



**HAL**  
open science

## Optimisation conjointe de la métrologie active et du routage : une approche markovienne

Maxime Mouchet, Sandrine Vaton, Olivier Brun, Pablo Belzarena, Isabel Amigo, Balakrishna Prabhu

### ► To cite this version:

Maxime Mouchet, Sandrine Vaton, Olivier Brun, Pablo Belzarena, Isabel Amigo, et al.. Optimisation conjointe de la métrologie active et du routage : une approche markovienne. Rencontres Francophones sur la Conception de Protocoles, l'Évaluation de Performance et l'Expérimentation des Réseaux de Communication, May 2018, Roscoff, France. hal-01785870

**HAL Id: hal-01785870**

**<https://hal.science/hal-01785870v1>**

Submitted on 4 May 2018

**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.

# Optimisation conjointe de la métrologie active et du routage : une approche markovienne

M. Mouchet<sup>1</sup>, S. Vaton<sup>1</sup>, O. Brun<sup>2</sup>, P. Belzarena<sup>3</sup>, I. Amigo<sup>1</sup> et B. Prabhu<sup>2</sup>

<sup>1</sup>IMT Atlantique, IRISA, UBL, Brest, France

<sup>2</sup>LAAS-CNRS, Université de Toulouse, CNRS, Toulouse, France

<sup>3</sup>IIE, FING, UdelaR, Montevideo, Uruguay

---

Nous nous intéressons à l'optimisation des stratégies de monitoring actif dans le but d'optimiser le routage dans un réseau overlay. Notre approche est d'exploiter la stabilité des différentes routes en caractérisant les délais par des chaînes de Markov cachées (HMM) et de formaliser le problème comme un processus de décision markovien (MDP). Nous introduisons un modèle HMM pour caractériser les séries de RTT dans un réseau comme Internet. Puis nous formalisons le problème d'optimisation conjointe de la métrologie et du routage comme un MDP que nous résolvons avec une méthode de type *Value Iteration*. Enfin nous validons notre approche dans différents contextes. En particulier nous considérons un overlay de 30 ancres du réseau de mesure public RIPE Atlas. Nous montrons qu'en exploitant correctement la stabilité des performances des routes, il est possible de réduire le budget de mesure d'au moins 90% en ne perdant quasiment rien en termes de performance de routage.

**Mots-clefs :** optimisation, routage, métrologie, RTT, HMM, MDP, RIPE Atlas

---

## 1 Introduction

Les protocoles de routage Internet souffrent de nombreux problèmes de disponibilité et de qualité des routes calculées. Mais l'ossification de l'Internet ne permet plus d'apporter de changements majeurs à son architecture. Une alternative consiste à utiliser un overlay de routage, réseau virtuel déployé entre des hôtes terminaux de l'Internet. Les nœuds du réseau peuvent mesurer la qualité des routes IP entre eux, par l'envoi de paquets sondes, et sont capables de coopérer pour acheminer les messages. Les solutions existantes utilisent l'approche *all-pair probing*, qui implique l'estimation de la qualité de tous les liens overlays à chaque époque de mesure. Si une telle approche garantit un routage optimal, elle ne passe pas à l'échelle, l'effort de mesure croissant en  $O(n^2)$  avec le nombre  $n$  de nœuds.

Dans cet article nous cherchons un compromis entre la latence des chemins utilisés pour le routage et l'effort de mesure nécessaire pour les découvrir. Une question similaire a été abordée dans [BWG16, BHV16]. Ces travaux supposent qu'un nombre donné de chemins de l'overlay doivent être mesurés à chaque époque de mesure, la seule question étant alors celle de la sélection de ces chemins. Au contraire, nous souhaitons développer un système réalisant un compromis optimal entre coût des mesures et latence des chemins utilisés pour le routage, sans contrainte sur le budget de monitoring.

Des travaux en tomographie réseau sur la sélection des chemins de bout-en-bout à mesurer pour minimiser le trafic de mesure tout en permettant d'évaluer tous les liens existent également [THP17]. En dépit de certaines similarités, ils diffèrent des nôtres dans la mesure où il est supposé que l'objectif est d'évaluer une métrique donnée sur tous les liens du réseau, indépendamment des décisions d'allocation de ressources.

Nous développons une approche markovienne pour l'optimisation conjointe de la métrologie active et du routage. Nous montrons, à partir de mesures réelles, comment il est possible de modéliser le délai sur chaque chemin par une chaîne de Markov cachée (HMM). Cette caractérisation des délais est ensuite utilisée pour formuler le problème d'optimisation conjointe de la métrologie et du routage comme un processus de décision markovien (MDP) qui peut être résolu numériquement. Les résultats numériques obtenus, en particulier pour un overlay de 30 ancrés du réseau RIPE, montrent qu'il est possible d'obtenir un routage quasi optimal malgré une réduction drastique de l'effort de mesure.

