



HAL
open science

Routage et apprentissage par renforcement

Alexis Bitailleu, Benoît Parrein, Guillaume Andrieux

► **To cite this version:**

Alexis Bitailleu, Benoît Parrein, Guillaume Andrieux. Routage et apprentissage par renforcement. Journées non thématiques GDR-RSD 2020, Jan 2020, Nantes, France. hal-02536945

HAL Id: hal-02536945

<https://hal.science/hal-02536945v1>

Submitted on 8 Apr 2020

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.

Routage et apprentissage par renforcement

Alexis Bitailou¹, Benoît Parrein¹, Guillaume Andrieux²

¹ Université de Nantes, LS2N (UMR 6004), France

² Université de Nantes, IETR (UMR 6164), France

Introduction et état de l'art

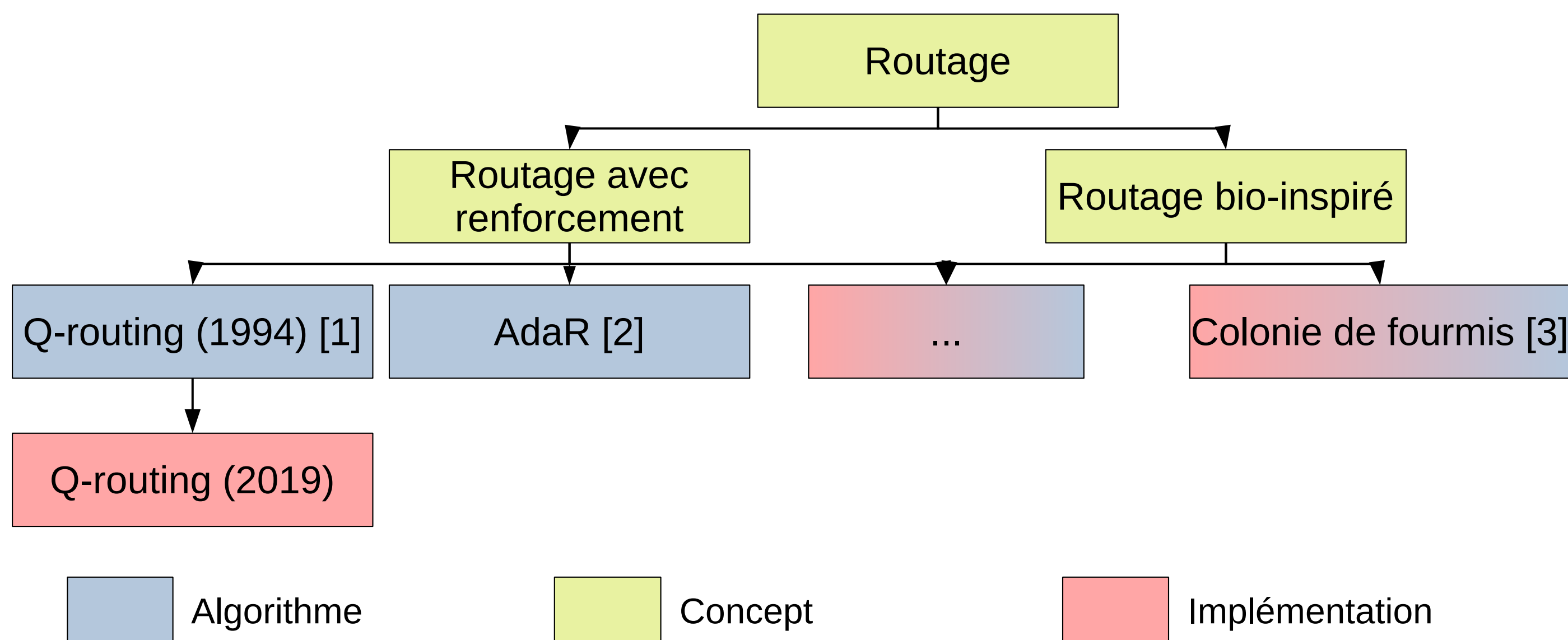


Figure 1 : les nouvelles approches du routage dans les années 90.

