

**Licence :** Ce texte a été écrit pour la préparation à l'agrégation externe de l'université de Bourgogne. Il est réutilisable librement, et adaptable selon les termes de la licence [Creative Commons by-nc-sa](#).

**Résumé :** Ce texte présente un modèle d'une partie du protocole informatique de transmission de données TCP. Sous des hypothèses simplificatrices, le modèle est une chaîne de Markov discrète que l'on peut étudier avec des outils classiques, en particulier des martingales. Dans la dernière partie, on observe le processus à grande<sup>1</sup> échelle et l'on décrit un processus continu limite.

**Mots clefs :** Chaînes de Markov, martingales, changement d'échelle

- *Il est rappelé que le jury n'exige pas une compréhension exhaustive du texte. Vous êtes laissé(e) libre d'organiser votre discussion comme vous l'entendez. Des suggestions de développement, largement indépendantes les unes des autres, vous sont proposées en fin de texte. Vous n'êtes pas tenu(e) de les suivre. Il vous est conseillé de mettre en lumière vos connaissances à partir du fil conducteur constitué par le texte. Le jury appréciera que la discussion soit accompagnée d'exemples traités sur ordinateur.*
- 

## Introduction

Ce texte propose une modélisation d'un aspect de la transmission des données sur Internet, le protocole TCP. Dans un cadre idéal, ce protocole se modélise par une chaîne de Markov ; on expliquera cette modélisation puis l'on essaiera de montrer quelques résultats sur ce modèle, *via* en particulier l'utilisation de martingales. Enfin, on verra comment le modèle discret, vu à une certaine échelle, peut s'apparenter à un modèle continu.

## 1 Le protocole TCP

On s'intéresse à un ordinateur A qui veut transmettre des données via internet à un ordinateur B. Ces données vont transiter par un certain nombre de nœuds du réseau (modem local, serveurs du fournisseur d'accès, etc.) avant de rejoindre B. Malheureusement, certains de ces nœuds peuvent être temporairement surchargés et peuvent refuser les connexions : les données sont alors perdues et il faut les ré-envoyer. Plus il y a de tentatives de connexion sur le réseau, plus ces pertes de données ont des chances de se produire.

L'ordinateur A n'a aucun moyen de connaître *a priori* le chemin que prendront ses données, ni la capacité (avant surcharge) des différents nœuds de ce chemin. Il doit donc essayer de trouver la bonne quantité de données à envoyer pour profiter au maximum des possibilités du réseau, sans pour autant le surcharger et provoquer des erreurs.

Une partie du protocole TCP (Transmission Control Protocol) fixe une stratégie pour savoir combien de données envoyer<sup>2</sup>. À quelques détails près, cette stratégie est la suivante. L'ordinateur garde en mémoire un entier, la « taille de fenêtre de transmission »  $X_n$ , qui représente le nombre

---

1. ou petite, suivant votre tolérance aux abus de langage

2. L'autre partie fixe la façon d'enrober les données brutes avec des indications pour qu'elles arrivent à bon port.

d'unités de données qu'il va envoyer au temps  $n$  (considéré comme un entier pour l'instant). Chaque unité de donnée peut être transmise avec succès, ou être perdue. Si toutes les unités sont transmises, l'ordinateur augmente légèrement son débit de transmission pour l'étape suivante :  $X_{n+1} = X_n + 1$ . Si il y a des pertes dans la transmission, l'ordinateur diminue fortement son débit en le divisant par deux :  $X_{n+1} = \lfloor X_n/2 \rfloor$  (où l'on note  $\lfloor x \rfloor$  la partie entière de  $x$ ).

Pour faire une étude simplifiée de l'efficacité de cette stratégie, on a besoin d'une modélisation des pertes. On fera ici l'hypothèse (peu réaliste) que chaque unité de donnée est perdue indépendamment des autres avec la même probabilité  $q \in ]0, 1[$ . On note  $p = 1 - q$  la probabilité que la transmission d'une unité de donnée réussisse : la probabilité de transmettre avec succès  $x$  unités de données est donc  $p^x$ .

Sous ces hypothèses, on a le résultat suivant.

**Théorème 1.** *La suite  $X_n$  est une chaîne de Markov à valeurs dans  $\mathbb{N}$ , de matrice de transition :*

$$\begin{cases} P(x, x+1) = p^x, \\ P(x, \lfloor x/2 \rfloor) = 1 - p^x, \\ P(x, y) = 0 \text{ si } y \notin \{x+1, \lfloor x/2 \rfloor\}. \end{cases}$$

*Elle est irréductible et apériodique.*

On cherche ensuite à savoir si la chaîne est récurrente, et si elle admet une mesure de probabilité stationnaire (ce qui correspond à la récurrence positive). Une idée possible est de chercher une fonction  $V$  positive, qui, en un certain sens, décroisse le long des trajectoires.