## 2 Modélisation statistique des RTTs

Notre objectif est de déterminer quand et sur quel chemin faire des mesures pour prendre des décisions de routage optimales. Un monitoring parcimonieux dans le temps n'est possible que si les performances des chemins sont suffisamment stables (ou au moins prévisibles) et si on est capable de prendre en compte cette stabilité dans la stratégie de monitoring.

Nous sommes partis de l'observation de traces réelles de RTT, sur NLNOG et sur RIPE Atlas. RIPE Atlas est une initiative de RIPE NCC (*Réseaux IP Européens*) pour mettre à disposition des opérateurs et des académiques une infrastructure de métrologie à l'échelle mondiale, sur laquelle il est possible de lancer des campagnes de mesure et de récupérer de très nombreux résultats (en accès public).

Nous proposons de caractériser les délais comme des chaînes de Markov cachées (Hidden Markov Model, HMM). Dans un modèle HMM les délais sont distribués selon un nombre fini de lois de probabilité, et la distribution change au bout d'un temps aléatoire (géométrique) en suivant une dynamique markovienne. On peut voir les modèles HMM comme une généralisation des modèles de mélange qui ont été proposés dans la littérature pour caractériser les séries de délais [FMF15]. Contrairement aux modèles de mélange, les HMM prennent en compte la dépendance temporelle des valeurs successives, ce qui est une propriété indispensable pour envisager un monitoring parcimonieux dans le temps. Par la suite on note  $L(t)$  le délai du chemin à l'instant discret  $t$  et  $S(t)$  une variable aléatoire qui représente l'état "caché" du système.  $S(t)$  est une chaîne de Markov à temps discret et  $L(t)$  dépend aléatoirement de  $S(t)$  seulement, par exemple selon une loi gaussienne.

L'estimation des paramètres du modèle peut être effectuée au sens du maximum de vraisemblance avec l'algorithme Expectation-Maximization (EM). Afin de déterminer l'ordre du modèle un critère de type AIC ou BIC peut être utilisé pour trouver un compromis entre la complexité d'un modèle et la vraisemblance des observations sous ce modèle. Alternativement des méthodes de Monte Carlo par chaîne de Markov (MCMC) peuvent être employées.

Sur la Figure 1 nous montrons la segmentation obtenue grâce à un modèle HMM sur une série de RTT entre deux ancrés RIPE Atlas pendant une semaine.

## 3 Optimisation du monitoring : une approche MDP

Considérons  $\mathcal{P}$  chemins entre une origine et une destination, et supposons que le délai  $L^i(t)$  du chemin  $i$  est une chaîne de Markov à  $K^i$  états  $\{l_1^i, \dots, l_{K^i}^i\}$ . A chaque instant on doit décider de mesurer ou non le délai sur le chemin  $i$  puis choisir de router le trafic par un des  $\mathcal{P}$  chemins. On note  $M^i(t) \in \{0, 1\}$  la décision de mesure sur le chemin  $i$  (1 signifiant "mesurer" et 0 "ne pas mesurer").

Le chemin choisi est celui dont le délai espéré est le plus faible. On note  $\gamma_{t,t}^{(i)}$  le filtre défini comme  $\gamma_{t,t}^{(i)}(j) = \mathbb{P}(L^i(t) = l_j^i | I_t^i)$  où  $I_t^i = \sigma(L^i(u), u \leq t / M^i(u) = 1)$ . Ce filtre représente la croyance sur l'état du délai au vu des mesures passées sur le chemin  $i$ . En raison de la propriété de Markov faible, la connaissance sur l'état du chemin  $i$  se résume à la connaissance du nombre de périodes de temps  $\tau^i$  écoulées depuis la

dernière mesure sur le chemin  $i$  et celle de la dernière valeur mesurée  $L_{\text{last}}^i$ . Le filtre peut être calculé comme  $\gamma_{t,t}^{(i)} = e_{L_{\text{last}}^i} (P^{(i)})^{\tau^i}$  où  $P^{(i)}$  est la matrice de transition de la chaîne de Markov, et  $e_{L_{\text{last}}^i}$  est un vecteur ligne dont les éléments sont tous 0 sauf celui en position  $L_{\text{last}}^i$  qui vaut 1. Quant à la valeur espérée du délai du chemin  $i$ , elle vaut  $E(L^i(t) | I_t) = \sum_j \gamma_{t,t}^{(i)}(j) l_j^i$ .

