



**HAL**  
open science

## Réseaux Stochastiques et Algorithmes

Philippe Robert

► **To cite this version:**

Philippe Robert. Réseaux Stochastiques et Algorithmes. Mathématiques [math]. Université Pierre et Marie Curie - Paris VI, 2006. tel-00166813

**HAL Id: tel-00166813**

**<https://theses.hal.science/tel-00166813>**

Submitted on 10 Aug 2007

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# DOCUMENT D'HABILITATION À DIRIGER DES RECHERCHES

## Réseaux Stochastiques et Algorithmes

Université Pierre et Marie Curie

*Spécialité* : Mathématiques Appliquées

*présenté le vendredi 1er décembre 2006 par*

PHILIPPE ROBERT  
INRIA-Rocquencourt

### RAPPORTEURS

**Maury Bramson** *University of Minnesota*  
**Ravi Mazumdar** *University of Waterloo*  
**Isi Mitrani** *University of NewCastle upon Tyne*

### JURY

**Philippe Bougerol** *Université Pierre et Marie Curie*  
**Francis Comets** *Université Denis Diderot*  
**Philippe Flajolet** *INRIA-Rocquencourt*  
**Ravi Mazumdar** *University of Waterloo*  
**Isi Mitrani** *University of NewCastle upon Tyne*



## Table des matières

Publications Présentées pour l'Habilitation	5
Un Bref Survol Historique	9
Chapitre 1. Renormalisation des Réseaux	11
1. Un Peu d'Histoire	11
2. Renormalisation Fluide des Réseaux Stochastiques	13
3. Contributions aux Méthodes de Renormalisation	18
4. Grands Réseaux	22
Chapitre 2. Algorithmes	27
1. Algorithmes en Arbre	27
2. Algorithmes de Contrôle de la Congestion	31
Chapitre 3. Publications	35
1. Preprints	35
2. Livres	35
3. Articles de Revues : Probabilités	35
4. Articles de Revues : Modèles Stochastiques	36
5. Conférences avec comité de lecture	37
6. Contributions au groupe de normalisation IEEE 802.14	38
Références Générales	41
Annexe A. Protocoles d'Accès à un Canal de Communication	45
1. Algorithmique des Réseaux d'Accès	45
2. Le Protocole en Arbre	47
Annexe B. Le protocole de transmission de données TCP	51
1. Introduction	51
2. Algorithmique de TCP	52



## Publications Présentées pour l’Habilitation

1. Nelson Antunes, Christine Fricker, Philippe Robert, and Danielle Tibi. Stochastic networks with multiple stable points. *Annals of Probability*, 2007. To appear.
2. Nelson Antunes, Christine Fricker, Philippe Robert, and Danielle Tibi. Analysis of loss networks with routing. *Annals of Applied Probability*, 16(4) :2007–2026, 2006.
3. Christian Gromoll, Philippe Robert, and Bert Zwart. Fluid limits for processor sharing queues with impatience, April 2006. <http://www.arxiv.org/abs/math.PR/0604215>.
4. Hanène Mohamed and Philippe Robert. A probabilistic analysis of some tree algorithms. *Annals of Applied Probability*, 15(4) :2445–2471, November 2005. <http://www.arxiv.org/abs/math.PR/0412188>.
5. Philippe Robert. On the asymptotic behavior of some algorithms. *Random Structures and Algorithms*, 27(2) :235–250, September 2005. <http://www.arxiv.org/abs/cs.DS/0502014>.
6. Fabrice Guillemin, Philippe Robert, and Bert Zwart. AIMD algorithms and exponential functionals. *Annals of Applied Probability*, 14(1) :90–117, 2004.
7. Philippe Robert. *Stochastic Networks and Queues*, volume 52 of *Stochastic Modelling and Applied Probability Series*. Springer, New-York, June 2003.
8. Vincent Dumas, Fabrice Guillemin, and Philippe Robert. A Markovian analysis of Additive-Increase Multiplicative-Decrease (AIMD) algorithms. *Advances in Applied Probability*, 34(1) :85–111, 2002.
9. Jean-François Dantzer and Philippe Robert. Fluid limits of string valued Markov processes. *Annals of Applied Probability*, 12(3) :860–889, 2002.
10. Jean-François Dantzer, Mostapha Haddani, and Philippe Robert. On the stability of a bandwidth packing algorithm. *Probability in the Engineering and Informational Sciences*, 14(1) :57–79, 2000.
11. Christine Fricker, Philippe Robert, Ellen Saada, and Danielle Tibi. Analysis of some networks with interaction. *Annals of Applied Probability*, 4(4) :1112–1128, 1994.
12. Alain Jean-Marie and Philippe Robert. On the transient behavior of some single server queues. *Queueing Systems, Theory and Applications*, 17 :129–136, 1994.



**Remerciements.** *Je tiens à remercier les rapporteurs Maury Bramson, Ravi Mazumdar et Isi Mitrani d'avoir accepté ce travail et de l'avoir mené de manière diligente. Je remercie Philippe Bougerol, Francis Comets et Philippe Flajolet d'avoir accepté de participer au jury de soutenance, c'est toujours avec plaisir que je passe la porte de leurs bureaux respectifs pour des discussions diverses et variées. Christine Fricker et Danielle Tibi ont été pendant toutes ces années mes co-auteurs les plus régulières, je les en remercie et espère que cela continuera. Fabrice Guillemin de France Telecom & RD et son équipe ont contribué et influencé les recherches que nous menons dans le cadre d'une collaboration régulière, je l'en remercie vivement. Les membres présents et passés de l'équipe de recherche RAP de l'INRIA, Nelson Antunes, Youssef Azzana, Abdelghani Ben-Tahar, Yousra Chabchoub, Vincent Dumas, Stelios Kouremenos, Hanène Mohamed, Florian Simatos et Bert Zwart ont contribué de façon essentielle à la vie de l'équipe de recherche, merci. Je remercie Virginie Collette qui assure le fonctionnement de notre équipe au quotidien avec efficacité et bonne humeur.*

*Jacques Neveu m'a introduit, il y a bien longtemps, dans le monde des probabilités. J'ai appris l'analyse des phénomènes aléatoires (de préférence aussi probabiliste et élégante que possible), mais aussi la nécessaire ouverture à d'autres disciplines scientifiques (comme l'informatique) et aux applications qui portent en elles une partie du renouvellement des thématiques de recherche. Je lui exprime toute ma gratitude.*



## Un Bref Survol Historique

Ce document présente de façon succincte un ensemble de travaux relatifs à l'étude des réseaux stochastiques. En regardant une vingtaine d'années en arrière, il est frappant de noter l'intégration progressive de ces études dans le domaine des probabilités appliquées en général.

Dans la première moitié du vingtième siècle, des mathématiciens d'Europe de l'Est comme Kolmogorov, Khinchin, Pollaczek, Takacs, ... et quelques ingénieurs en Europe de l'Ouest (Erlang, Engset, Palm, ...) contribuent principalement à ce domaine de recherche sur des réseaux de type téléphonique qui sont les seuls vrais réseaux de cette époque. Au tournant des années 1960, la production sur une échelle industrielle d'ordinateurs donne une impulsion à la modélisation probabiliste de réseaux de machines s'échangeant des flots de requêtes. Les problèmes concernent principalement les temps de traitements de requêtes sur un système centralisé ou encore les délais, le taux d'occupation des différents nœuds du réseau. C'est, mathématiquement, l'époque des réseaux à forme produit : à l'équilibre le réseau se comporte comme un ensemble de nœuds indépendants. Un effort important est produit jusque dans les années 1970 pour étendre ces résultats à des réseaux aussi généraux que possible et, par là même, essayer de comprendre l'origine de cette indépendance asymptotique. Sur le plan technique, les systèmes à un, voire plusieurs, serveurs sont étudiés de façon assez exhaustive par des méthodes analytiques notamment. L'ensemble des résultats obtenus est assez disparate en raison notamment de la grande variété des systèmes analysés, la nomenclature de Kendall ( $G/G/1$ ,  $M/G/k$ , ...) donnera un peu de cohérence au domaine sans toutefois le structurer complètement.

Au début des années 1970, des systèmes distribués de taille significative sont enfin utilisés : les réseaux locaux (ensemble d'ordinateurs reliés par un canal de communication sur lequel les temps de propagation sont très petits) et les premiers essais de ce qui sera plus tard l'Internet. La conception d'algorithme s'impose alors de plus en plus sur le devant de la scène. C'était déjà le cas pour les serveurs parallèles ou bien les systèmes d'exploitation (scheduling) dans la période précédente, mais le cadre distribué donne un relief particulier à ces questions : existe-t-il un algorithme commun à chaque nœud d'un tel système de sorte que l'ensemble "s'auto-organise" ? Mathématiquement la question est, par exemple, de déterminer le comportement *global* d'un système dont chacune des composantes est régie par un algorithme utilisant une information locale. L'ergodicité ou la transience d'un modèle markovien associé au système distribué est un des sujets de ce type d'étude.

Les années 1980-1990 voient le développement exponentiel de l'Internet et avec lui des remises en cause. Les processus de Poisson, si utiles pour l'analyse mathématique, ne peuvent pas y être utilisés sans précaution ("The failure of Poisson modeling"). En raison des transferts de gros fichiers sur le réseau, des

phénomènes de dépendance longue sont mis en évidence, ce qui obère l'utilisation de ces processus. Et bien sûr, il devient très difficile d'expliciter les distributions à l'équilibre des diverses caractéristiques dans un tel contexte. Rétrospectivement toutefois, les processus de Poisson (ou, plus généralement, les processus à dépendance faible) ont été écartés un peu trop vite. Il n'est pas clair qu'à une certaine échelle de temps ces processus ne soient pas encore pertinents. De toutes façons, pragmatiquement, l'hypothèse d'indépendance est dans une certaine mesure inéluctable : peu de techniques en dehors du cadre markovien (et donc avec de l'indépendance) permettent d'investiguer actuellement de façon explicite le comportement à l'équilibre de ces réseaux. Sur le plan mathématique, la dynamique d'un réseau tel qu'Internet est encore largement inconnue. Si les études concernant *un* nœud saturé sont nombreuses, et rigoureuses pour certaines, celles décrivant la coexistence de connexions Internet sur plusieurs nœuds sont pour la plupart assez empiriques. Pour l'instant, une analyse mathématique au sens strict est probablement hors de portée.

Sur le plan des méthodes, au début des années 1990, des techniques de renormalisation ont été introduites pour étudier les probabilités invariantes des grands réseaux avec perte (qui ont une forme produit mais dont l'expression fait intervenir des sommes combinatoires complexes), ainsi que les réseaux dont la mesure invariante n'est pas produit. Des études sur la convergence des processus vers des systèmes dynamiques déterministes et sur l'interaction d'échelles de temps différentes pour un même système sont apparues peu à peu. Les méthodes de convergence de processus, de calcul stochastique ont fait ainsi progressivement leur entrée dans l'étude de ces réseaux.

De cette façon, un domaine assez délimité, à la terminologie un peu énigmatique, autrefois appelé "*théorie des files d'attente*" s'est quelque peu désenclavé pour intégrer progressivement le domaine d'application des méthodes probabilistes classiques. En retour, les réseaux stochastiques fournissent un vaste champ d'expérience et de développement pour ces techniques. Comme on l'a aussi noté, les questions de conception d'algorithmes jouent un rôle de plus en plus grand dans les études. L'analyse mathématique d'un algorithme est d'une certaine façon la deuxième phase de l'étude d'un système, la première étant de cerner les contours algorithmiques du problème.

Le document est organisé de la façon suivante : le chapitre 1 est consacré aux méthodes de renormalisation dans l'étude des réseaux stochastiques. Le chapitre 2 passe en revue l'étude mathématique de plusieurs algorithmes distribués. Les annexes A et B de l'appendice rappellent le contexte ainsi que le fonctionnement des algorithmes examinés dans le chapitre 2.

## Renormalisation des Réseaux

### 1. Un Peu d'Histoire

**Les Réseaux avec Forme Produit.** Un réseau de Jackson est un ensemble de  $N$  files d'attente FIFO, i.e. avec la discipline de service premier arrivé premier servi, avec le fonctionnement suivant : pour  $1 \leq i \leq N$ , la  $i$ -ième file d'attente (appelée aussi le  $i$ -ième nœud du réseau) délivre un service exponentiel de paramètre  $\mu_i$  et les clients arrivent dans le réseau à la file  $i$  suivant un processus de Poisson de paramètre  $\lambda_i$ . Une fois servi par la file  $i$ , le client passe à la file  $j$  avec probabilité  $p_{ij}$  (avec  $p_{ii} = 0$ ) ou quitte définitivement le réseau avec la probabilité résiduelle.

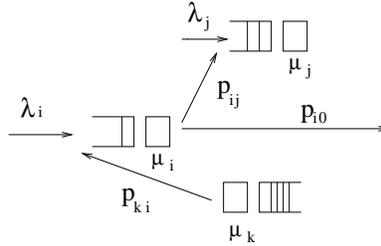


FIG. 1. Un réseau de Jackson

À ce réseau est associé un ensemble d'équations dites de trafic, i.e. un vecteur  $(\bar{\lambda}_i) \in \mathbb{R}_+^N$  solution des équations

$$\bar{\lambda}_i = \lambda_i + \bar{\lambda}_1 p_{1i} + \bar{\lambda}_2 p_{2i} + \dots + \bar{\lambda}_N p_{Ni}, \quad 1 \leq i \leq N.$$

Sous la condition que les clients sortent avec probabilité 1 du réseau, il est facile de montrer qu'il existe un unique vecteur positif  $(\bar{\lambda}_i, 1 \leq i \leq N)$  solution. Pour  $1 \leq i \leq N$ , la quantité  $\bar{\lambda}_i$  peut être interprétée comme le flux global "à l'équilibre" passant par le nœud  $i$ , ainsi  $\bar{\rho}_i = \bar{\lambda}_i / \mu_i$  est la charge moyenne par unité de temps passant par ce nœud. Le résultat classique de Jackson [36] (1957) montre que si, à chaque nœud  $i$  du réseau, la charge  $\bar{\rho}_i$  est strictement plus petite que 1, alors le processus de Markov décrivant le réseau a une probabilité invariante avec une forme produit. Autrement dit, sous la condition de charge plus petite que 1, le réseau est stable (le processus de Markov associé est ergodique) et, à l'équilibre, les files d'attente du réseau sont indépendantes. Ce réseau a, par la suite, été généralisé sous de multiples versions : réseaux de Gordon-Newell, Kelly, etc... Ces généralisations concernent les disciplines de services aux nœuds, les diverses classes de clients qui arrivent dans le réseau et qui peuvent avoir des routages dépendant de la classe, etc. En un mot, les extensions de Jackson sont plus hétérogènes, à une caractéristique près cependant : le taux de service délivré par un nœud. Dans toutes ces généralisations, le *taux de*

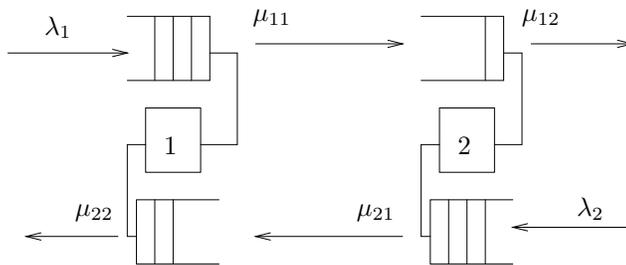


FIG. 2. Le réseau de Rybko-Stolyar/Lu-Kumar

*service peut varier pour les différents clients en attente, mais en aucun cas il ne dépend de la classe des clients.*

Jusqu’au début des années 1990, s’il était bien connu que la forme produit était un phénomène plutôt rare dans le domaine des réseaux stochastiques, il était généralement admis que la condition de charge plus petite que 1 à chaque nœud était la condition de stabilité pour le processus de Markov associé. L’étude de plusieurs exemples, décrits ci-dessous, a montré cependant que ce résultat était faux : Il y a des réseaux pour lesquels  $\bar{\rho}_A < 1$  pour tout nœud  $A$  et qui sont instables.

**Le Réseau de Rybko-Stolyar/Lu-Kumar.** Ce réseau multi-classe a deux serveurs et quatre files d’attente. Pour  $i = 1$  et  $2$ , les clients de classe  $i$  arrivent au nœud  $i$  suivant un processus de Poisson d’intensité  $\lambda_i$  où ils demandent un service de distribution exponentielle de paramètre  $\mu_{i1}$  puis passent à l’autre file d’attente pour être servis au taux  $\mu_{i2}$ . Voir la figure 2.

Ce réseau a été étudié par Rybko et Stolyar [55] (1992) et par Lu et Kumar [44] (1991) dans une version déterministe. Les clients de classe 1 sont prioritaires quand ils accèdent au serveur 2 : aucun client de classe 2 ne peut être servi si un client de classe 1 est présent dans la file 2. De façon symétrique les clients de classe 2 sont prioritaires dans la file 1. À l’intérieur d’une classe, le service se fait dans l’ordre des arrivées. Les charges des nœuds 1 et 2 valent donc respectivement

$$\rho_1 = \lambda_1/\mu_{11} + \lambda_2/\mu_{22} \text{ et } \rho_2 = \lambda_2/\mu_{21} + \lambda_1/\mu_{12}.$$

Les conditions  $\rho_1 < 1$  et  $\rho_2 < 1$  sont donc les conditions habituelles de stabilité rencontrées jusqu’alors. Rybko et Stolyar (1992) ont montré cependant qu’elles ne suffisent pas et qu’une condition supplémentaire “croisée”,

$$(1) \quad \lambda_1/\mu_{12} + \lambda_2/\mu_{22} < 1,$$

est nécessaire pour assurer l’ergodicité du processus de Markov associé dans  $\mathbb{N}^4$ . La figure 3 représente les trajectoires asymptotiques de  $((L_1(t), L_2(t)))$  quand les conditions  $\rho_1 < 1$  et  $\rho_2 < 1$  sont satisfaites mais pas la relation (1). On a fait l’hypothèse  $\mu_{21} = \mu_{11} = +\infty$  pour simplifier, ce qui signifie qu’à leur arrivée, les clients demandent un service nul dans la première file d’attente où ils arrivent.

