Plan

1. Passage à grande échelle
2. Diffusion à grande échelle
   - Algorithmes structurés
   - Algorithmes probabilistes
3. Systèmes pair à pair
   - Principales difficultés
   - Classification
   - Systèmes non structurés
   - Systèmes structurés

Diffusion à grande échelle

Besoins

- Nombre de sites inconnu
- Nombre et identité des sites variables
- Grand nombre de sites

Limites des approches classiques

- Ensemble bien identifié de sites (notion de groupe)
- Propriétés fortes (fiabilité, ordre, atomicité) néfastes au passage à l'échelle

Diffusion par inondation

Cf chapitre IV « problèmes génériques »

Diffuser(m), sur p

```
-- p = émetteur, m = message
∀ s ∈ voisins(p) ∪ {p} faire
    envoyer(⟨p,m⟩) à s
fin pour
```

Réception(⟨p,m⟩), sur q

```
si q n'a pas déjà délivré m alors
    si p ≠ q alors                      -- propagation
        ∀ s ∈ voisins(q) faire
            envoyer(⟨p,m⟩) à s
        fin pour
    fin si
    délivrer(m)
fin si
```

Groupes et diffusion

Notion de groupe

Cf chapitre « Tolérance aux fautes »

Limites

- Vision synchrone des arrivées et départs
- Propriétés fortes (fiabilité, ordre), coûteuses et non indispensables
- Taille d'un groupe limitée

Arbre de recouvrement

- Construire un arbre issu du site de diffusion et contenant tous les sites
- Approximation : graphe orienté acyclique avec détection de messages en doublon
- Difficulté : construire l'arbre. . .
- . . . mais c'est simple quand on a une table de routage hiérarchique, cf transparent 38

Algorithmes épidémiques

Diffusion à grande échelle

- Grand nombre de sites (> 100), voire très grand (> 10000)
- Nombre inconnu et variable de sites
- Rôle symétrique de tous les sites
- Présence de sites en panne
- Topologie d'interconnexion inconnue (a priori non directement maillée)

Approche

- Algorithmes probabilistes
- Propagation aléatoire

1. *Epidemic Algorithms for Replicated Database Maintenance*, Alan Demers et al. 6th Symposium on Principles of Distributed Computing. Aug. 1987.

Algorithme épidémique : Rumeur ♪♪♪

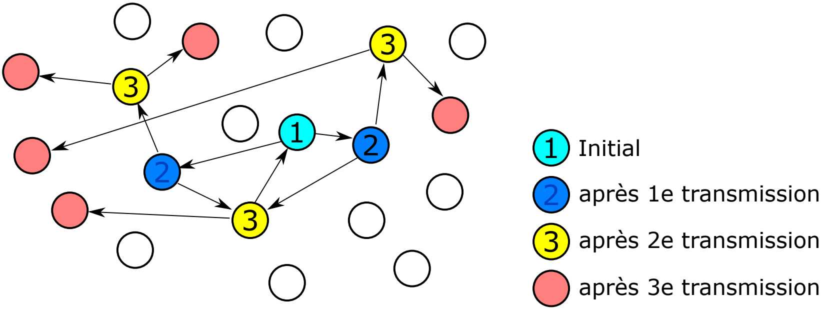
Rumeur (*Gossip*)

Contamination d'autres sites choisis aléatoirement avec une information supposée nouvelle.
États d'un site :

- Infectable : ne possède pas l'information
- Contagieux : apte à contaminer des sites infectables
- Immunisé : a cessé d'être contagieux
  - Immunisation en aveugle ou avec rétroaction (échec d'une tentative de contamination)
  - Compteur (de tentatives ou d'échecs)
  - Probabilité d'abandon après chaque tentative / échec

Rumeur : exemple

20 sites, nombre de sites contactés à chaque tour = 2, immunisation en aveugle à 1 tour