Le problème de sélection des chemins à mesurer (ou pas) peut être formulé comme un problème de décision markovien. On définit l'état de la chaîne de Markov contrôlée comme l'ensemble des états des chemins individuels, soit  $s = \{s^i\}_{i=\{1\dots\mathcal{P}\}} = \{(\tau^i, L_{\text{last}}^i)\}_{i=\{1\dots\mathcal{P}\}}$ . L'espace d'état résultant est discret, et également fini si  $\tau^i$  est limité à une valeur  $\tau_{\text{max}}^i$ . De même, l'action à un instant donné est définie comme l'ensemble des actions pour chaque chemin, soit  $M = \{M^i\}_{i=\{1\dots\mathcal{P}\}}$ . Suivant ces définitions,  $s$  est une chaîne de Markov à temps discret et à espace d'état discret, contrôlée par la décision de mesure  $M$ .

Les probabilités de transition de la chaîne de Markov contrôlée par la décision de mesure sont définies comme  $\mathbb{P}(s' | s, M) = \prod_i \mathbb{P}(s'_i | s_i, M^i)$ . Si  $M^i = 0$ , la transition  $s_i \rightarrow s'_i$  est de la forme  $(\tau^i, L_{\text{last}}^i) \rightarrow (\tau^i + 1, L_{\text{last}}^i)$  avec probabilité 1. Si  $M^i = 1$  la transition est de la forme  $(\tau^i, L_{\text{last}}^i) \rightarrow (0, l_j^i)$ ,  $j = 1 \dots K^i$  et les probabilités correspondantes sont  $e_{L_{\text{last}}^i} P_i^{\tau^i+1} e_j^T$  avec  $\bullet^T$  l'opérateur de transposition.

Enfin nous définissons la récompense instantanée du MDP comme l'opposé du délai minimum espéré sur les  $\mathcal{P}$  chemins, moins la somme des coûts de mesure :

$$R(s) = -D(s) - \sum_i c_i \mathbb{1}_{\tau_i=0} \quad \text{avec} \quad D(s) = \min_i \mathbb{E}(L^i(t) | I_t) = \min(\min_{i:\tau_i \geq 1} \mathbb{E}(L^i(t) | \tau^i, L_{\text{last}}^i), \min_{i:\tau_i=0} L_{\text{last}}^i)$$

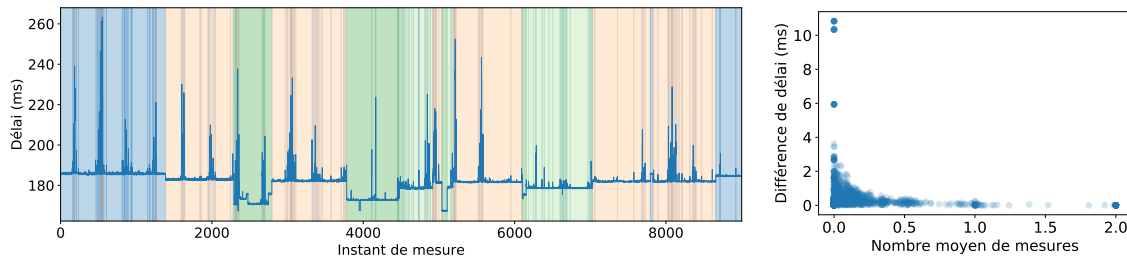
Soit  $\mathcal{S}$  l'ensemble des états possibles,  $\mathcal{M}$  l'ensemble des actions possibles, et  $\mu : \mathcal{S} \rightarrow \mathcal{M}$  la politique de décision. On définit la politique de mesure optimale  $M^*$  comme la politique qui maximise la récompense amortie à horizon infini (fonction de valeur)  $J_M(s) = \mathbb{E}_M(\sum_{t=1}^{\infty} \rho^t R(s_t) | s_0 = s)$  où  $0 < \rho < 1$  est un facteur de discount qui contrôle l'importance des récompenses futures et garantit la convergence de la série. Différentes méthodes existent pour trouver la politique  $M^*$  : algorithmes d'itération de politique (*Policy Iteration*) et des versions modifiées, d'itération de valeur (*Value Iteration*) [Kri16], apprentissage par renforcement (*Reinforcement Learning*). Nous avons utilisé l'algorithme *Value Iteration*.