Habituellement, dans les réseaux à forme produit, quand la condition de charge plus petite que 1 n’est pas satisfaite, le nombre de clients d’au moins un des nœuds tend vers l’infini presque sûrement. L’instabilité des réseaux ayant une charge plus petite que 1 est différente : la condition de charge plus petite que 1 implique que chaque nœud du réseau se vide une infinité de fois presque sûrement mais la durée entre les retours à l’état vide croît linéairement avec le temps. Ainsi le nombre

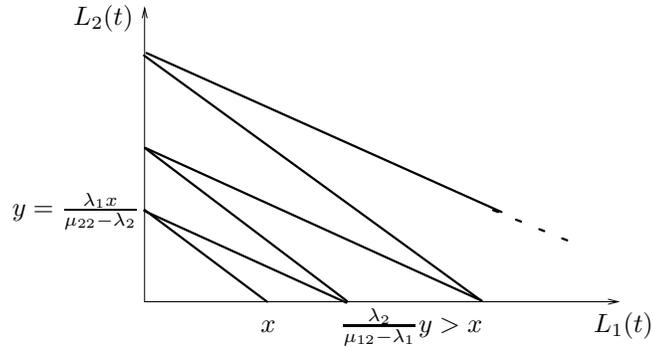


FIG. 3. Trajectoire fluide divergente du réseau de Rybko-Stolyar

de clients à chacun des nœuds du réseau oscille entre des valeurs de plus en plus grandes. Globalement en effet, le nombre total de clients tend presque sûrement vers l'infini.

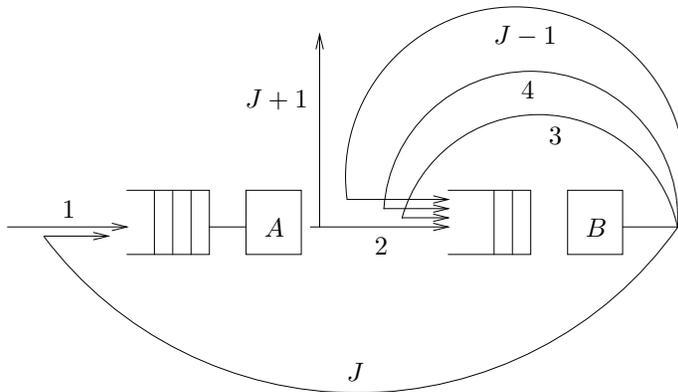
**Le Réseau de Bramson.** Ce réseau est constitué de deux files d'attente servant tous les clients dans l'ordre de leurs arrivées. Les clients arrivent suivant un processus de Poisson d'intensité  $\lambda$  à la file 1. Après cette étape, un client passe à la file 2 pour effectuer  $J - 2$  services exécutés séparément : après le  $k$ -ième service,  $1 \leq k \leq J - 2$ , le client se replace en fin de file d'attente pour recevoir le  $k + 1$ -ième service. Après l'étape  $J - 1$  à la file 2, le client rejoint la file 1 pour ensuite quitter définitivement le réseau. Pour  $1 \leq i \leq J$ , un client reçoit un service dont la durée a une distribution exponentielle de paramètre  $\mu_i$ . Le service d'un client dépend donc de son étape dans son trajet à travers le réseau. Voir la figure 4.

Il n'est pas difficile de constater qu'un processus de Markov de dimension finie ne peut décrire ce réseau puisqu'il faut connaître la classe de chaque client de la file d'attente pour déterminer le taux auquel il sera servi. L'étude de l'ergodicité d'un tel processus de Markov est très délicate, la condition exacte d'ergodicité n'est d'ailleurs pas connue, même dans des cas simples, dès que  $J \geq 3$ . Bramson [7, 8] a montré que le réseau était aussi instable avec des paramètres  $\lambda$ ,  $J$  et  $(\mu_j, 1 \leq j \leq J)$  pour lesquels les charges à chaque nœud  $\lambda/\mu_1 + \lambda/\mu_J$  et  $\lambda/\mu_2 + \dots + \lambda/\mu_{J-1}$  sont strictement plus petites que 1. La divergence du réseau est similaire à celle du réseau de Rybko et Stolyar.

Ces exemples montrent que l'hétérogénéité seule peut déstabiliser un réseau : même si, pour chaque nœud du réseau, la charge *moyenne* de travail qui arrive est strictement plus petite que sa capacité, le réseau peut osciller de telle sorte que le nombre total de requêtes dans le réseau diverge. Pour ces contre-exemples, chaque nœud du réseau se vide une infinité de fois mais globalement le réseau diverge, cette situation est impossible dans les réseaux classiques. Ces réseaux avec des trafics hétérogènes sont regroupés sous l'appellation *réseaux multi-classe*.

## 2. Renormalisation Fluide des Réseaux Stochastiques

Les travaux décrits précédemment ont conduit au développement de techniques de renormalisation pour étudier le comportement macroscopique de tels réseaux. L'espace d'états dénombrable  $\mathcal{S}$  est inclus dans un espace muni d'une norme  $\|\cdot\|$

FIG. 4. Le réseau de Bramson avec  $J$  étapes

(typiquement  $\mathbb{R}_+^d$ ) et est de plus localement fini, i.e. chaque boule de l'espace normé ne contient qu'un nombre fini de points de  $\mathcal{S}$ . Pour  $t \geq 0$  et  $x \in \mathcal{S}$ ,  $X(x, t)$  décrit l'état du réseau à l'instant  $t$  quand son état initial vaut  $x$ . On suppose que  $X(x, t)$  est intégrable. Le processus renormalisé  $\bar{X}$  associé est donné par

$$\bar{X}(x, t) = \frac{1}{\|x\|} X(x, \|x\|t), \quad x \in \mathcal{S}, t \geq 0.$$

Le temps est accéléré proportionnellement à la taille de l'état initial, et la variable spatiale est renormalisée avec l'inverse de cette taille. Ce type de renormalisation est tout à fait adapté aux processus de Markov qui sont localement des marches aléatoires, ce qui est le cadre naturel pour de nombreux réseaux de files d'attente. Remarquer que l'état initial du processus ( $\bar{X}(x, t)$ ) est de norme 1. Le comportement macroscopique de l'état du réseau s'étudie alors en faisant tendre  $\|x\|$  vers l'infini.

**DEFINITION 1.** *Une limite fluide du processus de Markov ( $X(t)$ ) est une valeur d'adhérence des processus ( $\bar{X}(x, t)$ ) quand  $\|x\|$  tend vers l'infini.*

Par exemple, si ( $X(x, t)$ ) est une marche aléatoire dans  $\mathbb{R}$  dont la moyenne des accroissements vaut  $\delta$ , la seule limite fluide positive possible est donnée par la fonction  $t \rightarrow 1 + \delta t$ . La renormalisation a gommé toutes les fluctuations pour ne garder que la dérive moyenne. La marche aléatoire ( $X(x, t)$ ) peut être vue comme une perturbation stochastique de la fonction  $t \rightarrow 1 + \delta t$ . Pour une large classe de réseaux, ce résultat peut être généralisé : l'état renormalisé du réseau converge vers la solution d'une équation différentielle *déterministe* ordinaire (avec des conditions aux bords éventuellement). L'état du réseau se décrit alors comme une perturbation stochastique d'une équation différentielle ordinaire.

Les idées de renormalisation sont anciennes, notamment en physique statistique, elles permettent d'étudier les comportements transitoires de systèmes de particules. Dans le cadre de processus diffusifs, ces perturbations ont été très étudiées. Voir Khasminski [35] (1960) et Freidlin et Wentzell [29] (1979) par exemple. Pour le cas des marches aléatoires réfléchies dans  $\mathbb{N}^d$ , les premiers travaux dans cette optique sont de Malyshev [46] et ses co-auteurs et de Dupuis et Williams [19] pour étudier l'ergodicité des diffusions avec réflexion sur les bords de l'orthant  $\mathbb{R}_+^d$ . Voir aussi Dai [14] et Stolyar [57].

**Système dynamique limite.** On suppose maintenant que l'espace d'états  $\mathcal{S}$  approxime un cône  $H$  de l'espace normé sous-jacent au sens où, pour tout  $\alpha \in H$ , il existe une suite  $(x_n)$  de  $\mathcal{S}$  telle que  $x_n/n$  converge vers  $\alpha$ . Pour  $\alpha \in H$ , on suppose de plus que la limite

$$T_t(\alpha) \stackrel{\text{def.}}{=} \lim_{n \rightarrow +\infty} \frac{1}{n} X(x_n, nt)$$

existe et ne dépend pas de la suite  $(x_n)$  considérée. Il n'est alors pas très difficile, via la propriété de Markov de  $(X(x, t))$ , de montrer que  $(T_t)$  possède la propriété de semi-groupe sur  $H$ ,

$$T_{s+t} = T_t \circ T_s, \text{ pour } s, t \geq 0.$$

Dans ce cadre, les limites fluides du processus de Markov sont données par des processus indexés par la boule unité de  $H$ ,

$$(T_t(\alpha)), t \geq 0, \quad \alpha \in H, \|\alpha\| = 1.$$

La renormalisation fluide consiste de ce point de vue à remplacer un processus de Markov  $(X(x, t))$  par un système dynamique  $(T_t(\alpha))$  qui est une description au premier ordre du processus initial. Pour certains exemples,  $(T_t(\alpha))$  est simplement le flot associé à une équation différentielle déterministe ordinaire. C'est le cas par exemple des réseaux de Jackson (via le problème de Skorokhod). Voir Chen et Mandelbaum [12].

**Stabilité des réseaux.** Rybko et Stolyar [55] donne un critère d'ergodicité pour des processus de Markov généraux en terme de limites fluides. Le théorème suivant intègre ce critère et un autre résultat dû à Filonov [24] dans le cas des chaînes de Markov, voir Robert [1].

**THEOREM 1** (Filonov/Rybko et Stolyar). *S'il existe un temps d'arrêt  $\tau$  et  $\varepsilon > 0$  tels que*

$$\limsup_{\|x\| \rightarrow +\infty} \frac{\mathbb{E}_x(\tau)}{\|x\|} < +\infty \text{ et } \limsup_{\|x\| \rightarrow +\infty} \mathbb{E}_x \left( \frac{\|X(\tau)\|}{\|x\|} \right) \leq 1 - \varepsilon,$$

*alors le processus de Markov  $X$  est ergodique.*

Autrement dit, si le processus de Markov est de norme plus petite que  $(1-\varepsilon)\|x\|$  en un temps d'arrêt dont la moyenne est de l'ordre de  $\|x\|$ , alors il est ergodique. Le corollaire habituellement utilisé de ce théorème se résume ainsi : s'il existe  $T > 0$  tel que toutes les limites fluides sont en 0 à l'instant  $T$ , alors le processus de Markov est ergodique. Donc, en première approximation, si le système dynamique des limites fluides a 0 comme point stable *global*, le réseau stochastique est stable. Ce résultat est en pratique très satisfaisant. Dai [14] a étendu ce critère en terme de limites fluides au cas où l'espace d'états est continu.

Pour résumer, concernant les réseaux, l'étude des limites fluides a principalement deux avantages :

- (1) Décrire le comportement macroscopique.
- (2) Donner un critère de stabilité.

La renormalisation est une des rares techniques possibles pour analyser les processus de Markov de réseaux non standards.

### Les Limitations des Résultats Actuels.

Les techniques de limites fluides se sont généralisées depuis une dizaine d'années,

elles ont contribué à une meilleure compréhension de la dynamique des réseaux avec des trafics hétérogènes. C'est actuellement un outil incontournable pour ce type d'étude. Il n'en reste pas moins que la vue que nous avons actuellement des réseaux multi-classe est encore très incomplète, de nombreux aspects importants sont encore obscurs : par exemple, le comportement d'un réseau multi-classe FIFO avec seulement deux nœuds n'est actuellement pas connu. Les points difficiles sont actuellement les suivants.

### A. Les Espaces de Chaînes de Caractères.

On considère un réseau de files d'attente FIFO où arrivent différentes classes de clients ayant des taux d'arrivée, des taux de service et des routages dépendant de leur classe. L'élément important pour ces réseaux est que l'évolution du nombre total de clients de chaque classe ne se décrit pas facilement. Pour représenter de façon markovienne l'évolution de ces réseaux, il est nécessaire de connaître la classe  $c \in C$  du client à la première place de la file d'attente, à la deuxième place, etc... L'état du nœud est donc représenté par une chaîne de caractères  $(c_i)$  où  $c_i$  est la classe du  $i$ -ième client dans la file d'attente. L'espace d'états  $\mathcal{C}$  est l'ensemble des suites finies de caractères à valeurs dans un espace fini  $C$ ,

$$\mathcal{C} = \{(c_i, 1 \leq i \leq n) : n \in \mathbb{N}, (c_i) \in C^n\}.$$

Cet espace est bien sûr dénombrable mais inclus dans un espace de dimension infinie  $C^{\mathbb{N}}$ . Si les clients de classe  $c \in C$  arrivent au taux  $\lambda_c$  et sont servis au taux  $\mu_c$ , pour un réseau composé d'une *seule* file d'attente, les transitions  $Q = (Q(c, d), c, d \in C)$  sont alors données par, pour  $a \in C$ ,

$$Q[c, c \cdot a] = \lambda_a, \quad Q[c, S(c)] = \mu_{c_1},$$

si  $c = (c_1, \dots, c_n)$ ,  $c \cdot a$  est la concaténation de  $c$  et  $a$ ,  $c \cdot a = (c_1, \dots, c_n, a)$  et  $S(c) = (c_2, \dots, c_n)$ .

Pour le réseau de Bramson décrit plus haut, l'espace d'états est

$$\mathcal{S} = \left\{ (c, d) : \begin{array}{l} n \in \mathbb{N}, c = (c_i, 1 \leq i \leq m) \in \{1, J+1\}^m, \\ m \in \mathbb{N}, d = (d_j, 1 \leq j \leq n) \in \{2, \dots, J\}^n. \end{array} \right\}$$

La matrice de sauts est donnée par,  $Q[(c, d), (c \cdot 1, d)] = \lambda$  et

$$\begin{aligned} Q[(c, d), (S(c), d \cdot 2)] &= \mu_1, \text{ si } c_1 = 1 \\ Q[(c, d), (S(c), d \cdot 2)] &= \mu_{J+1}, \text{ si } c_1 = J+1, \\ Q[(c, d), (c, S(d) \cdot (k+1))] &= \mu_k, \text{ si } d_1 = k < J, \\ Q[(c, d), (c \cdot (J+1), S(d))] &= \mu_k, \text{ si } d_1 = J. \end{aligned}$$

Si  $\|x\|$  est la longueur de la chaîne de caractères  $x$ ,  $\|\cdot\|$  constitue bien une norme et il y a  $|C|^n$  états de norme  $n \geq 1$ . À comparer avec les réseaux avec priorité, du type Rybko Stolyar où le nombre d'états de norme  $n$  est linéaire en  $n$ . Quand la norme tend vers l'infini, la composition interne de la chaîne de caractères a un très grand nombre de degrés de liberté et donc autant de possibilités pour la trajectoire qui part d'un vecteur de taille  $n$ .

Ces systèmes sont très délicats à étudier, nombre de notions sont encore à définir pour poser correctement les bases d'une définition correcte de la renormalisation de ces réseaux. Il n'est donc plus question d'envisager de résoudre, a priori, une équation différentielle déterministe ordinaire dans un espace  $\mathbb{R}^d$ . La renormalisation fluide en  $1/\|x\|$  de  $X(x, t\|x\|)/\|x\|$  n'est bien sûr pas pertinente dans ce contexte.

L'étude de Bramson suggère une renormalisation du type suivant : pour un nœud donné,

- l'état initial  $X_N(0) = x_N = (x_i, 1 \leq i \leq n)$  avec  $x_i \in \{1, \dots, J\}$  est "régulier" au sens où, pour  $u \geq 0$  et  $j \in \{1, \dots, J\}$ ,

$$h_j(0, u) \stackrel{\text{def.}}{=} \lim_{N \rightarrow +\infty} \frac{1}{N} \sum_{i=0}^{\lfloor uN \rfloor} \mathbb{1}_{\{x_i=j\}},$$

existe,  $h_j(0, \cdot)$  est une fonction positive. On peut interpréter  $h_j(0, \cdot)$  comme le profil de densité des éléments de classe  $j$  dans l'état initial.

