Urn models and peer-to-peer file sharing[1)]

Technical Research Center of Finland, VTT

Ilkka Norros and Hannu Reittu

1) to appear in *Physcomnet 08*, Berlin 2008

urns

- at time n, W_n white balls and B_n black balls
- draw randomly a ball from the urn
- reinforcement matrix:

$$A = \begin{pmatrix} \alpha & \beta \\ \beta & \alpha \end{pmatrix}$$

- when a color i is drawn color j is reinforced by A_{ij}

some cases

- Polýa urn: $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

- Has a 'random limit': $W_n\,/(n+B_0+W_0)$ is a martingale and convergences to a limiting r.v. with beta distribution $\beta(W_0,B_0)$
- say, $B_0=W_0=1$, uniform distribution in (0,1)
- 'self-organization' (automata, physics), 'lock-in' (economics)

Ehrenfests' urn:

- (equilibrium gas…), overal number of balls is const = N

$$\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}$$

- W_n tends to r.v. with Bin$(N,½)$

BERNARD FRIEDMAN'S URN

- $\beta > 0$

$$A = \begin{pmatrix} \alpha & \beta \\ \beta & \alpha \end{pmatrix}$$

- $W_n/(B_n+W_n)$ converges with pr. 1 to ½
- balancing!
- we rediscovered this: $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

79 pages of further material:

Probability Surveys
Vol. 4 (2007) 1–79
ISSN: 1549-5787
DOI: 10.1214/07-PS094

A survey of random processes with reinforcement*

Robin Pemantle

e-mail: pemantle@math.upenn.edu

Earlier work: file sharing in closed system

- File sharing using overlays(SpaSWiN 2007 and ITC20), 2 papers:

Toward modeling of a single file broadcasting in a closed network

Hannu Reittu and Ilkka Norros

On uncoordinated file distribution with non-altruistic downloaders

Ilkka Norros[1], Balakrishna Prabhu[2], and Hannu Reittu[1]

New topic, started in Amsterdam 2007, with SCALP-partners

- open system, with constant rate (λ) of arriving peers
- two chunks 0 and 1
- upon arriving, peer selects the first chunk to be downloaded, -> (0) and (1) nodes
- two 'algorithms':
- 1) deterministic last chunk ('DLC'): with prob. ½
- 2) use Friedman-urn ('F-a.'): $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

4+1 types of nodes:

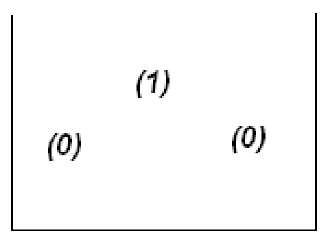
(0) has no chunks, takes chunk 0 first

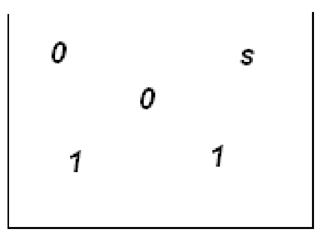
(1) has no chunks, takes chunk 1 first

1 has chunk 1

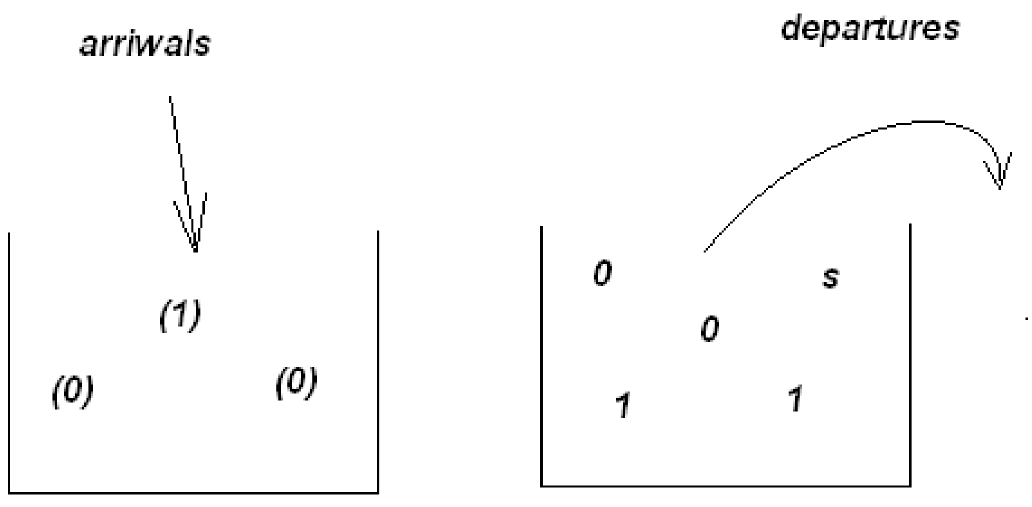
0 has chunk 0

s the seed node (permanent), has both chunks urns:

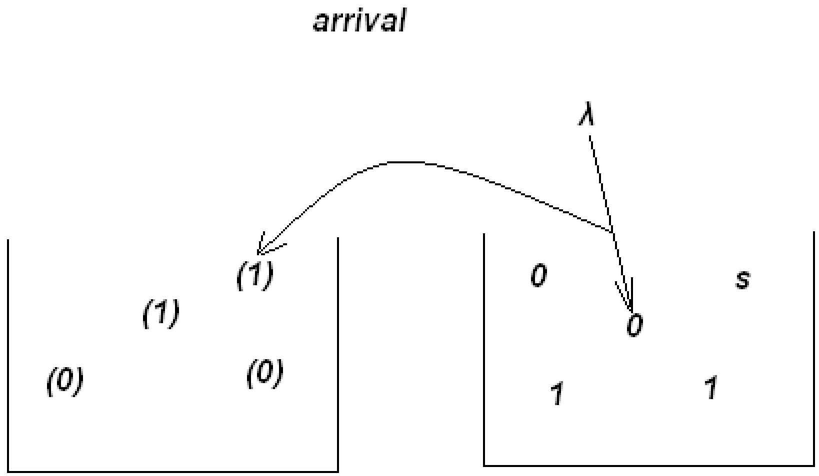




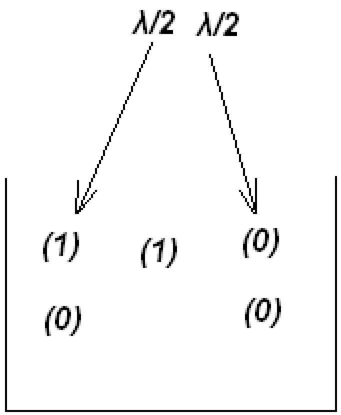
flows

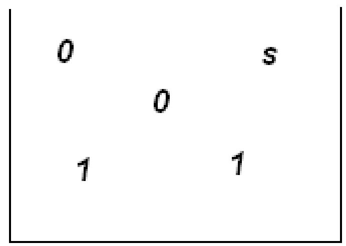


F-algorithm,
arriving peer contacts 0, becomes a (1) node

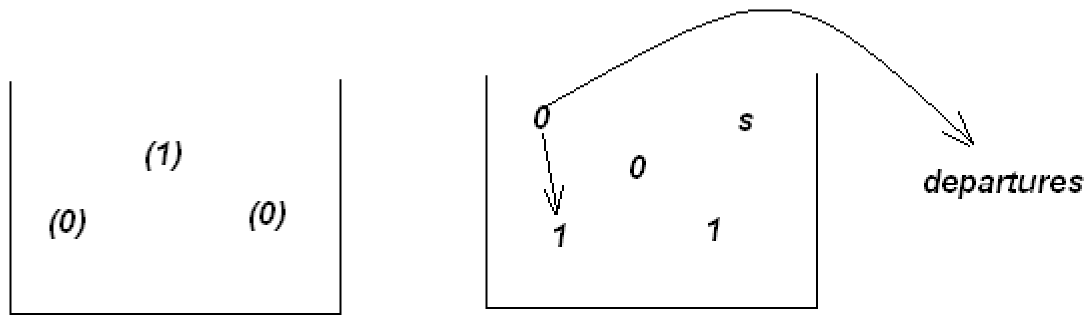


arrivals in DLC

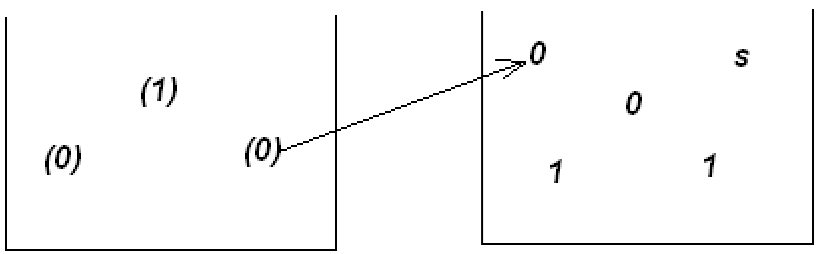




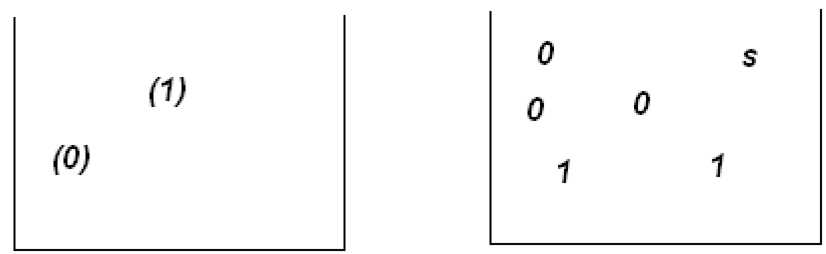
departure: 01 -> 1



A productive contact (0)0 -> 0+0



(0)0 -> 0+0



urn process

- Contact: xy, (uniformly randomly from urn(s))
- $x \in \{(0),(1),0,1\}$
- $y \in \{0,1,s\}$
- x downloader
- y uploader

arrivals in F-a:

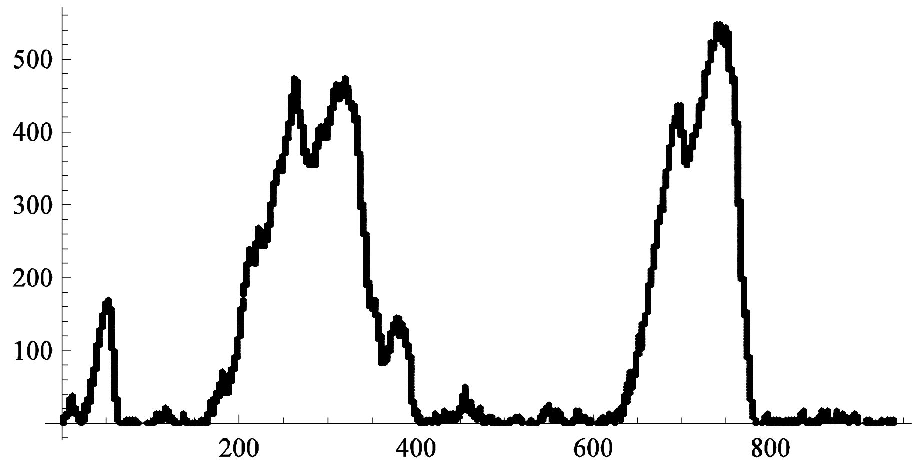
- Ø0 -> (1) + 0
- Ø1 -> (0) + 1

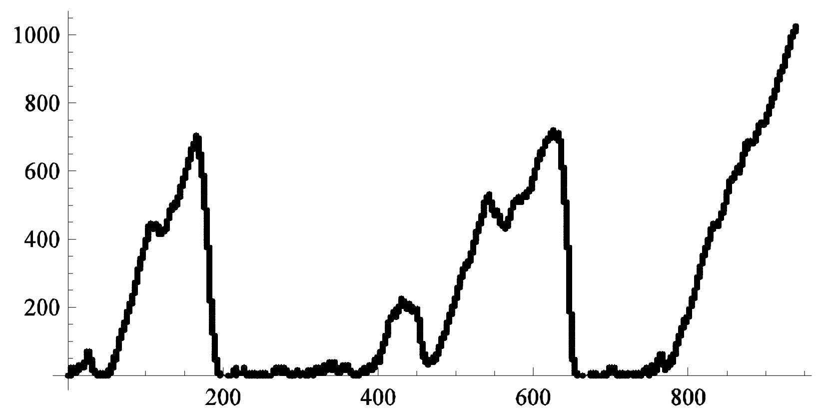
'chemistry' of contacts:

- (0)0 -> 0 + 0
- (0)s -> 0 + s
- (1)1 -> 1 + 1
- 01 -> 1
- 10 -> 0
- rest, say, (0)1-> (0)+1 - nothing happens,
- arrivals:
- Ø0 -> (1) + 0
- Ø1 -> (0) + 1

microscopic view on process, exp-delays between contacts

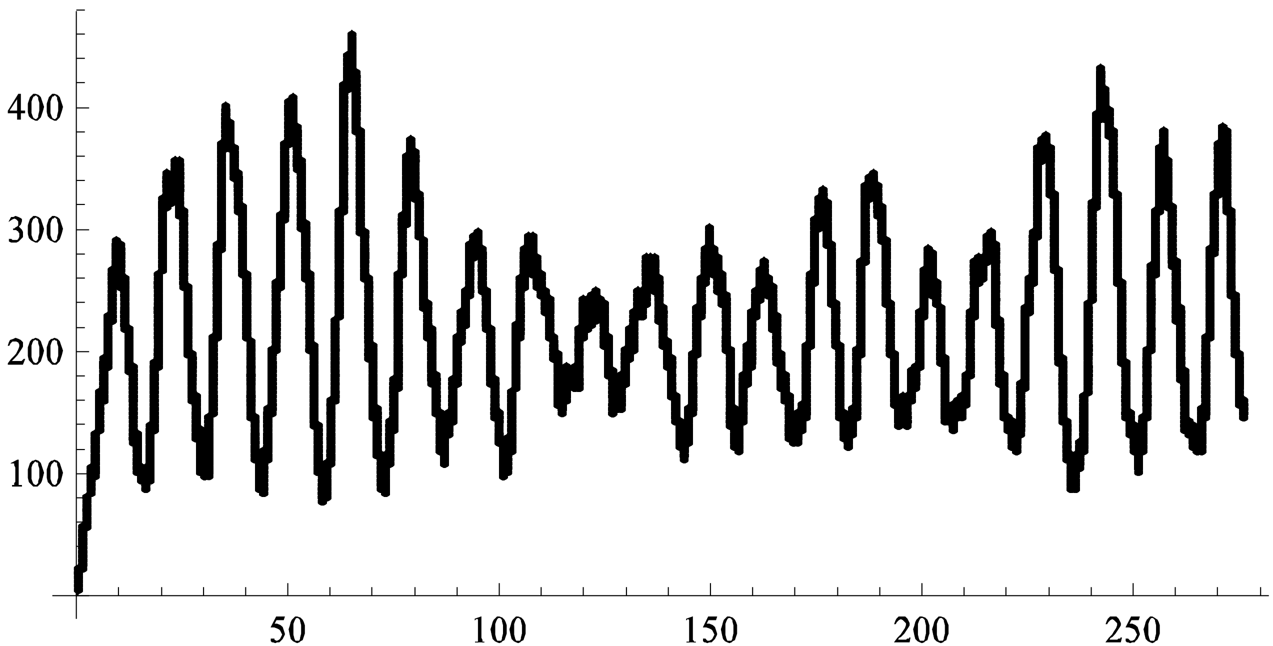
a sample paths of n_1 and n_0, $\lambda=20$ (DLC)