**Définition 1.** *On dit que  $V : \mathbb{N} \rightarrow \mathbb{R}_+$  est une fonction de Lyapunov si il existe  $c \geq 0$ , un entier  $k_0$  et une constante  $K < \infty$  tels que :*

$$\begin{cases} \forall x > k_0, & \mathbb{E}[V(X_1) \mid X_0 = x] \leq V(x) - c, \\ \forall x \leq k_0, & \mathbb{E}[V(X_1) \mid X_0 = x] \leq K. \end{cases} \quad (1)$$

On peut alors utiliser le théorème général :

**Théorème 2** (Critère de Foster Lyapunov). *Soit  $X_n$  est une chaîne de Markov sur  $\mathbb{N}$ , irréductible, et  $V$  une fonction de Lyapunov. On a les deux résultats suivants :*

- *Si  $V$  tend vers l'infini à l'infini, alors la chaîne est récurrente.*
- *Si la constante  $c$  est strictement positive, alors la chaîne est récurrente positive.*

Des éléments de preuve du théorème sont donnés dans la section suivante.

Dans notre cas, il est facile de vérifier que la fonction  $V(x) = x$  est une fonction de Lyapunov, avec une constante  $c > 0$ . La chaîne est donc récurrente positive, quelle que soit la valeur de  $p$ . On notera  $\mu_p$  la mesure de probabilité invariante.

La chaîne étant récurrente positive, elle finit par se comporter comme si elle était partie de sa loi invariante  $\mu_p$ . Certaines quantités intéressantes en pratique peuvent alors s'exprimer en fonction de  $\mu_p$ . On peut par exemple étudier le nombre moyen d'unités de données envoyées par unité de temps sur les  $n + 1$  premières étapes et montrer la convergence :

$$\frac{1}{n+1} \sum_{k=0}^n X_k \xrightarrow{n \rightarrow \infty} \sum_{k \in \mathbb{N}} k \mu_p(k). \quad (2)$$

## 2 Martingales et critère de Foster Lyapunov

Le critère de Foster–Lyapunov est classique et peut se montrer de nombreuses manières. Une approche possible est d'utiliser les martingales. On considère pour simplifier que  $k_0 = 0$  (on ne cherchera pas à déduire le cas général).

Intuitivement, l'équation (1) nous dit que, si l'on regarde l'évolution de  $V$  le long de la trajectoire  $X_n$ , la quantité  $V(X_n)$  a tendance à décroître, tant que l'on n'a pas atteint 0. Comme  $V$  est positive, elle ne peut pas décroître infiniment, et l'on doit donc atteindre rapidement 0.

Pour montrer la récurrence de la chaîne, il suffit de s'intéresser au temps de retour en 0 partant de 0. On introduit donc le temps  $\tau = \inf\{n \geq 1, X_n \leq k_0\} = \inf\{n \geq 1, X_n = 0\}$ . Ce temps aléatoire est un temps d'arrêt, à valeurs dans  $\mathbb{N} \cup \{\infty\}$ .

Plaçons-nous d'abord sous l'hypothèse que  $V$  tend vers l'infini à l'infini. L'objectif est de montrer, pour la chaîne partant de 0, que  $\tau$  est fini p.s. Pour deux entiers  $m, n$ , notons  $m \wedge n$  le minimum de  $m$  et  $n$ . Regardons, pour  $n \geq 1$ , la suite de variables aléatoires :

$$M_n = V(X_{n \wedge \tau}) : \omega \mapsto \begin{cases} V(X_n(\omega)) & \text{si } n \leq \tau(\omega) \\ V(X_{\tau(\omega)}(\omega)) & \text{si } n > \tau(\omega). \end{cases}$$

Dire que  $V$  « décroît le long des trajectoires » correspond à dire que  $M_n$  doit être une surmartingale (par rapport à la filtration naturelle  $\mathfrak{F}_n = \sigma(X_0, X_1, \dots, X_n)$ ). Pour le montrer, on peut établir successivement :

$$\begin{aligned} \mathbf{E}[M_{n+1} \mid \mathfrak{F}_n] - \mathbf{1}_{\tau \leq n} V(X_{n \wedge \tau}) &= \mathbf{1}_{\tau > n} \mathbf{E}[V(X_{n+1}) \mid \mathfrak{F}_n] \\ &= \mathbf{1}_{\tau > n} \sum_{k \geq 1} \mathbf{1}_{X_n=k} \mathbf{E}[V(X_1) \mid X_0 = k] \\ &\leq \mathbf{1}_{\tau > n} V(X_{n \wedge \tau}). \end{aligned}$$

On a donc bien :

$$\mathbf{E}[M_{n+1} \mid \mathfrak{F}_n] \leq V(X_{n \wedge \tau}) = M_n.$$

$M_n$  est une surmartingale positive, elle converge donc presque sûrement vers une variable  $M_\infty$  intégrable.