Si la dynamique du réseau est telle que la quantité

$$h_j(t, u) \stackrel{\text{def.}}{=} \lim_{N \rightarrow +\infty} \frac{1}{N} \sum_{i=0}^{\lfloor uN \rfloor} \mathbb{1}_{\{x_i(t)=j\}},$$

existe pour tout  $t \geq 0$ , l'étude du réseau pourrait alors se ramener à celle de l'évolution des profils  $t \rightarrow h_j(t, \cdot)$ . Ce point de vue a été développé au niveau fluide par Dumas [18] en utilisant des fonctions  $h_j$  constante par morceaux. Si cette approche est naturelle, elle n'est pas en revanche très commode en pratique. De plus, une question reste de toutes façons ouverte : que se passe-t-il si l'on part d'un état initial grand mais pas "régulier" ? L'étude des quelques cas connus suggèrent que même si l'état n'est pas régulier, la dynamique du réseau aurait tendance à le "régulariser". Cette remarque est plus spéculative que fondée sur des résultats concrets. Voir la section 3.1 pour un exemple de ce type d'approche.

Il y a très peu de travaux dans ce domaine (en dehors de ceux de Bramson [7]). Voir les travaux de Gajrat *et al.* [30] sur l'évolution de certaines chaînes de caractères qui étendent ceux de Dynkin et Maljutov [20] dans le cas des marches aléatoires sur le groupe libre. Les applications de ces travaux aux réseaux multi-classe sont cependant limitées : les réseaux correspondants ont un seul nœud et la dynamique ne dépend que d'un nombre borné de caractères au début de la chaîne.

## B. Transience et Divergence des Limites Fluides.

Le théorème de Filonov/Rybko et Stolyar qui établit une relation entre la stabilité du réseau et le fait que toutes ses limites fluides reviennent à 0. Pour la transience, le résultat correspondant serait d'établir une relation entre la divergence des limites fluides et la transience du réseau. Seuls quelques résultats assez partiels sont disponibles dans ce domaine, Dai [14] et Meyn [50] par exemple. Essentiellement, ces critères établissent que si la norme des limites fluides divergent vers  $+\infty$  uniformément par rapport au point départ (qui est, rappelons-le, sur la boule unité du cône associé), alors le processus de Markov est transient. Ce critère est l'analogue du critère d'ergodicité de Filonov/Rybko et Stolyar qui exprime que, si l'on revient vers 0 uniformément par rapport au point de départ (situé sur la boule unité, rappelons le), alors il y a ergodicité.

En pratique, pour les cas difficiles, ce type de critère n'est pas applicable. Typiquement la situation est la suivante : il est possible de trouver un sous-ensemble de l'espace d'états composés d'états relativement "réguliers" pour lequel on peut définir des limites fluides qui sont divergentes. En dehors de ces états, il n'est pas clair que les limites fluides divergent, si tant est que l'on puisse définir celles-ci. C'est par exemple le cas du réseau de Bramson [7] décrit précédemment. Sur une certaine classe d'états initiaux dont la norme tend vers l'infini, Bramson montre

que, sous certaines conditions, la norme du processus de Markov associé tend vers l’infini. Des techniques de grandes déviations lui permettent de montrer que les processus de Markov associés restent dans le voisinage de ces “limites fluides” et donc tendent vers l’infini avec probabilité positive. D’une certaine façon, c’est le problème de la stabilité des trajectoires fluides qui est la propriété importante ici. Pukhal’skiĭ et Rybko [54]) donne un critère de transience (avec un grand nombre de conditions) dans cet esprit.

### C. L’Aléatoire Résiduel.

La vue qui prévaut généralement sur les limites fluides pour l’étude des réseaux stochastiques peut se résumer ainsi : L’état d’un réseau est une perturbation stochastique d’une fonction déterministe. Autrement dit, la résolution d’une équation différentielle déterministe permet d’obtenir le comportement macroscopique du réseau (quitte à éliminer des “fausses solutions” au passage). Si cette approche est effective dans de nombreux cas de réseaux multi-classe, en particulier les réseaux avec des priorités, elle ne couvre pas complètement tous les cas pratiques. En effet, si la renormalisation gomme en effet toutes les fluctuations à la limite, *elle ne supprime pas toutes les composantes aléatoires*. Certaines des composantes aléatoires de ces réseaux ne font pas partie de la partie diffusive et donc restent après le passage à la limite. Cet aléatoire résiduel joue un rôle dans les questions de transience notamment : une limite fluide peut aller à l’infini sans que le système soit transient, il suffit que le processus décroche de façon aléatoire de la trajectoire qui va à l’infini. Voir Dumas [17] et Bramson [11].

**2.1. En Guise de Conclusion.** Si les renormalisations fluides sont maintenant un outil très courant dans l’étude des réseaux, il faut remarquer que peu de résultats généraux ont été obtenus dans ce domaine. Le développement passe encore maintenant par l’étude d’exemples pour dégager un ensemble consistant de techniques. À cet égard le programme de Bramson [9, 10] est tout à fait intéressant, il consiste à exhiber des fonctions de Lyapounov (qui ressemblent à des entropies) pour les limites fluides de réseaux de files d’attente multi-classe. Il est possible que ces fonctions aient une structure (qui pour l’instant n’est pas claire) qui pourrait permettre de mieux comprendre la dynamique de ces systèmes complexes.

## 3. Contributions aux Méthodes de Renormalisation

**3.1. Processus à Valeurs dans les Chaînes de Caractères.** Les réseaux multi-classe avec la discipline de service FIFO ont une représentation markovienne qui s’exprime en terme de vecteurs de chaînes de caractères qui interagissent. Comme il a été déjà expliqué, il y a actuellement peu de méthodes pour étudier ces objets, voire pas du tout. Le cas le plus simple est bien sûr celui du réseau à un nœud mais dans ce cas, la complexité de l’aspect multi-classe ne joue pas vraiment puisque la somme totale des services demandés par tous les clients en attente est une marche aléatoire réfléchie sur  $\mathbb{R}_+$  dont l’étude ne pose pas de difficulté. Le modèle multi-classe de difficulté minimale dans ce domaine est donc celui du réseau de Bramson décrit par deux chaînes de caractères qui est, pour l’instant, trop complexe.

L’approche que nous avons suivie ici a consisté à étudier l’évolution markovienne d’un processus à valeurs dans les chaînes de caractères mais dans un domaine connexe des réseaux de files d’attente multi-classe : les processus de bin-packing.

Cette classe de modèles, issue de l'informatique théorique, a de nombreuses applications dans la gestion de stocks, les systèmes de production et les réseaux de communication entre autres. Le modèle se décrit ainsi : une boîte de taille  $C$  sert au taux 1 les pièces qui se trouvent à l'intérieur. Les pièces arrivent suivant un processus de Poisson et la taille de chacune d'entre elles est une variable aléatoire à valeurs dans un espace fini. Comme la somme des tailles des pièces qui se trouvent dans la boîte ne peut excéder  $C$ , celles qui ne peuvent rentrer sont stockées dans une file d'attente dans l'ordre des arrivées. Une fois dans la boîte, une pièce reste un temps exponentiel de paramètre 1. Dans ce cadre, la classe d'un client est la taille de la pièce. Plusieurs disciplines de service ont été étudiées : Au moment du départ d'une pièce de la boîte ou de l'arrivée d'une pièce dans le système,

– **Next Fit.**

L'algorithme consiste à tester si la pièce en tête de file d'attente peut aller dans la boîte et à la rentrer le cas échéant.

– **First Fit.**

L'algorithme consiste à tester dans l'ordre des arrivées, si *chacune* des pièces dans la file d'attente peut rentrer. Voir la figure 5.

La description markovienne de Next-Fit est assez simple : il suffit de connaître les tailles des pièces dans la boîte (en nombre fini rappelons-le), la taille de celle en tête de file et le nombre total de pièces dans la file d'attente. Globalement, c'est un modèle unidimensionnel gouverné par une chaîne à espace d'états fini. Une étude, incluant l'expression explicite de la mesure invariante, a été menée par Kipnis et Robert [15].

La description markovienne de First-Fit fait, elle, intervenir la description complète de la file d'attente en tant que chaîne de caractères puisqu'à chaque départ de la boîte, il est nécessaire d'examiner, dans l'ordre des arrivées, toutes les pièces qui peuvent éventuellement être incluses dans la boîte. Noter que la dynamique fait donc intervenir toute la chaîne de caractères et non une partie bornée de celle-ci comme dans les travaux de Gajrat *et al.* [30].

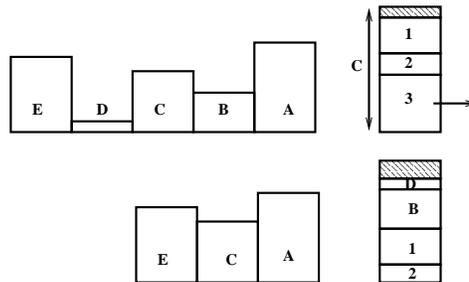


FIG. 5. File d'attente de l'Algorithme Next-Fit après le départ de 3

**Deux tailles de pièces : des limites fluides aléatoires.** Dans Dantzer, Haddani et Robert, le cas où il y a seulement deux tailles possibles 1 et  $a > 1$  de pièces est examiné. Si l'étude proprement dite du processus de Markov ne pose pas de problème, un phénomène intéressant a été montré. Pour cela, si  $L_i(t)$  est le

nombre de pièces de taille  $i \in \{1, a\}$  à l'instant  $t$  et si  $L_1(0) = N$  et  $L_a(0) = 0$ , alors si

$$\bar{L}_i(t) = \lim_{N \rightarrow +\infty} \frac{1}{N} L_i(NT), \quad i = 1, a,$$

$(\bar{L}_1(t), \bar{L}_a(t))$  est la limite fluide du couple  $(L_1(t), L_a(t))$ . Dans le cas où le système est transient, la fonction  $(\bar{L}_1(t), \bar{L}_a(t))$  diverge à la manière du réseau de Rybko et Stolyar, voir la figure 6, avec une différence intéressante : les pentes successives  $(Y_n)$  des chemins linéaires par morceaux sont *aléatoires* et forment une *chaîne de Markov* à espace d'états finis. La suite  $(Y_n)$  est l'aléatoire résiduel des limites fluides au sens de la partie B de la section 2. L'explication est assez simple, en revenant à l'échelle temporelle normale au moment des instants de changement de direction de la limite fluide, la pente de la section suivante est en fait fonction de la direction de fuite à l'infini d'une chaîne de Markov transiente dirigeant la dynamique locale. Le mécanisme de renormalisation "gomme" les perturbations stochastiques mais pas ce type de caractéristique aléatoire.

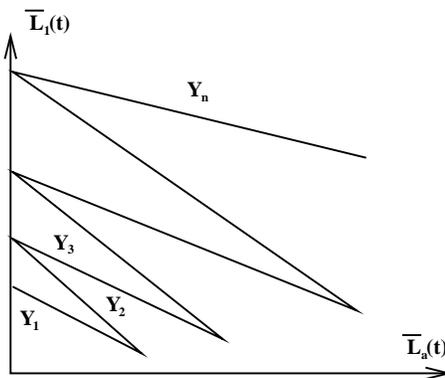


FIG. 6. Divergence de l'Algorithme de Bin-Packing dans le Cas Transient

**Trois tailles de pièces : Dynamique des Chaînes de Caractères.** L'article Dantzer et Robert [5] le cas où les tailles des pièces ont trois valeurs 1, 2, 3 avec les probabilités respectives  $p$ ,  $q$  et  $r$ . La taille de la boîte vaut 4. Ce modèle simple a priori permet de mettre en œuvre certaines des idées développées dans la section 2. Le cas des tailles de boîte arbitraire et de taille de pièce générales augmente de façon significative la combinatoire des divers cas à examiner.

On introduit la notion d'état régulier "smooth state" dans l'esprit des techniques utilisées pour les chaînes de Harris : il ne s'agit pas d'un sous-ensemble de l'espace des états. Cette notion est relative à la distribution de la chaîne de caractères. En prenant  $\mathcal{E}$  l'ensemble des distributions sur l'ensemble des chaînes de caractères de la forme

$$\prod_{i=1}^k \delta_3 \prod_{i=1}^l \left( \frac{q}{q+r} \delta_2 + \frac{r}{q+r} \delta_3 \right) \prod_{i=1}^m (p\delta_1 + q\delta_2 + r\delta_3), \quad k, l, m \in \mathbb{N},$$

une chaîne de caractères dont la distribution est dans  $\mathcal{E}$  est constituée de trois sections : une suite de 3, puis une suite i.i.d. de 2 et 3 avec les proportions respectives

$q/(q+r)$  et  $r/(q+r)$  et enfin une suite i.i.d. de 1, 2, 3 avec les proportions de  $p$ ,  $q$  et  $r$ . La distribution est dite *régulière* si elle est dans la clôture convexe de l'ensemble  $\mathcal{E}$ . L'approche suivie dans Dantzer et Robert [5] consiste essentiellement en deux étapes :

*Régularisation des États.* Il est montré que, quel que soit l'état initial  $x$ , il existe un temps d'arrêt  $U$  tel que l'état de la file d'attente soit régulier au sens défini précédemment. Il est de plus montré que la valeur moyenne de  $U$  et de l'état  $X(U)$  à l'instant  $U$  sont de l'ordre de  $\|x\|$ . L'état se régularise donc "rapidement".

*Système Dynamique Limite Aléatoire dans  $\mathbb{R}_+^2$ .* En regardant la suite  $(U_n)$  de temps d'arrêt induite par le temps d'arrêt  $U$ , l'étude asymptotique de la suite  $(X(U_n))$  se ramène à celle d'un produit de matrices aléatoires dans  $\mathbb{R}_+^2$ . De cette façon, modulo plusieurs estimations, l'étude complète de la transience et de l'ergodicité peut alors être menée. Ici aussi le processus renormalisé limite est aussi aléatoire.

En conclusion, ces modèles permettent de tester certaines des méthodes esquissées dans la section 2 dans un cadre assez simple avec une seule chaîne de caractères. Dans cette perspective, ils constituent une classe intéressante de modèles multi-classe. Il n'en reste pas moins que la combinatoire associée avec la composition de la boîte complique sérieusement l'étude dès que le nombre de tailles possibles est significatif.

**3.2. Processus à Valeurs Mesure.** Les processus à valeurs mesure interviennent de manière naturelle quand un ensemble non ordonné de points évolue suivant une certaine dynamique. Il est quelquefois commode de représenter ces points par une mesure de façon à pouvoir utiliser la topologie existante sur les mesures et les résultats associés.

Exemples.

– *La discipline de service processor-sharing.*

À l'instant  $t$ , chaque client de la file d'attente est servi à la vitesse  $1/L(t)$  si  $L(t)$  est le nombre de clients. Si  $x_1(t), x_2(t), \dots, x_{L(t)}(t)$  est l'ensemble des services résiduels à cet instant, en posant

$$Z(t) = \sum_1^{L(t)} \delta_{x_i(t)},$$

alors  $(Z(t))$  est un processus de Markov à valeurs dans les mesures ponctuelles sur  $\mathbb{R}_+^2$  satisfaisant l'équation différentielle stochastique

$$Z(t+dt) = \delta_\sigma \mathcal{N}_\lambda(dt, d\sigma) + T_{1/L(t-)} Z(t),$$

où  $\delta_a$  est la masse de Dirac en  $a \in \mathbb{R}_+$  et  $T_x$  est la translation de  $x$  sur les mesures ponctuelles et  $\mathcal{N}_\lambda(dt, d\sigma)$  est le processus d'arrivées des clients avec leur service (composante  $\sigma$ ). Avec la convention que  $\delta_0 \equiv 0$ , i.e. quand un client a un service nul, il disparaît de la mesure aléatoire. Cette représentation a été utilisée par Jean-Marie et Robert [10] (1994) pour l'étude de la renormalisation de cette file d'attente sous l'hypothèse de surcharge  $\lambda E(\sigma) > 1$ . Cet article étudie le profil limite de la mesure aléatoire décrivant la file d'attente processor-sharing en surcharge. Plus récemment, dans ce cadre markovien, Gromoll *et al.* [32] étudie le comportement de cette file d'attente quand  $\lambda E(\sigma) \nearrow 1$ .

- *Les files d’attente avec impatience.* Des clients arrivent à une unité de service avec un service et une durée maximum d’attente possible au terme de laquelle, ils quittent le système. Un client sera représenté par un couple  $(x, y)$  où  $x$  est la valeur du service résiduel et  $y$  l’impatience résiduelle. Comme précédemment, un processus de Markov à valeurs dans les mesures ponctuelles sur  $\mathbb{R}_+^2$  peut être utilisé. C’est l’approche utilisée par Doytchinov *et al.* [16] pour étudier la discipline de service qui consiste à servir le client avec la plus petite impatience.
- *Les systèmes de type Bin-Packing.* La mesure ponctuelle composée des tailles des pièces dans la boîte (voir la section précédente) décrit l’évolution de ce système pour la discipline Next-Fit. C’est la description utilisée par Kipnis et Robert [15].

Structurellement, la situation est plus établie que pour les chaînes de Markov à valeurs dans les chaînes de caractères. L’espace d’états est l’espace métrique complet  $\mathcal{M}(\mathbb{R}_+^d)$  des mesures de Radon sur  $\mathbb{R}_+^d$ . Des critères de compacité pour les processus à valeurs dans  $\mathcal{M}(\mathbb{R}_+^d)$  sont disponibles. Voir Dawson [15] par exemple.

