ being bounded → Phase transition at

$$\theta = y_0$$

P2P: mobile and partially cooperative users

Ratio of peers that never get file (=x(∞))



$\log_{10}(y_0/\theta)$

$$\theta = \mu/\beta$$

P2P: mobile and partially cooperative users
(Epidemic-like dissemination) (cont.)

A - Mean-field approach

B - Branching process

P2P: mobile and partially cooperative users
(Epidemic-like dissemination) (cont.)

- Branching process approximation

Applies when large fraction of population does not have file at t=0 (Y(0) big)

→ replace Y(t) by $Y(0) := y_0$ for all t

$(X_b(t)) \rightarrow X_b(t)+1$ rate $\lambda X(t) y_0$

$\rightarrow X_b(t)-1$ rate $\mu X(t)$

$\{X_b(t)\}_t$ = Markov branching process

Branching process approximation **(cont.)**

q_k = file extinction probability given $X_b(0)=k$

Can show that $q_k = \min(1, 1/\rho^k)$, $\rho := \lambda y_0/\mu$

→ Phase transition at $\rho = 1$

- Extinction certain if $\rho \leq 1$

Expected time before extinction $(\rho<1)= \dfrac{1}{\mu}\displaystyle\int_0^1 \frac{1-x^k}{\rho x^2-(1+\rho)x+1}dx$

$= -(1/\rho\mu)\ln(1-\rho)$ if $X_b(0)=1$

- $X(t) \leq_{st} X_b(t)$ for all t if $X(0) \leq X_b(0)$

P2P: mobile and partially cooperative users
(Epidemic-like dissemination) (cont.)

Can be used to investigate impact of measures against illegal file sharing

Measures would aim at decreasing contact rate (parameter λ) and cooperation degree (μ)