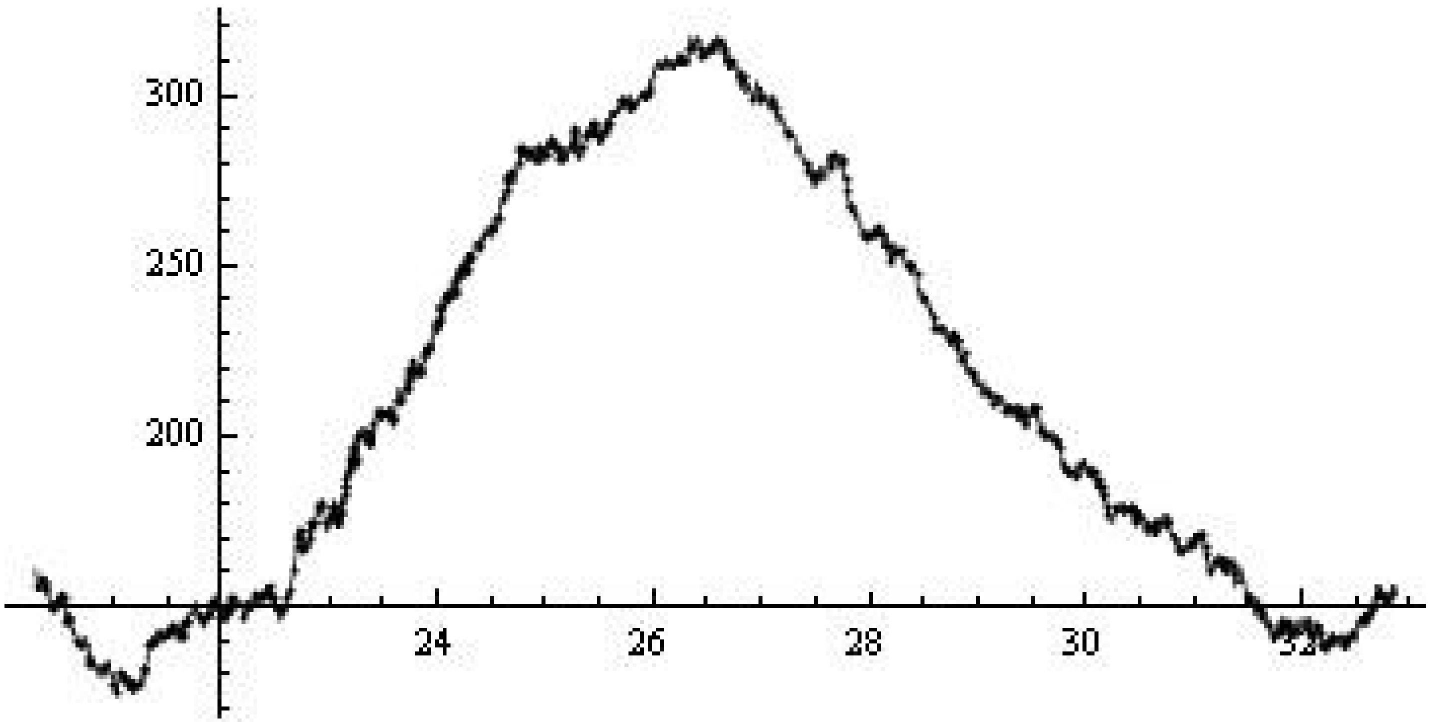


a sample path of n_1, $\lambda=200$

(F-a.) -> oscillations arround equilib. ?