La renormalisation donne, en général, un système dynamique qui vérifie une équation différentielle à valeurs mesure pour laquelle il existe peu de techniques de résolution. Dans le cas des exemples que l’on sait traiter, une équation plus ordinaire se cache derrière cette équation. En pratique, l’étude des points fixes de ce système dynamique est une première étape qui n’est pas toujours simple. Voir Gromoll *et al.* [32]. Un autre obstacle de ce cadre est la complexité technique liée aux critères de tension des suites de processus à valeurs mesures. Cela peut mener à un formalisme technique un peu disproportionné, voire lourd, au regard du problème.

Gromoll, Robert et Zwart [1] étudie la renormalisation d’une file d’attente avec impatience et utilisant la discipline processor-sharing. Le système dynamique limite à valeur mesure se décrit en utilisant la solution  $(z(t))$  de l’équation fonctionnelle

$$z(t) = z_0 H \left( \int_0^t \frac{1}{z(s)} ds, t \right) + \lambda \int_0^t H \left( \int_s^t \frac{1}{z(u)} du, t - s \right) ds,$$

où  $H$  est une fonction connue. Les points d’équilibre (des mesures) de la limite fluide s’expriment à partir d’une équation ordinaire de point fixe sur  $\mathbb{R}_+$ .

#### 4. Grands Réseaux

Le développement du réseau Internet et celui plus récent des réseaux mobiles, de capteurs et de réseaux pair à pair conduisent à des modèles de réseaux avec un très grand nombre  $N$  de nœuds (de l’ordre de plusieurs dizaines de milliers de nœuds) et par là même à un ensemble de renormalisations avec ce paramètre. Les questions abordées concernent les points d’équilibre des systèmes dynamiques associés à ces renormalisations ainsi que leur propriété de stabilité : minimum ou maximum local, point selle. Un premier exemple a été étudié récemment.

Antunes *et al.* [2, 1] étudient des réseaux de files d’attente à capacité limitée dans lequel circulent plusieurs classes de clients. Les clients de classe  $k$  arrivent aux taux  $\lambda_k$ , sont servis au taux  $\gamma_k$ , quittent le réseau au taux  $\mu_k$  et requièrent  $A_k$  places.

- Chaque nœud  $i$  a une capacité limitée  $C_i$  de telle sorte qu’un client de classe  $k$  qui ne peut y avoir  $A_k$  places libres est rejeté du réseau.
- Un client de classe  $k$  à un nœud est servi immédiatement au taux  $\gamma_k$ .

- Les clients de classe  $k$  se déplacent dans le réseau suivant une matrice de routage  $(p^{(k)}(i, j))$  dépendant de leur classe.

Ces réseaux sont utilisés pour représenter des réseaux de mobiles, un client de classe  $k$  est la connexion d'un utilisateur mobile demandant la bande passante  $A_k$  dans chaque cellule où celui-ci passe,  $\gamma_k$  est le degré de mobilité d'une cellule de communication à l'autre et  $1/\mu_k$  est le volume moyen total demandé par la connexion.

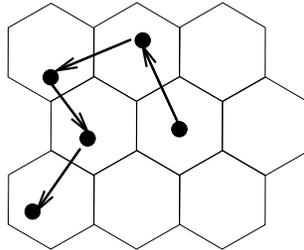


FIG. 7. Le mouvement d'un mobile parmi les cellules du réseau

Sur le plan mathématique, ces réseaux combinent les deux modèles les plus classiques de réseaux de files d'attente :

- (1) Les réseaux de Jackson par le routage des clients dans le réseau.
- (2) Les réseaux avec perte : chaque nœud a une capacité limitée.

Les processus de Markov associés n'héritent pourtant d'aucune des belles propriétés de ces deux classes de réseaux : en général ils ne sont pas réversibles comme le sont les réseaux avec perte classiques et ils n'ont pas de mesure invariante produit comme c'est le cas pour les réseaux de Jackson. Deux régimes limites ont été étudiés.

**Régime Limite de Kelly.** Ce régime limite consiste à accélérer les arrivées et les capacités des nœuds par un facteur  $N$  :  $\lambda_k \rightarrow \lambda_k N$  et  $C_i \rightarrow C_i N$ . C'est un régime limite classique pour étudier les réseaux avec perte, voir l'article de Kelly [38] par exemple. Si  $X_{i,k}^N(t)$  est le nombre de clients de classes  $k$  au nœud  $i$  à l'instant  $t$ , il n'est pas très difficile de montrer que le vecteur  $(X_{i,k}^N(t)/N, 1 \leq i \leq I, 1 \leq k \leq K)$  converge en distribution vers une fonction  $(x(t)) = (x_{i,k}(t), 1 \leq i \leq I, 1 \leq k \leq K)$  vérifiant l'équation

$$\frac{dx}{dt}(t) = F(x(t)) = [F_{i,k}(x(t))],$$

avec

$$F_{i,k}(x) = -(\gamma_k + \mu_k)x_{i,k} + \left( \lambda_k + \gamma_k \sum_j x_{j,k} p^{(k)}(j, i) \right) \tau_i(x)$$

et  $\tau_i(x) = 1$  si  $\sum_r x_{i,r} < c_i$ , et

$$\tau_i(x) = \min \left\{ \frac{\sum_r (\gamma_r + \mu_r) x_{i,r}}{\sum_r [\lambda_r q_r(i) + \gamma_r \sum_j x_{j,r} p^{(r)}(j, i)]}, 1 \right\}$$

sinon.

L'étude des points d'équilibre du système limite, l'ensemble des solutions de l'équation  $F(x) = 0$ , se révèle délicate en raison de la forme de la fonction  $F$  qui est définie par morceaux. Le résultat principal concerne la preuve de l'unicité du

point d'équilibre d'un tel système. La difficulté ici vient du fait que la fonction  $F$  ne semble pas avoir de bonnes propriétés de convexité qui donneraient directement un tel résultat. Les étapes principales de la preuve sont :

- Une formulation duale du problème de l'unicité ;
- Une inégalité de type entropique (apparemment nouvelle) ;
- Une représentation probabiliste d'un ensemble d'équations linéaires.

Même dans des cas avec deux types de clients et deux nœuds, l'expression de cet unique point fixe est assez compliquée.

**Limites Thermodynamiques.** Le deuxième régime limite consiste à supposer que tous les nœuds ont la même capacité et que leur nombre  $N$  tend vers l'infini. De plus les clients sont routés uniformément parmi les cellules du réseau lors de leurs déplacements. Ce qui revient à faire l'hypothèse d'un graphe complètement connecté. L'état d'un nœud du réseau est un élément de l'ensemble

$$\mathcal{X} = \left\{ n = (n_1, \dots, n_K) \in \mathbb{N}^K : \sum_{k=1}^K n_k A_k \leq C \right\}.$$

En raison de la symétrie du réseau, il est naturel de considérer la mesure empirique du système, i.e. pour  $n \in \mathcal{X}$ ,

$$Y_n^N(t) = \frac{1}{N} \sum_{i=1}^N \mathbb{1}_{\{X_i^N(t)=n\}}.$$

Il n'est pas difficile non plus de montrer que  $(Y_n^N(t), n \in \mathcal{X})$  converge en distribution vers une fonction déterministe  $(y_n(t), n \in \mathcal{X})$  vérifiant l'équation

$$y'(t) = V(y(t)),$$

où  $V$  est le champ de vecteur  $V(y) = (V_n(y), n \in \mathcal{X})$  défini par

$$\begin{aligned} V_n(y) = & \sum_{k=1}^K (\lambda_k + \gamma_k \langle \mathbb{1}_k, y \rangle) \left( y_{n-f_k} \mathbb{1}_{\{n_k \geq 1\}} - y_n \mathbb{1}_{\{n+f_k \in \mathcal{X}\}} \right) \\ & + \sum_{k=1}^K (\gamma_k + \mu_k) \left( (n_k + 1) y_{n+f_k} \mathbb{1}_{\{n+f_k \in \mathcal{X}\}} - n_k y_n \right), \quad n \in \mathcal{X}, \end{aligned}$$

avec  $\langle \mathbb{1}_k, y \rangle = \sum_{m \in \mathcal{X}} m_k y_m$  et  $f_k$  est le  $k$ -ième vecteur unité de  $\mathbb{R}^K$ . Le terme  $\gamma_k \langle \mathbb{1}_k, y \rangle$ , le taux d'arrivée des clients de classe  $k$  venant d'un nœud interne, est dû à la convergence de type champ moyen de ce système.

La fonction  $(y(t))$  est donc un système dynamique sur l'ensemble  $\mathcal{P}(\mathcal{X})$  des probabilités sur  $\mathcal{X}$ . L'ensemble de ses points d'équilibre se décrit assez simplement : il est constitué par les probabilités  $\nu_\rho$  sur  $\mathcal{X}$  définies par

$$\nu_\rho(n) = \frac{1}{K(\rho)} \prod_{\ell=1}^K \frac{\rho_\ell^{n_\ell}}{n_\ell!}, \quad n = (n_\ell) \in \mathcal{X},$$

où  $\rho = (\rho_k, 1 \leq k \leq K) \in \mathbb{R}_+^K$ , satisfait l'ensemble de relations

$$\lambda_k = \rho_k \left( \mu_k + \frac{\gamma_k}{K(\rho)} \sum_{n: n+f_k \notin \mathcal{X}} \prod_{\ell=1}^K \frac{\rho_\ell^{n_\ell}}{n_\ell!} \right), \quad 1 \leq k \leq K$$

et  $K(\rho)$  est la constante de normalisation,

$$K(\rho) = \sum_{m \in \mathcal{X}} \prod_{\ell=1}^K \frac{\rho_\ell^{m_\ell}}{m_\ell!}.$$

Si l'ensemble des points d'équilibre est indexé par un sous-ensemble de  $\mathbb{R}_+^K$ , le système dynamique  $(y(t))$  ne peut se réduire à un système dynamique sur  $\mathbb{R}_+^K$  ce qui réduirait de façon importante la dimension du problème. Noter de plus que le problème de la stabilité des points d'équilibre de  $(y(t))$  n'est pas réduit à un problème similaire sur  $\mathbb{R}_+^K$ .

L'approche qui a été suivie a consisté à introduire deux fonctions d'énergie, une,  $g$ , sur l'ensemble  $\mathcal{P}(\mathcal{X})$ , l'autre,  $\phi$ , sur  $\mathbb{R}_+^K$ . La fonction  $g$  est en fait une fonction de Lyapounov pour  $(y(t))$ . Le résultat principal montre que les extrema de  $g$  et  $\phi$  sont en correspondance biunivoque, *de même que leurs propriétés de stabilité*, ce qui achève la réduction de la dimension du problème de la stabilité. Un exemple de réseau avec plusieurs points d'équilibre stables est ensuite exhibé en utilisant la réduction obtenue. C'est à notre connaissance, le premier exemple de réseau stochastique où cette propriété est démontrée de façon rigoureuse. Une situation similaire a été rencontrée par Gibbens *et al.* [31] dans l'étude du routage dans des réseaux avec perte. Dans ce cas aussi, les points d'équilibre sont aussi indexés par un sous-espace de petite dimension. Gibbens *et al.* [31] utilise des approximations et des résultats numériques pour mettre en évidence des cas où plusieurs points d'équilibre stables coexistent.

Le résultat obtenu pour la limite thermodynamique suggère bien évidemment une propriété de métastabilité du processus de Markov. Malheureusement, c'est pour l'instant une question ouverte. Les techniques disponibles dans ce domaine ne concernent essentiellement que certains processus réversibles. Voir Bovier [5, 6] et Olivieri et Vares [52] par exemple.

**Conclusion.** Cette thématique de recherche des grands réseaux est appelée à jouer un rôle de plus en plus important en raison du déploiement d'architectures virtuelles de grande taille. Ces modèles sont tout à fait prometteurs et pourraient permettre une étude mathématique détaillée de réseaux de taille convenable (ce qui n'est pas le cas actuellement). Il y a là un champ immense de développement pour des méthodes de type physique statistique.



## Algorithmes

### 1. Algorithmes en Arbre

Dans cette partie, on présente une classe d'algorithmes ayant une structure d'arbre sous-jacente. En informatique théorique, ces algorithmes appartiennent la classe des algorithmes "diviser pour régner". Ceux qui nous intéressent ici peuvent se décrire de la façon suivante.

Si  $S$  est un ensemble fini, un algorithme en arbre  $A$  appliqué à l'ensemble  $S$  est défini comme suit :

---

L'ALGORITHME  $\mathcal{A}(S)$

— CONDITION DE TERMINAISON.

Si  $\text{card}(S) < D \quad \rightarrow \text{STOP}.$

— STRUCTURE D'ARBRE.

Si  $\text{card}(S) \geq D$ ,  $S$  est décomposé aléatoirement en une partition de  $G$  sous-ensembles  $S_1, S_2, \dots, S_G$  où  $G$  est une variable aléatoire de distribution donnée.

$\rightarrow$  APPLIQUER  $\mathcal{A}(S_1), \mathcal{A}(S_2), \dots, \mathcal{A}(S_G).$

---

La structure d'arbre est claire, un nœud contient les éléments sur lesquels l'algorithme s'applique. Au terme de l'algorithme, les feuilles de l'arbre final ne comptent que des ensembles de cardinal plus petit que  $D$ .

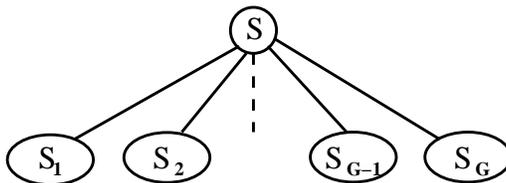


FIG. 1. Décomposition de l'Ensemble  $S$

Si l'ensemble  $S$  est de cardinal  $n$  et les sous-ensembles  $S_i$ ,  $1 \leq i \leq G$  de cardinal  $n_i$ , l'algorithme peut être vu comme la décomposition de l'entier  $n$  en la somme de  $G$  entiers et ainsi de suite jusqu'à ce que  $n$  s'exprime comme la somme d'entiers compris entre 0 et  $D - 1$ , le nombre de composantes étant le nombre de feuilles de l'arbre final.

En pratique l'ensemble  $S$  peut être celui des stations d'un réseau qui veulent transmettre sur un canal de communication. Voir l'annexe 1 pour une description détaillée d'un algorithme astucieux utilisé dans le cadre d'un système distribué.

*Le cas de l'arbre binaire.*

Une autre application est le cas où  $S$  est un ensemble d'éléments que l'on veut stocker de telle sorte qu'il soit facile de déterminer en un petit nombre d'opérations si un élément  $y$  appartient ou non à  $S$ . La technique classique consiste à utiliser une fonction de hachage  $h$  de  $S$  dans  $[0, 1]$  telle que, pour  $x \in S$ , on puisse supposer que l'élément  $h(x)$  soit uniformément distribué sur  $[0, 1]$  et que les variables  $h(x)$ ,  $x \in S$  soient indépendantes. Autrement dit, pour  $y \in X$ ,

$$(2) \quad h(y) = 0.X_1^y X_2^y \dots X_n^y \dots,$$

où les  $X_n^y \in \{0, 1\}$ ,  $n \geq 1$ , sont i.i.d. Ici  $D = G = 2$  et  $S_1$  est l'ensemble des éléments  $y$  pour lesquels  $X_1^y = 0$ ,  $S_2$  étant le reste de l'ensemble  $S$ . Les éléments de  $S$  sont ainsi stockés dans les feuilles de l'arbre associé. On parle dans ce cas de structure d'arbre digital associé. Voir Flajolet [26] pour un survol des nombreux aspects mathématiques et algorithmiques de ces structures de données.

Mathématiquement ces algorithmes sont reliés à plusieurs classes de processus classiques.

– *Processus de Fragmentation.*

On peut voir un algorithme en arbre comme un processus de décomposition d'entiers en entiers inférieurs à  $D$ . Une version continue de ce processus consiste à “casser” une particule de masse  $x$  en plusieurs morceaux puis à recommencer de façon récursive sur chacun de ces morceaux. Ces processus ont été étudiés de façon extensive ces dernières années. Les problèmes considérés sont du nature un peu différente de ceux du cas entier : il s'agit d'obtenir les propriétés de régularité des processus de Markov associés, d'étudier les relations de dualité avec les processus de coalescence, de déterminer le taux de décroissance vers 0 des morceaux obtenus, etc. Voir Bertoin [1] et Miermont [51] et les références de ces travaux.

Un modèle analogue à celui de la décomposition du cas discret a été étudié par Bertoin et Martínez [3] : le processus de fragmentation s'arrête sur un morceau dès que sa taille est plus petite que  $\varepsilon$ . Des théorèmes limites sont obtenus pour le nombre de morceaux en fin de processus quand  $\varepsilon$  tend vers 0. Noter que cela correspond au nombre de feuilles de l'arbre du cas discret.

– *Décomposition récursives aléatoires.*

Dans le cas de l'arbre binaire, l'algorithme en arbre peut être vu comme un processus de découpage de l'intervalle  $[0, 1]$ . On lance  $n$  points au hasard sur  $[0, 1]$ , si  $n \leq D$  c'est terminé, sinon  $n \geq 2$  et on coupe l'intervalle en  $x = 1/2$  et on recommence sur chacun des deux morceaux. Le nombre d'intervalles ainsi obtenus correspond à la taille de l'arbre de l'algorithme. Cette analogie est bien sûr valable pour un processus aléatoire arbitraire de découpage.