Plaçons-nous sur l'événement  $A = \{M_n \text{ converge vers } M_\infty < \infty\} \cap \{\tau = \infty\}$ . Sur  $A$ ,  $V(X_n) = M_n$  converge vers une limite finie. Comme  $V$  tend vers l'infini à l'infini, on en déduit que  $\sup_{n \in \mathbb{N}} X_n$  est fini (toujours sur  $A$ ). Comme la chaîne est irréductible,  $\mathbf{P}[\sup_n X_n < \infty] = 0$ . Par conséquent, la chaîne est récurrente.

Pour montrer la récurrence positive quand  $c$  est non-nul, on introduit une nouvelle quantité :

$$N_n = V(X_{n \wedge \tau}) + c(n \wedge \tau).$$

Là encore, on montre que  $N_n$  est une surmartingale positive. Elle converge donc p.s. Une nouvelle application du lemme de Fatou donne :

$$\mathbf{E}[N_\infty] \leq \mathbf{E}[N_1] = \mathbf{E}[V(X_1) + c] < \infty. \quad (3)$$

$N_\infty$  est donc finie p.s. Comme  $V$  est positive, ceci implique que  $\tau$  est fini p.s., donc

$$N_\infty = V(X_\tau) + c\tau = V(0) + c\tau.$$

En reprenant l'espérance, et en utilisant (3), on en déduit  $\mathbf{E}[\tau] < \infty$  et la chaîne est bien récurrente positive.

---

## 3 Limite continue

### 3.1 Renormalisation

Dans la pratique, le nombre de paquets envoyés est très grand et l'unité de temps pour la transmission des paquets très petite. On cherche alors à renormaliser le processus, c'est-à-dire qu'on le regarde à une nouvelle échelle. Supposons par exemple que l'ordinateur fasse des transmissions toutes les millisecondes (notées ms), et que l'unité de donnée soit l'octet (noté o). Ce qui nous intéresse est de tracer l'évolution du débit sur des intervalles de quelques secondes (s), ce débit étant *a priori* de l'ordre du kilo-octet<sup>3</sup> (ko). Notons  $\epsilon = 10^{-3}$  le rapport entre les différentes unités :  $1 \text{ ms} = 10^{-3} \text{ s}$ , et  $1 \text{ o} = 10^{-3} \text{ ko}$ .

On suppose que le taux de réussite d'une transmission d'un octet est un certain  $p_\epsilon$ , et on note  $X^\epsilon$  la chaîne de Markov en temps discret définie dans la première partie, avec le paramètre  $p_\epsilon$ .

Regardons maintenant ce qui se passe lorsque l'on change d'échelle. Si  $t$  est un temps (en seconde), appelons  $Y_t^\epsilon$  la taille de la fenêtre de transmission, en kilo-octets, au temps  $t$ .

Au temps  $t$  secondes =  $t/\epsilon$  ms, cette taille est de :

$$X_{t/\epsilon}^\epsilon \text{ octets} = \frac{1}{\epsilon} X_{t/\epsilon}^\epsilon \text{ kilo-octets}.$$

On pose donc :

$$Y_t^\epsilon = \frac{1}{\epsilon} X_{t/\epsilon}^\epsilon.$$

Cette définition n'est pour l'instant valable que pour des réels  $t$  tels que  $t/\epsilon \in \mathbb{N}$ . Pour des temps ne correspondant pas à des multiples exacts de la milliseconde, on peut tout de même poser :

$$Y_t^\epsilon = \frac{1}{\epsilon} X_{\lfloor t/\epsilon \rfloor}^\epsilon.$$

Il est naturellement possible de définir  $Y_t^\epsilon$  pour toutes les valeurs de  $\epsilon$  (et pas seulement  $\epsilon = 10^{-3}$ ). On dispose donc ainsi d'une famille de processus  $Y^\epsilon$ .