Dans certains cas simples (deux chemins markoviens, ou un chemin markovien et un chemin déterministe) il est possible de trouver une forme analytique pour la politique optimale quand  $\rho = 0$  (c'est-à-dire qu'on ne tient compte que de la récompense immédiate). Ces solutions analytiques se ramènent à des politiques à seuil sur les valeurs des filtres  $\gamma$  et peuvent être interprétées comme la décision de mesurer quand l'incertitude sur l'état des chemins est élevée [SOM<sup>+</sup>18].

## 4 Résultats et conclusion

Nous avons simulé un overlay de 30 noeuds à partir de mesures RIPE Atlas. Pour chaque paire origine-destination nous avons conservé les deux chemins les plus courts la majorité du temps, le *chemin direct* et le *chemin alternatif*. Choisir dynamiquement le chemin de plus court délai permet de réduire le délai moyen jusqu'à 15% par rapport à l'utilisation du seul chemin direct ou alternatif.

La politique MDP avec un facteur de discount de 0.9 a été apprise, pour chaque paire, et pour différentes valeurs de  $c$ . Nous avons ensuite simulé le routage en appliquant la politique MDP, et une politique de mesure constante. Le nombre de mesures réalisées et le délai ont été enregistrés à chaque instant. Lorsque  $c \rightarrow 0$ , alors le nombre moyen de mesures par pas de temps tend vers 2, ce qui équivaut à mesurer les deux chemins en permanence. La différence de délai avec une politique de mesure constante en fonction du nombre de mesures observées est mise en évidence dans la figure 1. Chaque point représente une paire origine-destination. On remarque qu'un nombre très faible de mesures permet d'obtenir un délai proche d'une politique de mesure constante. Cela s'explique par la stabilité des routes pour lesquelles la probabilité



**FIGURE 1** : Gauche : Segmentation d’une série de RTT entre 2 ancres RIPE Atlas. Les couleurs identiques représentent des états identiques d’une HMM. Droite : Délai supplémentaire par rapport à la politique *toujours mesurer* en fonction du nombre moyen de mesures par pas de temps.

de changer d’état est faible, ce qui permet de prédire avec peu d’erreur le délai moyen pendant plusieurs pas de temps. Pour un coût de mesure fixé à 1, nous avons observé un nombre moyen de mesure de 0.17 (ce qui équivaut à un lien mesuré tout les 6 pas de temps) pour un délai supplémentaire par rapport à une politique de mesure constante de seulement 0.8 ms (0.26%). Cela représente une réduction de 91% du nombre de mesures par rapport à une politique de mesure constante.

La formalisation du problème de minimisation conjointe du délai et du coût de monitoring sous forme de MDP, rendue possible par la modélisation markovienne du délai, permet de trouver des politiques de mesure légères dont les performances sont vérifiées sur des données réelles. Un problème de cette approche est la complexité exponentielle de l’espace d’état avec le nombre de chemins, ce qui rend difficile l’obtention de la politique avec plus de deux chemins dans une paire origine-destination. Nos travaux futurs s’orientent vers la recherche de méthodes de résolution plus efficaces. Nous nous intéressons également à la mise en oeuvre de notre approche dans un contexte SDN et à l’identification de nouveaux cas d’usage du formalisme de métrologie parcimonieuse.

**Remerciements** : Ce travail a reçu un soutien financier du programme STIC AmSud dans le cadre du projet PROVE (*optimal Path discovery in Routing OVERlays*, 2016-2017).

## Références

- [BHV16] O. Brun, H. Hassan, and J. Vallet. Scalable, self-healing, and self-optimizing routing overlays. In *IFIP Networking 2016*, Vienna, Austria, May 17-19 2016.
- [BWG16] O. Brun, L. Wang, and E. Gelenbe. Big data for autonomic intercontinental overlays. *IEEE Jour. Selected Areas in Communications (special Issue on Emerging Technologies in Communications - Big data)*, 34 :575–584, 2016.
- [FMF15] R. Fontugne, J. Mazel, and K. Fukuda. An empirical mixture model for large-scale RTT measurements. In *IEEE Conference on Computer Communications (INFOCOM)*, 2015.
- [Kri16] V. Krishnamurthy. *Partially Observed Markov Decision Processes : From Filtering to Controlled Sensing*. Cambridge University Press, 2016.
- [SOM<sup>+</sup>18] S.Vaton, O.Brun, M.Mouchet, P.Belzarena, I.Amigo, B.Prabhu, and T.Chonavel. Joint minimization of monitoring cost and delay in overlay networks : optimal policies with a markovian approach. *Journal of Networks and System Management*, under revision–2018.
- [THP17] D. Z. Tootaghaj, T. He, and T. La Porta. Parsimonious tomography : Optimizing cost-identifiability trade-off for probing-based network monitoring. In *IFIP Performance 2017*, 2017.