Ces décompositions récursives aléatoires ont été étudiées du point de vue de la géométrie (dimension de Hausdorff entre autres) par de nombreux auteurs : Mauldin and Williams [49], Waymire and Williams [60]. Hambly and Lapidus [34] and Falconer [22] considèrent des décompositions de l'intervalle  $[0, 1]$  du point de vue des longueurs des sous-intervalles obtenus. L'intervalle  $[0, 1]$  est représenté par une suite décroissante de longueurs d'intervalles  $(L_n)$  dont la somme vaut 1. Cette description s'apparente à la représentation classique des processus de fragmentation. Voir Pitman [53].

– *Cascades multiplicatives et martingales de Mandelbrot.*

Ces martingales introduites par Mandelbrot [47] puis étudiées par Kahane and Perrière [37] interviennent assez naturellement dans toutes ces procédures récursives. Voir Liu [43].

**Étude du Coût Asymptotique de l’Algorithme.** Si  $C_n$  est la taille de l’arbre pour un ensemble  $S$  de taille  $n$ ,  $C_n$  est le nombre d’étapes pour que l’algorithme se termine. Le comportement asymptotique de  $C_n$  est bien sûr une des questions importantes dans ce contexte. Le temps moyen de traitement par élément de l’ensemble  $S$  vaut  $\mathbb{E}(C_n)/n$ , on peut donc espérer une loi des grands nombres dans cette situation. Ce n’est cependant pas le cas. Dans le cas simple de l’arbre binaire par exemple la suite  $\mathbb{E}(C_n)/n$  oscille autour de  $2/\log 2$  avec une amplitude de l’ordre de  $10^{-5}$ , i.e.

$$\mathbb{E}(C_n) \sim nF(\log_2 n),$$

où  $F$  est la fonction périodique de période 1 définie par

$$F(y) = 4 \int_0^{+\infty} 2^{-\{y - \log_2(x)\}} e^{-x} dx, \quad y \in \mathbb{R}.$$

Ce phénomène, a priori curieux pour un probabiliste, est bien connu des informaticiens. Knuth [39] (1973) l’a montré dans le cas de l’arbre binaire par des méthodes d’analyse complexe.

L’analyse classique de ces algorithmes reposent sur des transformées successives et leurs inversions. Voir Dumas *et al.* [27] pour une approche générique (avec de nombreux exemples) ainsi que les livres de Mahmoud [45] et Szpankowski [58]. On décrit brièvement les étapes de cette approche. La propriété de séparation des lois de Poisson rend naturelle l’utilisation de la *transformée de Poisson*

$$C(x) = \sum_{n \geq 0} C_n \frac{x^n}{n!},$$

il est ensuite facile de vérifier que la fonction  $C$  vérifie une relation du type

$$(3) \quad C(x) = \sum_i C(p_i x) + h(x),$$

où  $(p_i)$  est un vecteur de probabilité et  $h(x)$  est une fonction connue. La deuxième transformation est la *transformée de Mellin* :

$$C^*(s) = \int_0^{+\infty} C(x) x^{s-1} dx,$$

où  $s$  appartient à une bande de convergence  $\mathcal{B}$  du plan complexe. La propriété importante ici est que si on pose  $C_\mu(x) = C(\mu x)$ , alors

$$C_\mu^*(s) = \mu^{-s} r^*(s) C^*(s),$$

ce qui permet, avec l’équation (3), d’exprimer explicitement  $C^*$  en fonction de  $h^*$ . *Inversion de Mellin.* On utilise ensuite un résultat qui lie le comportement asymptotique de  $C(x)$  au voisinage de  $+\infty$  à celui de la première singularité à droite de la bande de convergence de  $C^*$ . Voir Flajolet *et al.* [27].

*Inversion de Poisson.* Cette étape finale consiste à montrer que les comportements à l’infini de  $x \rightarrow C(x)$  et  $(C_n)$  sont similaires. Le cas où il y a une oscillation correspond au cas où il y a une suite de singularités régulièrement espacées sur un axe parallèle à l’axe imaginaire.

**Le Théorème de Renouvellement dans l'Étude Asymptotique.** Nous avons étudié ces questions sous l'angle probabiliste dans Robert [5] et Mohamed et Robert [4] pour une classe assez générale d'algorithmes de division. La méthode d'analyse consiste à donner une interprétation probabiliste à l'équation (3) satisfaite par l'équation de Poisson, à faire ensuite une inversion probabiliste de celle-ci, puis à interpréter l'expression obtenue à l'aide d'une marche aléatoire. Le résultat-clé dans ces questions est le théorème de renouvellement pour obtenir le comportement asymptotique de  $(\mathbb{E}(C_n)/n)$ , l'existence ou non d'oscillations étant liée au fait que le support de la marche aléatoire soit latticiel ou non. Dans certains cas, en utilisant cette représentation probabiliste du coût, il est possible d'obtenir un théorème central limite en utilisant le théorème central limite classique.

Dans le prolongement de ces travaux, nous nous sommes intéressés au cas avec immigration, i.e. quand l'ensemble  $S$  reçoit chaque unité de temps un groupe d'éléments. Fayolle *et al.* [23] (voir aussi Flajolet et Mathys [48]) étudie cette situation dans le cas de l'arbre binaire, l'équation (3) pour la transformée de Poisson devient

$$C(x) = C(\lambda + px) + C(\lambda + qx) + h(x),$$

ici les transformations affines  $\sigma_1$  (resp.  $\sigma_2$ )  $x \rightarrow \lambda + rx$ , avec  $r = p$  (resp.  $q$ ) compliquent l'étude asymptotique de la transformée de Poisson. L'analyse se fait à partir d'un schéma itératif faisant intervenir toutes les compositions possibles de ces deux fonctions qui ne commutent pas

$$(4) \quad C(x) = \sum_{n \geq 0} \sum_{(a_i) \in \{1,2\}^n} h(\sigma_{a_1} \circ \sigma_{a_2} \cdots \sigma_{a_n}(x)).$$

Mohamed et Robert [3] considère une réécriture probabiliste de l'équation précédente qui permet alors d'analyser le comportement asymptotique de la suite  $(C_n)$  dans ce cas. Les techniques développées dans Mohamed et Robert [4] s'appliquent avec un nouvel ingrédient : un processus auto-régressif. La mesure invariante de ce processus permet de déterminer certaines constantes du système. Ici encore, c'est le théorème de renouvellement qui permet d'obtenir le résultat limite.

Pour cet exemple, l'approche probabiliste présente peut-être un avantage au sens où une partie conséquente de Fayolle *et al.* [23] est consacrée à l'estimation asymptotique de la somme (4) quand  $x$  tend vers l'infini, i.e. déterminer les trajectoires (compositions de  $\sigma_1$  et  $\sigma_2$ ) de cette somme qui sont significatives. Le théorème de renouvellement donne en fait directement la réponse, ce qui simplifie cette partie.

L'importance du théorème de renouvellement dans toutes ces questions ne semble pas avoir été relevée dans la littérature du domaine. Dans Erdős *et al.* [21] la version analytique du théorème de renouvellement est utilisée pour étudier un cas assez particulier de processus de séparation. En revanche, dans un autre domaine, celui du dénombrement des orbites périodiques de systèmes dynamiques, suivant la terminologie de Lalley, un résultat "cousin" du théorème de renouvellement a été démontré et utilisé d'une façon similaire par Lalley [40, 41] pour déterminer le comportement asymptotique de ces dénombrements.

**Cas des systèmes dynamiques.** Dans le cas où les éléments de l'ensemble  $S$  sont les mots d'un dictionnaire représentés par un expression similaire à (2), la suite  $(X_n^x)$  n'est plus nécessairement i.i.d. (les lettres d'un mot présentent des dépendances), aussi est-il naturel de considérer que la représentation (2) est toujours valide mais que la suite  $(X_n^x)$  n'est plus i.i.d. mais seulement stationnaire. Il s'agit

bien sûr d'étudier le comportement asymptotique de la suite  $(C_n)$  associée. Dans le cas où la suite  $(X_n^x)$  est générée par l'itération d'une fonction  $\phi$  sur  $[0, 1]$ , ces questions ont été étudiées de façon extensive par Vallée et ses co-auteurs, voir Vallée [59], Clément *et al.* [13] et les références de ces articles. L'approche consiste à étudier les propriétés spectrales de l'opérateur de transfert de Ruelle associé à ce système dynamique :

$$\mathcal{G}_s(f)(z) = \sum_{i=1}^r \phi_i(z)^s f \circ \phi_i(z)$$

où les  $(\phi_i)$  sont les fonctions continues définissant (par morceaux) la fonction  $\phi$  sur  $[0, 1]$ . Les fonctions de coût associées sont liées à des expressions de la forme

$$(5) \quad C_n(f) = \sum_{\alpha \in \Sigma^*} f(np_\alpha)$$

où  $\Sigma^*$  est l'ensemble des vecteurs de taille finie  $\alpha = (\alpha_i)$ ,  $\alpha_i \in \{0, 1\}$ ,

$$p_\alpha = \mathbb{P}(X_i^x = \alpha_i, \forall i)$$

est la probabilité du vecteur  $\alpha$  pour la suite stationnaire et  $f$  est une fonction telle que  $f(0) = 0$ . Le cas de la taille totale de l'arbre associé correspond à la fonction  $f(x) = (1 - (x + 1)e^{-x})$ .

L'approche qui a été suivie dans Robert [2] pour étudier le cas d'une suite stationnaire quelconque consiste aussi à "probabiliser" l'expression (5), i.e. à écrire  $C_n(f)$  comme l'espérance d'une variable aléatoire et à étudier le comportement asymptotique de celle-ci. Le théorème de renouvellement, mais cette fois pour les suites stationnaires, est aussi l'ingrédient-clé de la convergence. Le résultat utilisé est celui de Guivarc'h et Hardy [33] et il est intéressant de noter que la fonction importante pour ces questions est la fonction entropie  $h$  du système définie pour  $\omega = (\omega_i)$ , par

$$h(\omega) = -\log \mathbb{P}[X_0 = \omega_0 \mid X_i = \omega_i, i < 0].$$

Si celle-ci n'est pas co-homologue à une fonction latticielle alors, sous certaines conditions, on a

$$\lim_{n \rightarrow +\infty} \frac{1}{n} C_n(f) = \frac{1}{H} \int_0^{+\infty} \frac{f'(u)}{u} du,$$

où  $H = \mathbb{E}(h)$  est l'entropie du système.

## 2. Algorithmes de Contrôle de la Congestion

L'algorithme principale de contrôle de la congestion du protocole TCP (décrit dans l'annexe B) peut être représenté par la suite  $(W_n)$  telle que

$$W_{n+1} = \begin{cases} W_n + 1 & \text{si aucune perte parmi les } W_n \text{ paquets} \\ \max(\lfloor \delta W_n \rfloor, 1) & \text{sinon,} \end{cases}$$

où  $\delta \in (0, 1)$  et  $\lfloor x \rfloor$  désigne la partie entière de  $x$ . Si les pertes de paquets sont indépendantes les unes des autres, il est clair que  $(W_n)$  est une chaîne de Markov. Rappelons (voir l'appendice) que  $W_n$  est le nombre maximal de paquets que la source TCP peut avoir en circulation dans le réseau lors du  $n$ -ième cycle. On s'intéresse au comportement asymptotique de  $(W_n)$ . La probabilité invariante de cette chaîne de Markov n'est, à ce jour, pas connue, les sauts  $x \rightarrow \lfloor \delta x \rfloor$  compliquant singulièrement la résolution des équations d'équilibre. Si  $\alpha$  est la probabilité de perte d'un paquet, un moyen d'étudier cette chaîne de Markov consiste à étudier

le cas où  $\alpha$  tend vers 0. Un petit raisonnement élémentaire permet de comprendre comment renormaliser correctement cette chaîne de Markov avec le taux de perte.

**Heuristique.** Pour  $n \geq 1$ ,  $t_n^\alpha$  désigne l'indice du  $n$ -ième paquet perdu. L'hypothèse d'indépendance des pertes entraîne que la suite  $(\alpha t_n^\alpha)$  converge en distribution vers la suite des points d'un processus de Poisson de paramètre 1.

En supposant que  $W_0 = 1$ , s'il n'y a pas de perte jusqu'à transmission de la  $N$ -ième fenêtre, la croissance linéaire de l'algorithme AIMD entraîne que le nombre de paquets transmis à cette époque est de l'ordre de  $N^2/2$  (la surface d'un triangle). Le temps  $T$  nécessaire pour qu'un paquet soit perdu doit donc être tel que

$$(6) \quad \frac{T^2}{2} \stackrel{\text{dist.}}{=} t_2^\alpha - t_1^\alpha \sim E_1/\alpha \quad \text{d'où} \quad T \sim \frac{\sqrt{2E_1}}{\sqrt{\alpha}},$$

Le nombre de fenêtres est donc de l'ordre de  $1/\sqrt{\alpha}$ , l'inverse de la racine carrée du taux de perte. À cet instant, la taille de la fenêtre de congestion est bien entendu de cet ordre de grandeur. Cette discussion montre que le fameux phénomène de "l'inverse de la racine carrée du taux de perte" rencontré ici et là dans la littérature s'explique en fait très facilement de façon intuitive.

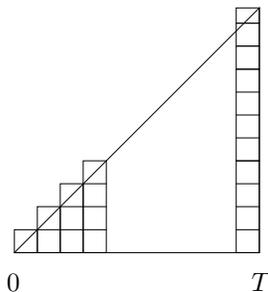


FIG. 2. Croissance de la fenêtre TCP

Il est donc naturel d'introduire le processus renormalisé  $(\sqrt{\alpha}W_{[t/\sqrt{\alpha}]})$ , son étude est détaillée dans Dumas *et al.* [10]. Ce processus converge en distribution vers un processus de Markov  $W(t)$  dont le générateur infinitésimal  $\Omega$  est donné par, si  $f$  est classe  $C^1$ ,

$$\Omega(f)(x) = f'(x) + x(f(\delta x) - f(x)), \quad x \geq 0.$$

La suite  $(V_n)$  des valeurs de  $(W(t))$  juste après les instants de sauts forme une chaîne de Markov. L'ingrédient qui simplifie l'étude consiste à remarquer que la suite des carrés  $(V_n^2)$  a même loi qu'une suite auto-régressive  $(X_n)$ , telle que

$$X_{n+1} = \delta X_n + 2E_n,$$

où  $(E_n)$  est une suite i.i.d. de variables exponentielles de paramètre 1. La loi à l'équilibre de  $(X_n)$  est donc donnée par la distribution de la variable

$$(7) \quad V = 2 \sum_{n \geq 1} \delta^n E_n.$$

Outre sa présence dans TCP, cette variable aléatoire joue un rôle dans l'étude asymptotique de plusieurs algorithmes : algorithmes de comptage rapide, Ethernet (voir Flajolet et Martin [28]), les algorithmes de parcours de points sur un

cercle (voir Litvak et van Zwet [42]), etc. L'étude de sa distribution ne pose pas de problème, voir Dumas *et al.* [10] ou Bertoin *et al.* [2]. De cette façon, il est alors possible d'obtenir la distribution invariante de  $(W(t))$  et donc de  $(\sqrt{\alpha}W_n)$  quand  $\alpha$  tend vers 0.

**Le cas des pertes non-indépendantes.** La propriété de Markov de la suite  $(W_n)$  est due au fait que les pertes de paquets sont indépendantes. En pratique ce n'est pas le cas, une perte de paquets entraîne avec forte probabilité d'autres pertes. En faisant l'hypothèse que les pertes surviennent par groupes et que globalement les pertes sont rares, il est possible de montrer la convergence en loi du processus  $(\sqrt{\alpha}W_{\lfloor t/\sqrt{\alpha} \rfloor})$ . Voir Guillemin *et al.* [7]. Le processus auto-régressif correspondant est défini par

$$X_{n+1} = \delta^{Z_n} X_n + 2E_n,$$

si  $Z_n$  est le nombre de pertes dans le  $n$ -ième groupe. Sa probabilité est définie par

$$(8) \quad V_Z = 2 \sum_{n \geq 0} \delta^{S_n} E_n,$$

si  $S_n = Z_1 + Z_2 + \dots + Z_n$ . Les moments fractionnaires de  $V$  qui jouent un rôle pour déterminer le débit asymptotique de l'algorithme ont une expression explicite. La distribution de  $V$  n'est, elle, en général pas connue. Quand la distribution de  $Z_1$  a une fonction génératrice rationnelle, il est possible (Guillemin *et al.* [7]) d'exprimer la densité de  $V$  avec des fonctions de type hypergéométriques (en fait dans le cadre du  $q$ -calcul).

**Lien avec les fonctionnelles exponentielles.** La variable  $V_Z$  ci-dessus peut aussi s'exprimer comme

$$V_Z = 2 \int_0^{+\infty} \delta^{\xi(s)} ds,$$

où  $\xi(s) = \sum_{i=1}^{N_s} Z_i$  et  $(N_s)$  est un processus de Poisson de paramètre 1 (où  $E_n$  est la distance entre le  $n$ -ième et le  $(n+1)$ -ième point de ce processus ponctuel). Quand  $(\xi(t))$  est un mouvement brownien avec drift négatif (et même plus généralement un processus de Lévy), ces variables jouent un grand rôle en mathématiques financières. Voir le livre de Yor [61] à ce sujet. Dans le cas de TCP,  $(\xi(t))$  est un processus de Poisson composé. Le cas où  $(\xi(t))$  est un processus de Poisson (qui correspond à la variable  $V$  de l'équation (7)) a aussi été étudié par Bertoin *et al.* [2, 4].