On suppose que la transmission d'un octet a de très fortes chances de réussir :  $p_\epsilon = 1 - \epsilon^2$ . On peut montrer dans ces conditions la proposition suivante :

**Proposition 1.** *Quand  $\epsilon$  tend vers 0, le processus  $Y^\epsilon$  « converge » vers un processus limite  $Y$ . En particulier, pour  $t$  fixé,  $Y_t^\epsilon$  converge en loi vers une variable aléatoire  $Y_t$ .*

On n'explicitera pas en quel sens la convergence du processus a lieu.

### 3.2 Construction explicite du processus limite

Le processus limite  $Y_t$  a une construction explicite. Pour la voir, supposons que l'on part de  $Y_0^\epsilon = x_0$ . Notons  $T^\epsilon$  le premier temps où le processus  $Y^\epsilon$  « descend » (fait une transition d'un état  $x$  vers l'état  $\lfloor x/2 \rfloor$ ).

**Théorème 3.** *Pour le processus partant de  $x_0$ , quand  $\epsilon$  tend vers zéro,  $T_\epsilon$  converge en loi vers une variable aléatoire  $T$  de fonction de répartition :*

$$F_{x_0}(t) = \mathbf{P}[T \leq t] = 1 - \exp\left(-\int_0^t (x_0 + s) ds\right) = 1 - \exp\left(-x_0 t - t^2/2\right),$$

et de « taux de défaillance »  $D_{x_0}(t) = \frac{F'_{x_0}(t)}{1 - F_{x_0}(t)} = x_0 + t$ .

---

3. On considère ici que le kilo-octet fait  $10^3$  octets, contrairement à l'usage, selon lequel  $1 \text{ ko} = 1024 \text{ o}$ .

*Démonstration.* On montre la convergence des fonctions de répartition. Soit  $t$  un réel positif, on étudie :

$$\mathbf{P} [T_\epsilon > t] .$$

Notons  $k_\epsilon = \lfloor t/\epsilon \rfloor$ . Alors l'événement  $\{T_\epsilon \geq t\}$  correspond au fait que la chaîne discrète partant de  $x_0/\epsilon$  va vers le haut pendant les  $k_\epsilon$  premières étapes.

Ceci arrive avec probabilité :

$$p_\epsilon^{\lfloor x_0/\epsilon \rfloor} p_\epsilon^{\lfloor x_0/\epsilon \rfloor + 1} \dots p_\epsilon^{\lfloor x_0/\epsilon \rfloor + k_\epsilon} .$$

Comme  $p_\epsilon = 1 - \epsilon^2$ , cette quantité tend vers  $\exp(-tx_0 - \frac{1}{2}t^2)$  quand  $\epsilon$  tend vers 0. On en déduit l'expression de la fonction de répartition et du taux de défaillance.  $\square$

Si on suppose alors que la propriété de Markov, valable pour les chaînes discrètes  $Y^\epsilon$ , se transmet au processus limite, on peut alors voir le processus limite de la façon suivante :

- On part de  $x_0$ ,
- on attend un temps aléatoire  $T$ , de fonction de répartition  $F_{x_0}$ , pendant lequel on augmente linéairement :  $Y_t = x_0 + t$  pour  $t < T$ ,
- au temps  $T$  on est divisé par deux :  $Y_T = \frac{x_0 + T}{2}$ ,
- on continue la construction en partant du nouveau point de départ, indépendamment de ce qui s'est passé auparavant.

## 4 Suggestions

- On pourra illustrer informatiquement l'évolution du processus discret.
- Comment peut-on simuler la variable aléatoire  $T$  de la dernière partie, et ainsi illustrer l'évolution du processus continu ?
- On pourra détailler la preuve du théorème 1, les preuves de la deuxième partie ou encore la construction de la limite continue.
- On pourra préciser des valeurs possibles de  $x_0$ ,  $c$  et  $K$  permettant de montrer que  $V : x \mapsto x$  est une fonction de Lyapunov. D'autres choix sont-ils possibles pour  $V$  ?
- Supposons que l'on observe une trajectoire du processus, partant de  $x$ , jusqu'au premier temps  $l$  où il y a une perte, et que l'on cherche à estimer  $p$ . On pourra montrer que la vraisemblance s'écrit :

$$V(p; x, l) = p^{(l-1)x + \frac{l(l-1)}{2}} (1 - p^{x+l})$$

et en déduire un estimateur du paramètre  $p$ .

- On pourra préciser sous quelles hypothèses et en quel sens la convergence a lieu dans l'équation (2), et l'illustrer informatiquement. Comment le débit moyen évolue-t-il en fonction du paramètre  $p$  ?
- La mesure invariante  $\mu_p$  est-elle réversible ?