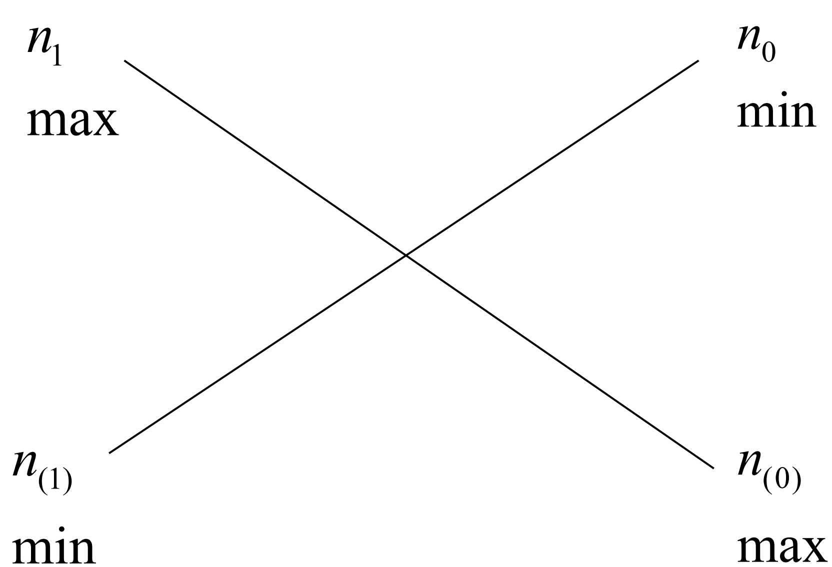
one hump in more details:



all components, similar

period ≈ 15, independent on λ

oscillations in phase, or in antiphase:



diff. equations for DLC
1 corresponds to seed

$$\dot{n}_{(1)} = \frac{\lambda}{2} - \frac{n_{(1)}(n_1 + 1)}{n_1 + n_0 + 1}$$

$$\dot{n}_{(0)} = \frac{\lambda}{2} - \frac{n_{(0)}(n_0 + 1)}{n_1 + n_0 + 1}$$

$$\dot{n}_1 = \frac{n_{(1)}(n_1 + 1)}{n_1 + n_0 + 1} - \frac{n_1(n_0 + 1)}{n_1 + n_0 + 1}$$

$$\dot{n}_0 = \frac{n_{(0)}(n_0 + 1)}{n_1 + n_0 + 1} - \frac{n_0(n_1 + 1)}{n_1 + n_0 + 1}$$

some comments:

- stable focus around $(\lambda,\lambda,\lambda,\lambda)$
- without seed, no stability!
- trajectories have nothing in common with random system!

The same for F-a.

$$\dot{n}_{(1)} = \lambda \frac{(n_0 + 1/2)}{n_1 + n_0 + 1} - \frac{n_{(1)}(n_1 + 1)}{n_1 + n_0 + 1}$$

$$\dot{n}_{(0)} = \lambda \frac{(n_1 + 1/2)}{n_1 + n_0 + 1} - \frac{n_{(0)}(n_0 + 1)}{n_1 + n_0 + 1}$$

$$\dot{n}_1 = \frac{n_{(1)}(n_1 + 1)}{n_1 + n_0 + 1} - \frac{n_1(n_0 + 1)}{n_1 + n_0 + 1}$$

$$\dot{n}_0 = \frac{n_{(0)}(n_0 + 1)}{n_1 + n_0 + 1} - \frac{n_0(n_1 + 1)}{n_1 + n_0 + 1}$$

- has also stable focus, around all=λ
- random system seems to oscillate around approx. same equilibrium

performance? (F-a)

τ ~ average time to make a contact, (emulates chunk download delay)

$N(t)$ - number of nodes in the system at time *t*

$$N = \lim_{t\to\infty} \frac{1}{t} \int_o^\infty N(s)ds$$

we have (?) $N = 4\tau\lambda$

Little's Theorem:

average time for a peer to obtain the file, $T = \frac{N}{\lambda} = 4\tau$

lower bound for *T* is 2τ

F-algorithm, interesting conjectures ???:

- regular oscillations of all components, around equilib.
- seed is not necessary after the start
- oscillations are in phase or in antiphase λ
- fequency (4π?), independent on arrival rate,
- mean delay of a peer = 4τ (only 2 excess contacts!)

Thank You!