## Publications

### 1. Preprints

1. Christian Gromoll, Philippe Robert, and Bert Zwart. Fluid limits for processor sharing queues with impatience, April 2006. To appear in "Mathematics of Operation Research".
2. Philippe Robert. A probabilistic analysis of tree algorithms (III) : The stationary case, 2006. In preparation.
3. Hanène Mohamed and Philippe Robert. A probabilistic analysis of tree algorithms (II) : The dynamic case, 2006. In preparation.

### 2. Livres

1. Philippe Robert. *Stochastic Networks and Queues*. Stochastic Modelling and Applied Probability Series. Springer-Verlag, New York, 2003. xvii+398 pp.
2. Philippe Robert. *Réseaux de Communication, Algorithmes et Probabilités*. École Polytechnique, Centre de mathématiques appliquées, route de Saclay 91120 Palaiseau, September 2005. vi+190 pp.

### 3. Articles de Revues : Probabilités

1. Nelson Antunes, Christine Fricker, Philippe Robert, and Danielle Tibi. Stochastic networks with multiple stable points. *Annals of Probability*, 2007. To appear.
2. Nelson Antunes, Christine Fricker, Philippe Robert, and Danielle Tibi. Analysis of loss networks with routing. *Annals of Applied Probability*, 16(4) :2007–2026, 2006.
3. Nelson Antunes, Christine Fricker, Fabrice Guillemin, and Philippe Robert. Perturbation analysis of a variable  $M/M/1$  queue : A probabilistic approach. *Advances in Applied Probability*, 38(1) :263–283, 2006.
4. Hanène Mohamed and Philippe Robert. A probabilistic analysis of some tree algorithms. *Annals of Applied Probability*, 15(4) :2445–2471, November 2005.
5. Philippe Robert. On the asymptotic behavior of some algorithms. *Random Structures and Algorithms*, 27(2) :235–250, September 2005.
6. Fabrice Guillemin, Philippe Robert, and Bert Zwart. Tail asymptotics for processor sharing queues. *Advances in Applied Probability*, 36(2) :525–543, 2004.
7. Fabrice Guillemin, Philippe Robert, and Bert Zwart. AIMD algorithms and exponential functionals. *Annals of Applied Probability*, 14(1) :90–117, 2004.

8. Christine Fricker, Philippe Robert, and Danielle Tibi. A degenerate central limit theorem for single resource loss systems. *Annals of Applied Probability*, 13(2) :561–575, 2003.
9. Jean-François Dantzer and Philippe Robert. Fluid limits of string valued Markov processes. *Annals of Applied Probability*, 12(3) :860–889, 2002.
10. Vincent Dumas, Fabrice Guillemin, and Philippe Robert. A Markovian analysis of Additive-Increase Multiplicative-Decrease (AIMD) algorithms. *Advances in Applied Probability*, 34(1) :85–111, 2002.
11. Vincent Dumas and Philippe Robert. On the throughput of a resource sharing model. *Mathematics of Operation Research*, 26(1) :163–173, 2001.
12. Christine Fricker, Philippe Robert, and Danielle Tibi. On the rates of convergence of Erlang’s model. *Journal of Applied Probability*, 36(4) :1167–1184, 1999.
13. Christine Fricker, Philippe Robert, Ellen Saada, and Danielle Tibi. Analysis of some networks with interaction. *Annals of Applied Probability*, 4(4) :1112–1128, 1994.
14. Vadim Malyshev and Philippe Robert. Phase transition in a loss load sharing model. *Annals of Applied Probability*, 4(4) :1161–1176, 1994.
15. Claude Kipnis and Philippe Robert. A dynamic storage process. *Stochastic Processes and their Applications*, 34 :155–169, 1990.
16. Philippe Robert. Sur un processus de vie et de mort de particules sur  $[0,1]$ . *Annales de l’institut Henri Poincaré*, 23(2) :225–235, 1987.

#### 4. Articles de Revues : Modèles Stochastiques

1. H. Christian Gromoll, Philippe Robert, Bert Zwart, and Richard Bakker. The impact of renegeing in processor sharing queues. *Sigmetrics Performance Evaluation Review*, 34(1) :87–96, 2006.
2. Nelson Antunes, Christine Fricker, Fabrice Guillemin, and Philippe Robert. Integration of streaming services and TCP data transmission in the Internet. *Performance Evaluation*, 62(1-4) :263–277, October 2005.
3. Isi Mitrani and Philippe Robert. On the ASTA property in a feedback processor-sharing queue. *Performance Evaluation*, 58(1) :81–85, October 2004.
4. Jean-François Dantzer, Isi Mitrani, and Philippe Robert. Large scale and heavy traffic asymptotics for systems with unreliable servers. *Queueing Systems, Theory and Applications*, 38(1) :5–24, 2001.
5. Jean-François Dantzer, Mostapha Haddani, and Philippe Robert. On the stability of a bandwidth packing algorithm. *Probability in the Engineering and Informational Sciences*, 14(1) :57–79, 2000.
6. Samuel Kotz, Hosam Mahmoud, and Philippe Robert. On generalized Pólya urn models. *Statistics & Probability Letters*, 49(2) :163–173, 2000.
7. Edward Coffman, Philippe Robert, and Alexander Stolyar. The interval packing process of linear networks. *Performance Evaluation*, 27(3), 1999.
8. Philippe Robert. Space-time properties of some storage processes. *Queueing Systems, Theory and Applications*, 29(2-4) :189–192, 1998.

9. E. G. Coffman, Jr., E. N. Gilbert, Albert G. Greenberg, F. T. Leighton, Philippe Robert, and Alexander L. Stolyar. Queues served by a rotating ring. *Communications in Statistics Stochastic Models*, 11(3) :371–394, 1995.
10. Alain Jean-Marie and Philippe Robert. On the transient behavior of the processor-sharing queue. *Queueing Systems, Theory and Applications*, 17 :129–136, 1994.
11. E. G. Coffman, Jr., Shlomo Halfin, Alain Jean-Marie, and Philippe Robert. Stochastic analysis of a slotted FIFO communication channel. *IEEE Transactions on Information Theory*, 39(5) :1555–1567, September 1993.
12. E. G. Coffman, Guy Fayolle, Philippe Jacquet, and Philippe Robert. Largest-first sequential selection with a sum constraint. *Operations Research Letters*, 9 :141–146, May 1990.
13. Philippe Robert and Rolf Schassberger. On the M/G/1 foreground-background processor-sharing policy. *Queueing Systems, Theory and Applications*, 4 :281–286, 1989.
14. Philippe Robert, Isi Mitrani, and Peter J. B. King. An intermittently served discrete time queue with applications to meteor scatter communications. *Queueing Systems, Theory and Applications*, 3 :25–40, 1988.

### 5. Conférences avec comité de lecture

1. Yousra Chabchoub, Christine Fricker, Fabrice Guillemin, and Philippe Robert. Deterministic versus probabilistic packet sampling in the Internet. In *Proceedings of ITC'20*, June 2007.
2. Nelson Antunes, Christine Fricker, Philippe Robert, and Danielle Tibi. Metastability of CDMA cellular systems. In *Mobicom'2006*, Los Angeles, September 2006. ACM.
3. Christian Gromoll, Philippe Robert, Bert Zwart, and Richard Bakker. The impact of renegeing in processor sharing queues. In *ACM-Sigmetrics*, Saint Malo, June 2006. ACM/IFIP WG 7.3.
4. Nelson Antunes, Christine Fricker, Fabrice Guillemin, and Philippe Robert. Perturbation analysis of the area swept under the the queue length process of a variable M/M/1 queue. In *Performance'05*, Juan les Pins, 2005. IFP WG 7.3.
5. Youssef Azzana, Fabrice Guillemin, and Philippe Robert. A stochastic model for topology discovery of tree networks. In *Proceedings of ITC'19*, Beijing, 2005.
6. Nadia ben Azzouna, Fabrice Guillemin, Stéphanie Poisson, Philippe Robert, Christine Fricker, and Nelson Antunes. Inverting sampled ADSL traffic. In *Proceedings of IEEE ICC'05*, Seoul, 2005.
7. Vincent Dumas, Fabrice Guillemin, and Philippe Robert. Admission control of leaky bucket regulated sources in a queueing system with priority. In *ITC'18*, Berlin, September 2003.
8. Jacqueline Boyer, Fabrice Guillemin, Philippe Robert, and Bert Zwart. Heavy tailed M/G/1-PS queues with impatience and admission control in packet networks. In *Infocomm'2003*, San Francisco, USA, 2003.

9. Fabrice Guillemin, Philippe Robert, and Bert Zwart. Performance of TCP in the presence of correlated packet loss. In *15th ITC Specialist Seminar on Internet Traffic Engineering and Traffic Management*, Wurzburg, July 2002.
10. Vincent Dumas, Fabrice Guillemin, and Philippe Robert. Limit results for Markovian models of TCP. In *Globecom'2001*, San Antonio, November 2001.
11. Alain Dupuis, Fabrice Guillemin, and Philippe Robert. Modeling max-min fairness for elastic flows in telecommunication networks. In *ITC'17*, Salvador da Bahia, September 2001.
12. Cédric Adjih, Philippe Jacquet, and Philippe Robert. Differentiated admission control in large networks. In *Infocomm'2000*, pages 1455–1461, Tel Aviv, 2000.
13. Christine Fricker, Philippe Robert, Ellen Saada, and Danielle Tibi. Analysis of a network model. In *Cellular automata and cooperative systems (Les Houches, 1992)*, pages 159–172. Kluwer Acad. Publ., Dordrecht, 1993.
14. C. Fricker and P. Robert. Memory reference model for the cache memories analysis. In *Performance'90 : International Symposium on Computer Performance*, Edinburgh, September 1990.
15. E. G. Coffman, S. Halfin, Alain Jean-Marie, and Philippe Robert. Stochastic analysis of a slotted fifo communication channel. In *4<sup>th</sup> International Conference on Data Communication Systems and their Performances*, Barcelone, June 1990.
16. E. G. Coffman, S. Halfin, Alain Jean-Marie, and Philippe Robert. Analysis of a slotted communication channel with next-fit message assignment. In *ORSA-TIMS conference*, Osaka, July 1989.
17. Philippe Robert. A queueing analysis of a go-back-N protocol. In *Int. seminar on Performance of distributed and Parallel systems*, Kyoto, December 1988. North Holland.
18. Philippe Robert, Isi Mitrani, and Peter J. B. King. Analysis of a meteor scatter communication protocol. In P.J. Courtois, editor, *Performance '87 : International Symposium on Computer Performance*, Bruxelles, 1987. North Holland.
19. C. Kipnis and Philippe Robert. On the maximal throughput of a storage process. In *2nd Symp. on Computer and Information Sciences*, Istanbul, 1987. Nova Science.
20. F. Baccelli and Philippe Robert. Analysis of update response times in a distributed data base. In S. K. Tripathi, editor, *Performance '83 : International Symposium on Computer Performance*, Washington, 1983. ACM, North Holland.

### 6. Contributions au groupe de normalisation IEEE 802.14

1. Philippe Jacquet, Paul Mühlethaler, and Philippe Robert. CATV slotted multiple access MAC. In *IEEE plenary 802 meeting, IEEE 802.14*, Montreal, November 1995.
2. Philippe Jacquet, Paul Mühlethaler, and Philippe Robert. Using QoS parameters to support various traffics on CATV multiple access MAC. In *IEEE plenary 802 meeting, IEEE 802.14*, Montreal, November 1995.

3. Philippe Jacquet, Paul Mühlethaler, and Philippe Robert. Performant implementations of tree collision resolution on CATV network. In *IEEE plenary 802 meeting, IEEE 802.14*, Los Angeles, May 1996.
4. Philippe Jacquet, Paul Mühlethaler, and Philippe Robert. A unified approach for CATV networks capable of mixing ATM and non ATM traffics with CBR and non CBR traffics. In *IEEE plenary 802 meeting, IEEE 802.14*, Twente, November 1996.
5. Philippe Jacquet, Paul Mühlethaler, and Philippe Robert. Simulation of a stack algorithm with priorities and reservations. In *IEEE plenary 802 meeting, IEEE 802.14*, Twente, November 1996.
6. Philippe Jacquet, Paul Mühlethaler, and Philippe Robert. Framing protocols on upstream channel in CATV networks : asymptotic average delay analysis. In *IEEE plenary 802 meeting, IEEE 802.14*, Los Angeles, May 1996.
7. Philippe Jacquet, Paul Mühlethaler, and Philippe Robert. Asymptotic average access delay analysis : adaptive p-persistence versus tree algorithm. In *IEEE plenary 802 meeting, IEEE 802.14*, Los Angeles, May 1996.
8. Philippe Jacquet, Paul Mühlethaler, and Philippe Robert. A slotted medium access control scheme designed for CATV network. Technical report, INRIA, INRIA-Rocquencourt, France, 2001.
9. Philippe Jacquet, Paul Mühlethaler, and Philippe Robert. Framing protocols on upstream channel in CATV networks : Asymptotic average delay analysis. Technical report, INRIA, INRIA-Rocquencourt, France, 2001.



## Références Générales

1. Jean Bertoin, *Homogeneous fragmentation processes*, Probability Theory and Related Fields **121** (2001), 301–318.
2. Jean Bertoin, Philippe Biane, and Marc Yor, *Poissonian exponential functionals,  $q$ -series,  $q$ -integrals, and the moment problem for log-normal distributions*, Seminar on Stochastic Analysis, Random Fields and Applications IV, Progr. Probab., vol. 58, Birkhäuser, Basel, 2004, pp. 45–56.
3. Jean Bertoin and Servet Martínez, *Fragmentation energy*, Advances in Applied Probability **37** (2005), no. 2, 553–570.
4. Jean Bertoin and Marc Yor, *Exponential functionals of Lévy processes*, Probability Surveys **2** (2005), 191–212 (electronic).
5. Anton Bovier, *Markov processes and metastability*, Lecture Notes TUB, 2003.
6. ———, *Metastability and ageing in stochastic dynamics*, Dynamics and Randomness II (Santiago de Chile) (A. Maas, S. Martinez, and J. San Martin, eds.), Kluwer, Dordrecht, 2004, pp. 17–81.
7. Maury Bramson, *Instability of FIFO queueing networks*, Annals of Applied Probability **4** (1994), no. 2, 414–431.
8. ———, *Instability of FIFO queueing networks with quick service times*, Annals of Applied Probability **4** (1994), no. 3, 693–718.
9. ———, *Convergence to equilibria for fluid models of FIFO queueing networks*, Queueing Systems, Theory and Applications **22** (1996), no. 1-2, 5–45.
10. ———, *Convergence to equilibria for fluid models of head-of-the-line proportional processor sharing queueing networks*, Queueing Systems, Theory and Applications **23** (1996), no. 1-4, 1–26.
11. ———, *Stability of two families of queueing networks and a discussion of fluid limits*, Queueing Systems, Theory and Applications **28** (1998), no. 1-3, 7–31.
12. H. Chen and A. Mandelbaum, *Discrete flow networks : bottleneck analysis and fluid approximations*, Mathematics of Operation Research **16** (1991), no. 2, 408–446.
13. J. Clément, P. Flajolet, and B. Vallée, *Dynamical sources in information theory : a general analysis of trie structures*, Algorithmica **29** (2001), no. 1-2, 307–369, Average-case analysis of algorithms (Princeton, NJ, 1998).
14. J. G. Dai, *On positive Harris recurrence of multiclass queueing networks : a unified approach via fluid limit models*, Ann. Appl. Probab. **5** (1995), no. 1, 49–77. MR MR1325041 (96c :60113)

15. Donald A. Dawson, *Measure-valued Markov processes*, École d'Été de Probabilités de Saint-Flour XXI—1991, Lecture Notes in Math., vol. 1541, Springer, Berlin, 1993, pp. 1–260.
16. Bogdan Doytchinov, John Lehoczky, and Steven Shreve, *Real-time queues in heavy traffic with earliest-deadline-first queue discipline*, *Annals of Applied Probability* **11** (2001), no. 2, 332–378.
17. V. Dumas, *Approches fluides pour la stabilité et l'instabilité de réseaux de files d'attente stochastiques à plusieurs classes de clients*, Ph.D. thesis, École Polytechnique, December 1995.
18. Vincent Dumas, *Diverging paths in FIFO fluid networks*, *IEEE Transactions on Automatic Control* **44** (1999), no. 1, 191–194.
19. P. Dupuis and R.J. Williams, *Lyapounov functions for semimartingale reflecting Brownian motions*, *Annals of Applied Probability* **22** (1994), no. 2, 680–702.
20. E. B. Dynkin and M. B. Maljutov, *Random walk on groups with a finite number of generators*, *Doklady Akademii Nauk SSSR* **137** (1961), 1042–1045.
21. P. Erdős, A. Hildebrand, A. Odlyzko, P. Pudaite, and B. Reznick, *The asymptotic behavior of a family of sequences*, *Pacific Journal of Mathematics* **126** (1987), no. 2, 227–241.
22. Kenneth Falconer, *Techniques in fractal geometry*, John Wiley and Sons, 1997.
23. G. Fayolle, P. Flajolet, and M. Hofri, *On a functional equation arising in the analysis of a protocol for a multi-access broadcast channel*, *Advances in Applied Probability* **18** (1986), 441–472.
24. Y. Filonov, *A criterion for the ergodicity of discrete homogeneous Markov chains*, *Akademiya Nauk Ukrainskoi SSR. Institut Matematiki. Ukrainskii Matematicheskii Zhurnal* **41** (1989), no. 10, 1421–1422.
25. Philippe Flajolet, *Évaluation de protocoles de communication : aspects mathématiques*, *Journées de la Société Mathématique de France*, janvier 1988, <http://algo.inria.fr/flajolet/Publications/smf.ps.gz>, pp. 1–22.
26. ———, *The ubiquitous digital tree*, STACS 2006. Annual Symposium on Theoretical Aspects of Computer Science (Marseille) (B. Durand and W. Thomas, eds.), *Lecture Notes in Computer Science*, vol. 3884, February 2006, pp. 1–22.
27. Philippe Flajolet, Xavier Gourdon, and Philippe Dumas, *Mellin transforms and asymptotics : harmonic sums*, *Theoretical Computer Science* **144** (1995), no. 1-2, 3–58, Special volume on mathematical analysis of algorithms.
28. Philippe Flajolet and Nigel Martin, *Probabilistic counting for data base applications*, *Journal of Computer Systems* **31** (1985), 182–209.
29. M. I. Freidlin and A. D. Wentzell, *Random perturbations of dynamical systems*, second ed., Springer-Verlag, New York, 1998, Translated from the 1979 Russian original by Joseph Szücs.
30. A.S. Gaïrat, V.A. Malyshev, M. V. Men'shikov, and K.D. Pelikh, *Classification of Markov chains describing the evolution of random strings*, *Russian Mathematical Surveys* **50** (1995), no. 2, 237–255.
31. R. J. Gibbens, P. J. Hunt, and F. P. Kelly, *Bistability in communication networks*, *Disorder in physical systems*, Oxford Sci. Publ., Oxford Univ. Press, New York, 1990, pp. 113–127.