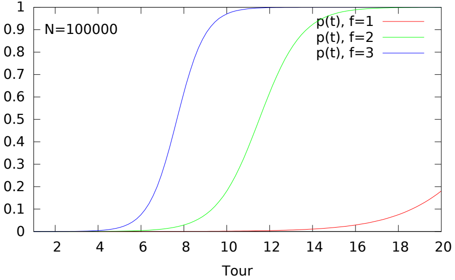


Rumeur : performance ♪♪♪

Modèle simple : *infect forever* = pas d'immunisation

- N = nombre de sites
- f = nombre de sites contactés à chaque tour par chaque site
- $I(t)$ = nombre de sites infectés (contagieux ou immunisés) après le t-ième tour
- $p(t) = I(t)/N$ = proportion de sites infectés au t-ième tour

Alors $p(t) = \frac{1}{1+(N-1)*e^{-f*t}}$



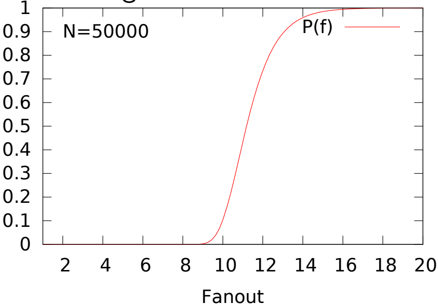
Rumeur : performance ♪♪♪

Modèle simple : *infect and die* = immunisation après un seul tour

- N = nombre de sites
- f = nombre de sites contactés par chaque site (*fanout*)
- P = probabilité que tous les sites finissent par être infectés (à l'∞)

Alors $P \approx e^{-e^{\log N - f}}$

Inversion autour de $f = \log N$

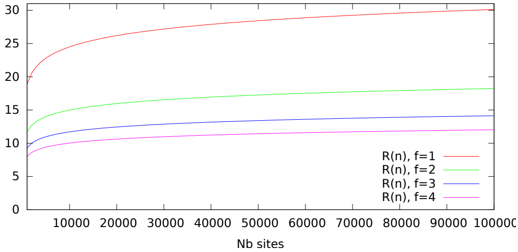


Rumeur : performance

Nombre de tours pour tout contaminer :

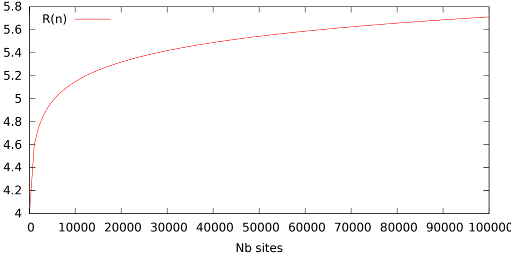
- Modèle sans immunisation (*infect forever*) :

$$R = \log_{f+1} n + \frac{1}{f} \log n + O(1)$$



- Modèle immunisation immédiate (*infect and die*) :
  (fin de l'épidémie avec tous contaminés)

$$R = \frac{\log n}{\log \log n} + O(1)$$



Rumeur : coût

- Diffusion probabiliste
- Existence de sites non informés (poches d'ignorance si le choix des sites est plutôt local)
- Coût individuel très faible (très peu de sites à contacter en comparaison du grand nombre de sites présents)
- Coût global élevé : grand nombre de messages (échecs de contamination de plus en plus probable quand la diffusion progresse)

Algorithme épidémique : Anti-entropie ♪♪♪

Anti-entropie

Périodiquement, chaque site contacte aléatoirement un autre site. Les deux sites échangent alors des informations, et leurs objets sont mis en cohérence.

L'algorithme converge vers l'égalité des copies : cohérence à terme.
Nombre de tours pour tout contaminer $= O(\log N)$

Graphes petit monde ♪♪♪

Les algorithmes précédents n'ont pas de notion de voisinage : choix arbitraire d'un site quelconque (graphe complet).

Graphe petit monde *(small-world graphs)*

Graphe connexe vérifiant :

- Grand nombre de nœuds ($N > 10000$)
- Faible connectivité des nœuds (de l'ordre de 5 à 10)
- Distance entre deux nœuds quelconques $\approx \log N$

- Remarquablement adapté aux algorithmes épidémiques, résistant aux pannes (de sites et de liens)
- Apparaît spontanément (ex : réseau routier : maillage local, rocade, autoroutes)
- ...ou pas (ex : réseau aérien avec quelques gros hubs à forte connectivité)