32. H. C. Gromoll, Amber Puha, and R. J. Williams, *The fluid limit of a heavily loaded processor sharing queue*, Annals of Applied Probability **12** (2002), no. 3, 797–859.
33. Y. Guivarc’h and J. Hardy, *Théorèmes limites pour une classe de chaînes de Markov et applications aux difféomorphismes d’Anosov*, Annales de l’Institut Henri Poincaré. Probabilités et Statistique **24** (1988), no. 1, 73–98.
34. B.M. Hambly and Michel L. Lapidus, *Random fractal strings : their Zeta functions, complex dimensions and spectral asymptotics*, Transactions of the AMS, To appear.
35. R. Z. Has’minskiĭ, *Stochastic stability of differential equations*, Sijthoff & Noordhoff, Alphen aan den Rijn, 1980, Translated from the Russian by D. Louvish.
36. James R. Jackson, *Networks of waiting lines*, Operations Research **5** (1957), 518–521.
37. J.P. Kahane and J. Peyrière, *Sur certaines martingales de Benoît Mandelbrot*, Advances in Mathematics **22** (1976), 131–145.
38. F.P. Kelly, *Blocking probabilities in large circuit-switched networks*, Advances in Applied Probability **18** (1986), 473–505.
39. Donald E. Knuth, *The art of computer programming. Volume 3*, Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1973, Sorting and searching, Addison-Wesley Series in Computer Science and Information Processing.
40. Steven P. Lalley, *Renewal theorems in symbolic dynamics, with applications to geodesic flows, non-Euclidean tessellations and their fractal limits*, Acta Mathematica **163** (1989), no. 1-2, 1–55.
41. ———, *Probabilistic methods in certain counting problems of ergodic theory*, Ergodic theory, symbolic dynamics, and hyperbolic spaces (Trieste, 1989), Oxford Sci. Publ., Oxford Univ. Press, New York, 1991, pp. 223–258.
42. Nelly Litvak and Willem R. van Zwet, *On the minimal travel time needed to collect  $n$  items on a circle*, Annals of Applied Probability **14** (2004), no. 2, 881–902.
43. Q. Liu, *On generalized multiplicative cascades*, Stochastic Processes and their Applications **86** (2000), 263–286.
44. S.H. Lu and P.R. Kumar, *Distributed scheduling based on due dates and buffer priorities*, IEEE Transactions on Automatic Control **36** (1991), 1406–1416.
45. Hosam M. Mahmoud, *Evolution of random search trees*, Wiley-Interscience Series in Discrete Mathematics and Optimization, John Wiley & Sons Inc., New York, 1992, A Wiley-Interscience Publication.
46. V. A. Malyshev, *Networks and dynamical systems*, Advances in Applied Probability **25** (1993), no. 1, 140–175.
47. Benoît Mandelbrot, *Multiplications aléatoires et distributions invariantes par moyennes pondérées*, Comptes Rendus de l’Académie des Sciences Série I **278** (1974), 289–292 and 355–358.
48. P. Mathys and P. Flajolet,  *$Q$ -ary collision resolution algorithms in random access systems with free or blocked channel access*, IEEE Transactions on Information Theory **31** (1985), 244–254, Special Issue on Random Access Communication.

49. R. Daniel Mauldin and S.C. Williams, *Random recursive constructions : Asymptotic geometric and topological properties*, Transactions of the AMS **295** (1986), no. 1, 325–346.
50. Sean P. Meyn, *Transience of multiclass queueing networks via fluid limit models*, Annals of Applied Probability **5** (1995), no. 4, 946–957.
51. Grégory Miermont, *Coalescence et fragmentation stochastique, arbres aléatoires et processus de Lévy*, Ph.D. thesis, Université de Paris VI, December 2003.
52. Enzo Olivieri and Maria Eulália Vares, *Large deviations and metastability*, Encyclopedia of Mathematics and its Applications, vol. 100, Cambridge University Press, Cambridge, 2005.
53. Jim Pitman, *Exchangeable and partially exchangeable random partitions*, Probability Theory and Related Fields **102** (1995), 145–158.
54. A. A. Pukhal'skiĭ and A. N. Rybko, *Nonergodicity of queueing networks when their fluid models are unstable*, Problemy Peredachi Informatsii **36** (2000), no. 1, 26–46.
55. A. N. Rybko and A. L. Stolyar, *On the ergodicity of random processes that describe the functioning of open queueing networks*, Problems on Information Transmission **28** (1992), no. 3, 3–26.
56. W. Richard Stevens, *TCP-IP illustrated vol. 1 : the protocols*, Professional Computing Series, Addison-Wesley, 1994.
57. A. L. Stolyar, *On the stability of multiclass queueing networks : a relaxed sufficient condition via limiting fluid processes*, Markov Processes and Related Fields **1** (1995), no. 4, 491–512.
58. Wojciech Szpankowski, *Average case analysis of algorithms on sequences*, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience, New York, 2001.
59. B. Vallée, *Dynamical sources in information theory : fundamental intervals and word prefixes*, Algorithmica **29** (2001), no. 1-2, 262–306, Average-case analysis of algorithms (Princeton, NJ, 1998).
60. Edward C. Waymire and S.C. Williams, *A general decomposition theory for random cascades*, Bulletin of the AMS **31** (1994), no. 2, 216–222.
61. Marc Yor, *Exponential functionals of Brownian motion and related processes*, Springer-Verlag, Berlin, 2001, With an introductory chapter by Hélyette Geman, Chapters 1, 3, 4, 8 translated from the French by Stephen S. Wilson.

## Protocoles d'Accès à un Canal de Communication

Ce chapitre donne un aperçu sur la problématique des protocoles d'accès et sur la façon dont l'algorithme en arbre peut être défini dans le cadre d'un système distribué. Voir Flajolet [25].

### 1. Algorithmique des Réseaux d'Accès

Les réseaux de communication considérés ici sont constitués d'émetteurs, appelés aussi sources ou stations. Les émetteurs se partagent un unique canal de communication. Les temps de propagation des signaux sur celui-ci entre les nœuds sont très courts. C'est le cadre d'un *réseau local* de communication, en anglais LAN (Local Area Network).

Chaque émetteur reçoit, de façon sporadique éventuellement, des messages à transmettre sur le canal en direction d'un autre émetteur. Le canal unique impose une contrainte pour la transmission des messages : Si au moins deux stations essaient de transmettre chacune leur message au même moment, les signaux se superposent sur le canal, chaque émetteur sait alors que son message n'a pas été transmis à cause de la collision.

Il n'y a pas, a priori, d'hypothèses sur la topologie ou sur le nombre d'émetteurs. Le réseau est ouvert, de nouvelles stations peuvent se rajouter, d'autres peuvent être retirées. De plus, le réseau peut être hétérogène, certaines stations n'ayant quasiment pas de messages à transmettre alors que d'autres ont des flots réguliers à émettre.

Cette absence de contrainte sur la configuration du réseau a un impact important sur la conception d'une solution au problème de la transmission des messages. Quand la topologie est connue, ou quand le nombre de stations est déterminé, les solutions algorithmiques peuvent utiliser ces informations pour améliorer l'efficacité des protocoles de communication. Ce cadre technique est assez courant, on peut citer

- *Les réseaux sans fil.* Ils se situent dans cette problématique : plusieurs ordinateurs mobiles se partagent localement une fréquence pour émettre des paquets vers d'autres mobiles ou vers le réseau Internet classique.
- *Le réseau câblé.* Ce réseau est initialement un réseau de diffusion. La tête de station transmet à tous les nœuds les canaux télévisés. La possibilité technique de pouvoir émettre en sens inverse (i.e. des nœuds vers la tête de station) a commencé à être utilisée au milieu des années 1990 pour développer un accès Internet à haut débit pour les particuliers. L'avantage de cette solution réside dans le débit d'un tel accès qui est supérieur à celui offert par le réseau téléphonique classique.
- *Les réseaux locaux.* Le réseau est constitué de stations de travail connectées à un câble qui les relie au réseau extérieur. Cette situation a été le cadre

classique d'application des protocoles d'accès pendant de nombreuses années. Les récents développements technologiques des réseaux locaux ont toutefois quasiment supprimé le problème de l'accès au canal dans ce cadre. Ils sont mentionnés principalement en raison de leur importance historique.

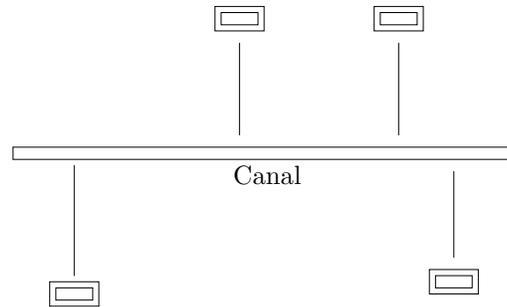


FIG. 1. Stations sur un canal de communication

**Les contraintes algorithmiques.** Les émetteurs ne peuvent se concerter pour l'accès au canal (en imaginant par exemple que les émetteurs soient dispersés sur les îles d'un atoll par exemple). Ils ne peuvent constituer une file d'attente, même virtuelle, pour résoudre les éventuels conflits d'accès. Un tel réseau est un *système distribué* : il ne peut pas y avoir de centre de contrôle régulant les transmissions. Ce chapitre est consacré aux algorithmes qui peuvent être utilisés par les émetteurs pour transmettre leurs messages. On parlera aussi de protocole de communication, c'est l'ensemble de toutes les règles qui régissent la transmission des messages. Un algorithme dans un tel cadre devra respecter les propriétés suivantes :

- (1) C'est un algorithme *distribué*. Les émetteurs utilisent le *même* algorithme.
- (2) Les émetteurs ont la possibilité d'écouter constamment le canal. La décision d'émission ou de non-émission est prise au début d'une unité de temps.
- (3) L'information d'un émetteur sur l'ensemble du réseau ne provient que du canal. Elle est *ternaire* : seuls trois événements sur le canal peuvent être distingués :
  - (B) un blanc : pas d'essai de transmission sur le canal.
  - (S) un succès : un seul émetteur transmet sur le canal.
  - (C) une collision : au moins deux émetteurs essaient une transmission.

En particulier un émetteur ne peut connaître le nombre d'émetteurs ayant des messages à transmettre. Il est clair que l'algorithme simpliste qui consiste à essayer la transmission dès que l'émetteur a un message à émettre est absurde. Le canal est alors définitivement bloqué quand deux émetteurs ont un message à transmettre.

### Cadre mathématique de l'étude

Plusieurs hypothèses simplificatrices sont faites. Elles consistent essentiellement à négliger certains aspects techniques, secondaires pour l'étude purement algorithmique. Par exemple, en pratique la détection d'une collision n'est pas toujours facile. Cela entraîne des contraintes sur les choix de fréquence, de taille du réseau, etc. De la même façon, la nomenclature (longueur, composition des entêtes,...) des

messages qui sont transmis est totalement ignorée alors qu'elle joue un rôle important dans la mise en œuvre effective des algorithmes. Avec ces simplifications, les problèmes algorithmiques liés à la transmission des messages sont plus faciles à envisager. En outre, les phénomènes principaux qui président aux choix algorithmiques sont conservés et mieux mis en évidence.

- **Le temps est discret.** Une horloge régule le canal de communication en unités de temps. Au début de chaque unité de temps, chaque émetteur ayant un message à transmettre détermine s'il essaie de transmettre.
- **Transmission.** La durée de transmission d'un message est d'une unité de temps.
- **Les arrivées.** Le modèle du nombre infini de sources. Pour  $n \geq 1$ , on note  $A_n$  le nombre de nouvelles demandes de transmission arrivées pendant la  $n$ -ième unité de temps. Les variables aléatoires  $A_n$ ,  $n \geq 1$  sont supposées indépendantes et identiquement distribuées.

En pratique, c'est le modèle du nombre fini de sources qui est pertinent : il y a  $p$  stations et chacune d'elles ne peut tenter d'émettre qu'un message à la fois. Chaque station a une file d'attente où sont stockés les messages en attente de transmission. Le modèle du nombre infini de sources a cependant le mérite d'être plus simple et de mieux mettre en évidence les faiblesses ou les qualités des algorithmes.

## 2. Le Protocole en Arbre

**Algorithmique.** Cet algorithme a été conçu, de façon indépendante, par Capetanakis au MIT (Massachusetts Institute of Technology) et Tsybakov et Mikhailov à l'Université de Moscou.

**DEFINITION 2** (L'algorithme en Arbre). *Un émetteur ayant un message à transmettre a un compteur  $C$ . Initialement,  $C = 0$ .*

*Au début de chaque unité de temps,*

- *Si  $C = 0$ , alors il y a essai de transmission.*
  - (1) *Si succès, c'est terminé.*
  - (2) *Si collision, tirage pile ou face :*
    - si pile,  $C = 0$ ,*
    - sinon,  $C = 1$ .*
- *Si  $C > 0$ , pas d'essai de transmission. L'émetteur écoute cependant le canal :*
  - (1) *Si succès ou pas d'essai de transmission,  $C \rightarrow C - 1$ .*
  - (2) *Si collision sur le canal,  $C \rightarrow C + 1$ .*

### Caractéristiques du protocole en arbre.

- Les seuls émetteurs ayant le droit de tenter une transmission sont ceux dont le compteur vaut 0. En cas de collision, ceux qui ont tenté d'émettre sont séparés, au hasard, en deux groupes en mettant la valeur du compteur d'une partie d'entre eux à 1. Ce mécanisme vise bien sûr à n'avoir, le plus souvent possible, qu'un élément avec le compteur 0.
- Pendant ses périodes d'inactivité, un émetteur continue à écouter le canal et module son compteur en fonction des événements se déroulant sur celui-ci.

En première approximation, plus la valeur du compteur d'un émetteur est élevée et plus (probablement) le nombre d'émetteurs en concurrence est élevé. Les collisions répétées indiquant un grand nombre de tentatives simultanées.

**Exemple.** On suppose qu'il y a initialement quatre messages A, B, C, D et pas de nouvelles arrivées.

Temps	1	2	3	4	5	6	7	8	9
État du Canal	Non	Non	$\emptyset$	Non	Ok	Ok	Non	Ok	Ok
Compteur Niveau 0	ABCD	AB		AB	A	B	CD	C	D
Niveau 1		CD	AB	CD	B	CD		D	
Niveau 2			CD		CD				

Une structure de données bien connue en informatique peut être utilisée pour représenter cet algorithme. Il s'agit de l'*arbre binaire*, c'est un arbre dont les sommets sont de degré égal à 2 ou sont isolés. À un sommet de cet arbre est associé un sous-ensemble  $S$  des messages.

- Si  $S$  est vide ou réduit à un point, le sommet n'a pas de descendants.
- Sinon, les messages tirant la valeur pile sont associés au nœud gauche, les autres au nœud droit.

La figure 2 représente l'exemple précédent de cette façon. Remarquer que le temps nécessaire pour transmettre ces quatre messages correspond au nombre de nœuds de l'arbre binaire associé.

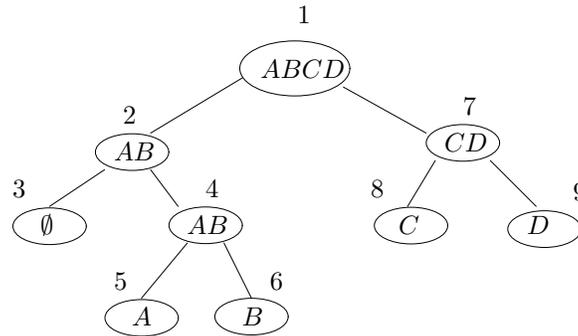


FIG. 2. Le protocole en arbre

**Une variante avec l'arbre ternaire.** En reprenant la représentation en arbre binaire, l'algorithme en arbre disperse les messages initialement en collision sur les feuilles de l'arbre binaire. Une transmission avec succès ayant lieu quand il n'y a qu'un message sur la feuille de l'arbre. Dans cette optique, il est naturel d'augmenter la capacité de ce schéma de dispersion en augmentant le degré de l'arbre.

DEFINITION 3 (L'algorithme avec un Arbre de Degré  $d \geq 2$ ).

Initialement le compteur d'un message est fixé à  $C = 0$ . Au début de chaque unité de temps,

- Si  $C = 0$ , alors il y a essai de transmission.
  - (1) Si succès, c'est terminé.
  - (2) Si collision, la valeur du compteur est obtenue après un tirage aléatoire sur  $\{0, 1, \dots, d\}$

- Si  $C > 0$ , pas d'essai de transmission.
- (1) Si succès ou pas d'essai de transmission,  $C \rightarrow C - 1$ .
- (2) Si collision sur le canal,  $C \rightarrow C + d - 1$ .

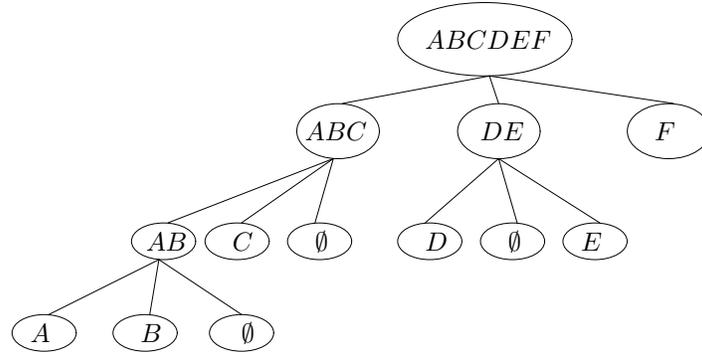


FIG. 3. Le protocole avec l'arbre ternaire

En représentant verticalement le vecteur d'état de l'algorithme, on obtient la configuration suivante pour l'arbre ci-dessus : Quand  $d = 3$  si  $\lambda < 0.40159$  alors la

Temps	1	2	3	4	5	6	7	8	9	10	11	12	13	14
État	6	3	2	1	1	0	1	0	2	1	0	1	1	0
	0	2	1	1	0	1	0	2	1	0	1	1	0	0
	0	1	0	0	1	0	2	1	0	1	1	0	0	0
	0	0	2	1	0	2	1	0	0	1	0	0	0	0
	0	0	1	0	2	1	0	0	0	0	0	0	0	0
	0	0	0	2	1	0	0	0	0	0	0	0	0	0
	0	0	0	1	0	0	0	0	0	0	0	0	0	0

TAB. 1. État de l'algorithme en arbre pour l'exemple de la figure 3

chaîne de Markov décrivant le protocole en arbre est ergodique. Il est facile de comprendre pourquoi un degré élevé pour l'arbre n'est pas souhaitable : si la dispersion s'accroît, le nombre de feuilles vides aussi et donc le nombre d'unités de temps inutilisées pour cause de silence.



## Le protocole de transmission de données TCP

Ce chapitre présente une classe d’algorithmes qui gèrent les phénomènes de congestion du réseau Internet. Ces algorithmes se sont imposés de façon quasi-universelle en raison de leur remarquable impact sur les performances globales des réseaux, et en raison aussi (et surtout), de la simplicité de leur conception.

### 1. Introduction

Le réseau de télécommunication considéré dans ce chapitre est constitué par

- des stations qui peuvent être à la fois *source* (émetteurs) ou *destination* (récepteur) de messages.
- des *routeurs* qui sont des nœuds assurant l’interconnexion entre les stations.

C’est le modèle du réseau Internet actuel. Dans ce réseau circulent des *paquets* qui vont d’un nœud du réseau, la source, à un autre nœud, la destination. Le paquet est l’unité élémentaire d’information qui transite dans le réseau. Pour transmettre un ensemble de données (un fichier) d’une source S à une destination D, la source S segmente le fichier en paquets qui sont envoyés successivement à D via le réseau. Cet ensemble de paquets est le *flot* associé à cette transmission de données. Un paquet est constitué d’un nombre borné de bits du type [ENT-INFO] avec un entête ENT où sont indiquées, entre autres, les adresses de la source et de la destination du paquet. Le reste du paquet, INFO, contient la partie données proprement dite.

#### Les phénomènes de congestion.

Les paquets sont envoyés de la source à un routeur qui détermine le prochain nœud du réseau vers lequel va transiter le paquet, et ainsi de suite jusqu’à la destination finale. Dans la figure 1, un paquet émis de la source X transite par trois routeurs pour arriver à la destination Z. Un routeur est un nœud ayant l’information sur la géographie locale du réseau. Il reçoit un flux de paquets qu’il oriente suivant les destinations de ceux-ci. Le routeur dispose d’une *mémoire tampon* (MT) où sont stockés les paquets qui sont en attente de traitement. Quand la mémoire tampon du routeur est pleine, les paquets arrivant au routeur sont perdus. Le réseau est alors saturé localement et tous les flots qui transitent à cet endroit ne sont plus correctement acheminés.

#### Problématique.

Un protocole de transmission de données est un ensemble de règles pour la transmission des messages utilisé par *toutes* les stations du réseau tel que :

- (1) Chaque source n’utilise que l’information *locale* au nœud pour transmettre les paquets. En particulier, il est impossible de connaître directement l’état de congestion des routeurs par lesquels vont transiter les paquets. Une fois

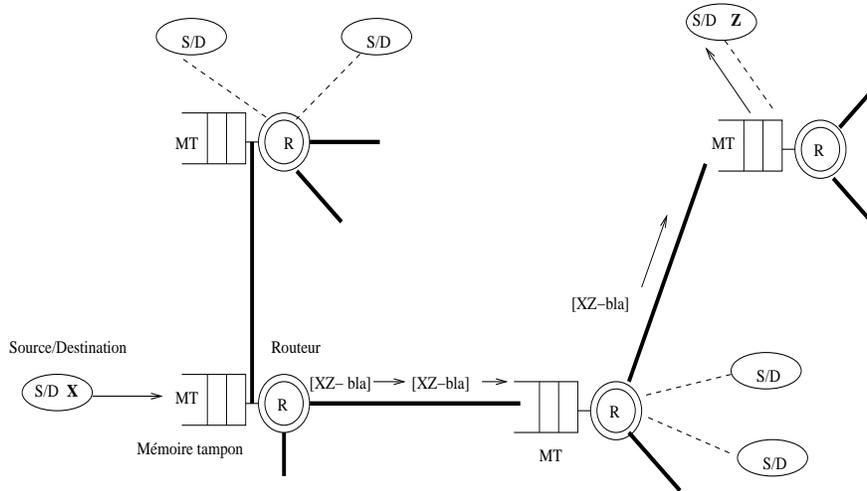


FIG. 1. Une architecture de réseau : routeurs et sources/destinations.

encore, c'est le cadre de l'*algorithmique distribuée* qui est pertinent. Il n'y a pas de contrôleur global du réseau Internet.

- (2) Tous les paquets transmis sont reçus avec succès par la destination. Par conséquent, des procédures pour détecter les pertes et pour réenvoyer des paquets perdus doivent faire partie du protocole.

Une propriété de robustesse connexe est sous-jacente : les paquets sont acheminés même si plusieurs routeurs sont définitivement en panne. Elle ne dépend toutefois pas du protocole lui-même mais de la façon dont la paquets sont routés dans le réseau. Cet aspect n'est pas envisagé dans ce chapitre.

Pour résumer, un protocole de transmission de données est un ensemble d'algorithmes distribués pour transmettre de façon *fiable* des données dans un réseau susceptible de perdre une partie des paquets qui y transitent.

## 2. Algorithmique de TCP

Le principe général du protocole est régi par des mécanismes d'acquittement, la destination fait savoir à la source quand elle a reçu un paquet donné. La source module ou accroît ses émissions suivant les informations reçues. La durée d'un aller et retour pour un paquet entre la destination et la source est notée RTT (Round Trip Time). La source sait donc au bout d'un temps RTT si un paquet envoyé a été reçu ou non. La source dispose d'une variable, la fenêtre de congestion  $W$ , qui donne le nombre maximum de paquets que la source peut envoyer sans s'occuper de la réponse de la destination.

**L'algorithme AIMD.** Si  $W = w$  la source envoie consécutivement  $w$  paquets au maximum sans attendre une réponse de la destination. Si tous les  $w$  paquets sont transmis avec succès, à l'issue de cette transmission, la variable  $W$  vaudra alors

$W = w + 1$ . De cette façon la source augmente progressivement, de façon linéaire, son débit tant que les paquets sont transmis avec succès.

Si, au contraire, la source détecte une perte de paquet, la variable  $W$  est alors mise à la valeur  $\lfloor W/2 \rfloor$  (où  $\lfloor x \rfloor$  désigne la partie entière de  $x \in \mathbb{R}$ ). Donc si  $W$  paquets sont transmis sans erreur, la valeur de la fenêtre de congestion à l'issue de ce cycle vaut  $W + 1$ , sinon elle vaut  $W/2$ . L'algorithme est dit à accroissements additifs et à décroissance multiplicative (AIMD en version anglaise : *additive increase, multiplicative decrease*). Dans le protocole TCP, l'algorithme est aussi appelé "Congestion Avoidance" et est schématisé ainsi

$$(9) \quad W \longrightarrow \begin{cases} W + 1 & \text{si aucune perte parmi les } W \text{ paquets} \\ \max(\lfloor W/2 \rfloor, 1) & \text{sinon.} \end{cases}$$

L'idée de base est très simple. La source n'augmente ses envois de paquets que de façon très progressive dans le réseau. En se rappelant que les connexions à un instant donné sont en compétition pour l'accès aux routeurs du réseau, une augmentation trop rapide de la taille de la fenêtre de chacune des connexions engendrerait plus rapidement des phénomènes de débordement des mémoires tampon et par conséquent impliquerait un plus grand nombre de connexions ayant des paquets perdus. À l'inverse, si une perte est détectée, les envois sont réduits de façon drastique en divisant par deux la taille de la fenêtre de congestion. S'il y a une perte, cela peut être dû au fait que la mémoire tampon d'un routeur est saturée par les multiples connexions TCP passant par ce routeur. Réduire la fenêtre de congestion de chacune de ces connexions de quelques unités ne décongestionnerait pas significativement ce goulet d'étranglement. Diviser par deux la taille de la fenêtre de congestion de chacune des connexions revient à "refroidir" brutalement l'activité du réseau dans ce secteur.

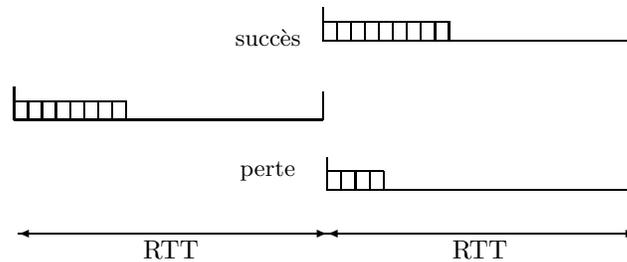


FIG. 2. La fenêtre de congestion pour TCP sur deux allers et retours

On peut ainsi donner une définition plus générale.

DEFINITION 4. *La taille  $W$  de la fenêtre de congestion d'un algorithme AIMD (Additive Increase, Multiplicative Decrease) de type  $(\beta, \delta)$  de transmission de paquets évolue de la façon suivante*

$$(10) \quad W \longrightarrow \begin{cases} W + \beta & \text{si pas de perte parmi les } W \text{ paquets transmis} \\ \max(\lfloor \delta W \rfloor, 1) & \text{sinon,} \end{cases}$$

où  $\beta \in \mathbb{N}$  et  $\delta \in ]0, 1[$ .

**2.1. Les autres algorithmes du protocole.** Les procédures supplémentaires ajoutées à l'algorithme AIMD qui vient d'être décrit visent à gérer au mieux les pertes supposées être détectées par la source. La source a deux moyens de détecter une perte :

- A. Par un moyen adéquat, elle estime initialement le temps qu'il faut pour acheminer un paquet et avoir la réponse de la destination. Avec cette donnée, une valeur maximum  $T_{out}$  (timeout) est établie. Si au bout de ce temps il n'y a aucune nouvelle du paquet, ce paquet est considéré comme définitivement perdu.
- B. Les paquets envoyés par la source sont numérotés. À chaque paquet reçu, la destination envoie à la source le numéro du prochain paquet attendu. Si la source reçoit plusieurs fois le même numéro  $N$ , cela implique que des paquets d'indices plus grand que  $N$  ont été reçus mais pas  $N$  lui-même. Ce type de situation peut arriver si un paquet est perdu à cause d'un phénomène local de congestion ou si certains paquets en dépassent d'autres en utilisant des chemins différents pendant leur transmission (ce qui est possible sur le réseau actuellement).

L'algorithme "slow start". Cet algorithme est utilisé dans le cas d'une perte définitive (cas A précédent). La taille de la fenêtre de congestion vaut à cet instant  $W$ . Une variable SSTHRESH ("slow start threshold") est initialisée à la valeur  $\lfloor W/2 \rfloor$  et la variable  $W$  est fixée 1. Ensuite, tant qu'il n'y a pas de perte, pour chaque paquet reçu correctement par la destination  $W$  est incrémenté de 1 tant que  $W$  ne dépasse pas la valeur SSTHRESH. Quand cette valeur est atteinte par  $W$ , l'algorithme AIMD décrit précédemment est ensuite utilisé (retour à la normale).

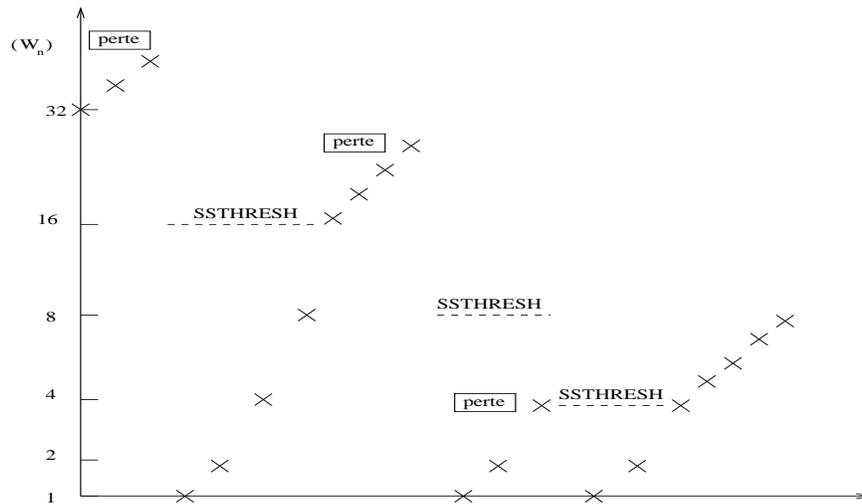


FIG. 3. L'algorithme "Slow start".

Remarquer que pour l'algorithme "slow start" tant que  $W$  est plus petit que SSTHRESH, à la fin d'un cycle de  $W$  paquets transmis avec succès la taille de la fenêtre de congestion vaut  $2W$ . Sur chaque cycle,  $W$  double jusqu'à rattraper la valeur SSTHRESH pour ensuite croître de 1 sur chaque cycle. Ce mécanisme peut

être vu comme un test rapide de la dégradation qui a conduit à la perte initiale dans le réseau.

*Les algorithmes “fast retransmit” et “fast recovery”.* Ces algorithmes sont utilisés dans le cas d’une perte “réparable” (cas B). Un paquet est égaré mais la durée critique de transmission  $T_{\text{out}}$  n’est pas écoulée. Dès que la source reçoit trois fois le même numéro de paquet, celui-ci est retransmis aussitôt (“fast retransmit”) et la variable  $W$  est changée en  $\lfloor W/2 \rfloor$ . Ensuite la source continue à envoyer les paquets comme si la situation était normale (“fast recovery”). Dans l’exemple de la table 1, le paquet 4 est perdu et la source s’en aperçoit à l’instant  $t = 9$ , elle le retransmet aussitôt. Puis la destination cesse d’envoyer le numéro 4 à l’instant où elle le reçoit pour envoyer le numéro du prochain paquet non reçu attendu (le numéro 9 dans ce cas).

Temps		1	2	3	4	5	6	7	8	9	10	11	12	13	14
Source	Émission	1	2	3	4	5	6	7	8	4	9	10	11	12	
	Réception					2	3	4	4	4	4	4	4	9	10
Destination	Réception			1	2	3	∅	5	6	7	8	4	9	10	11
	Émission			2	3	4	4	4	4	4	4	9	10	11	

TAB. 1. Les algorithmes “fast retransmit” et “fast recovery” pour la perte du paquet numéro 4, avec  $RTT = 4$

### Remarque.

La description de TCP qui vient d’être faite correspond à la version Reno du protocole (qui est, pour l’instant, la plus répandue). Tous les détails ne figurent pas dans la description qui vient d’être faite. De plus, il existe d’autres versions de ce protocole. Les principes de base de tous ces protocoles sont cependant contenus dans ces quatre algorithmes qui régissent les envois de paquets. Voir Stevens [56] pour un exposé plus détaillé.