Q-learning est un algorithme d'apprentissage par renforcement créé par Watkins et Dayan [4]. Pour prendre une décision, ils définissent la fonction Q qui associe un état (S), une action (A) et une récompense (R) telle que :

$$Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \alpha(R_{t+1} + \gamma \max_a Q(S_{t+1}, a))$$

Q-learning + Routage = *Q-routing*

En 1994, Boyan et Littman [1] intègrent *Q-learning* à leur algorithme de routage et adaptent la fonction Q afin de minimiser la latence :

$$Q_x(d, y) \leftarrow Q_x(d, y) + \eta(q + s + \min_{z \in \text{neighbour of } y} Q_y(d, z) - Q_x(d, y))$$

D'après leurs simulations :

À charge faible : *Q-routing* < plus court chemin (Bellman-Ford)

À charge élevée : *Q-routing* >> plus court chemin (Bellman-Ford)

Paramètres de simulation

- *Q-routing* vs. Bellman-Ford sur Qualnet 8.2
- Lien filaire 10 Mb/s, délai de propagation fixe
- Flux à débit constant (CBR)

Le scénario de test :

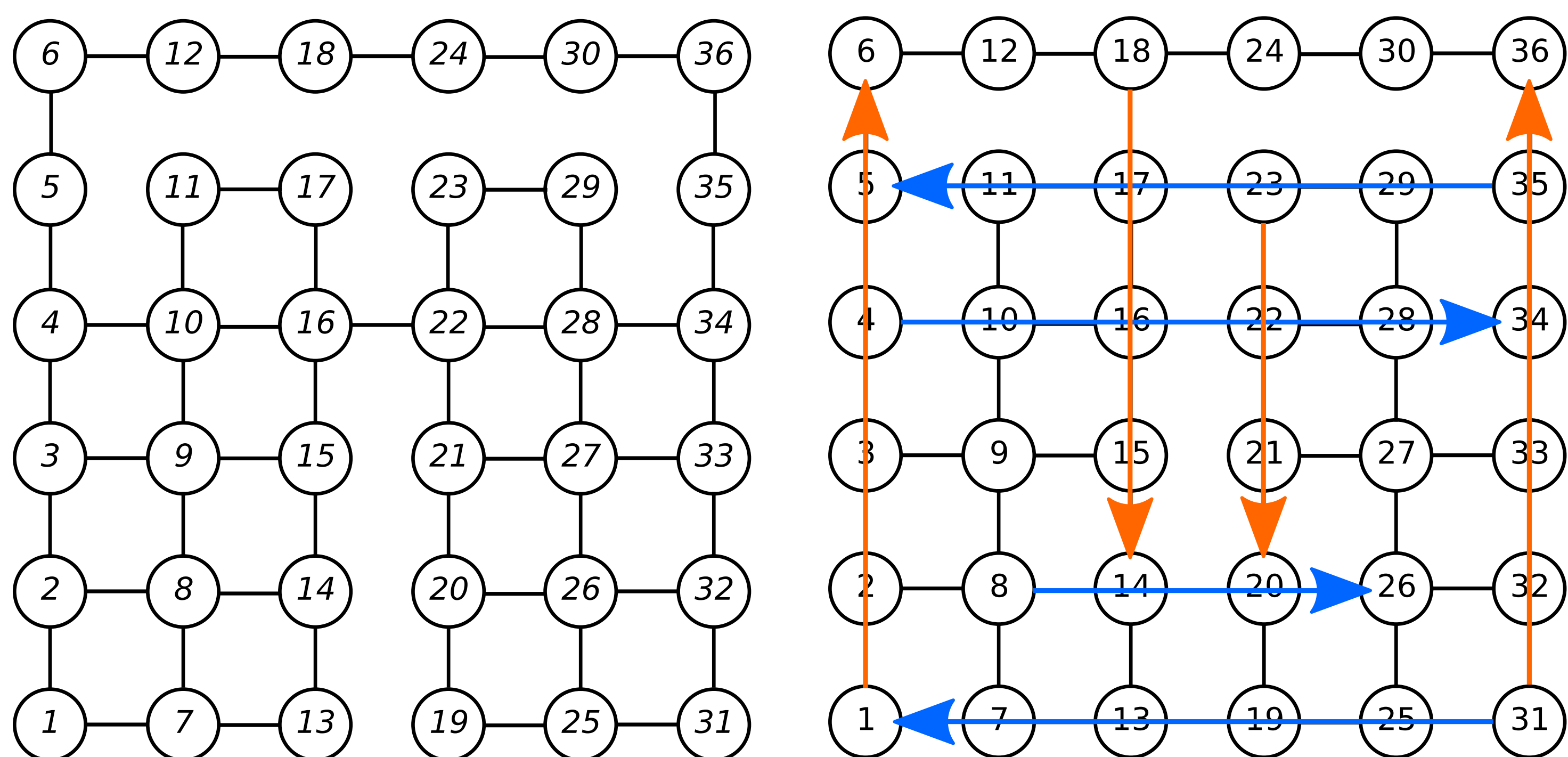


Figure 2 : Grille irrégulière de Boyan et Littman [1]

Figure 3 : Placement des flux CBR sur la grille

Résultats

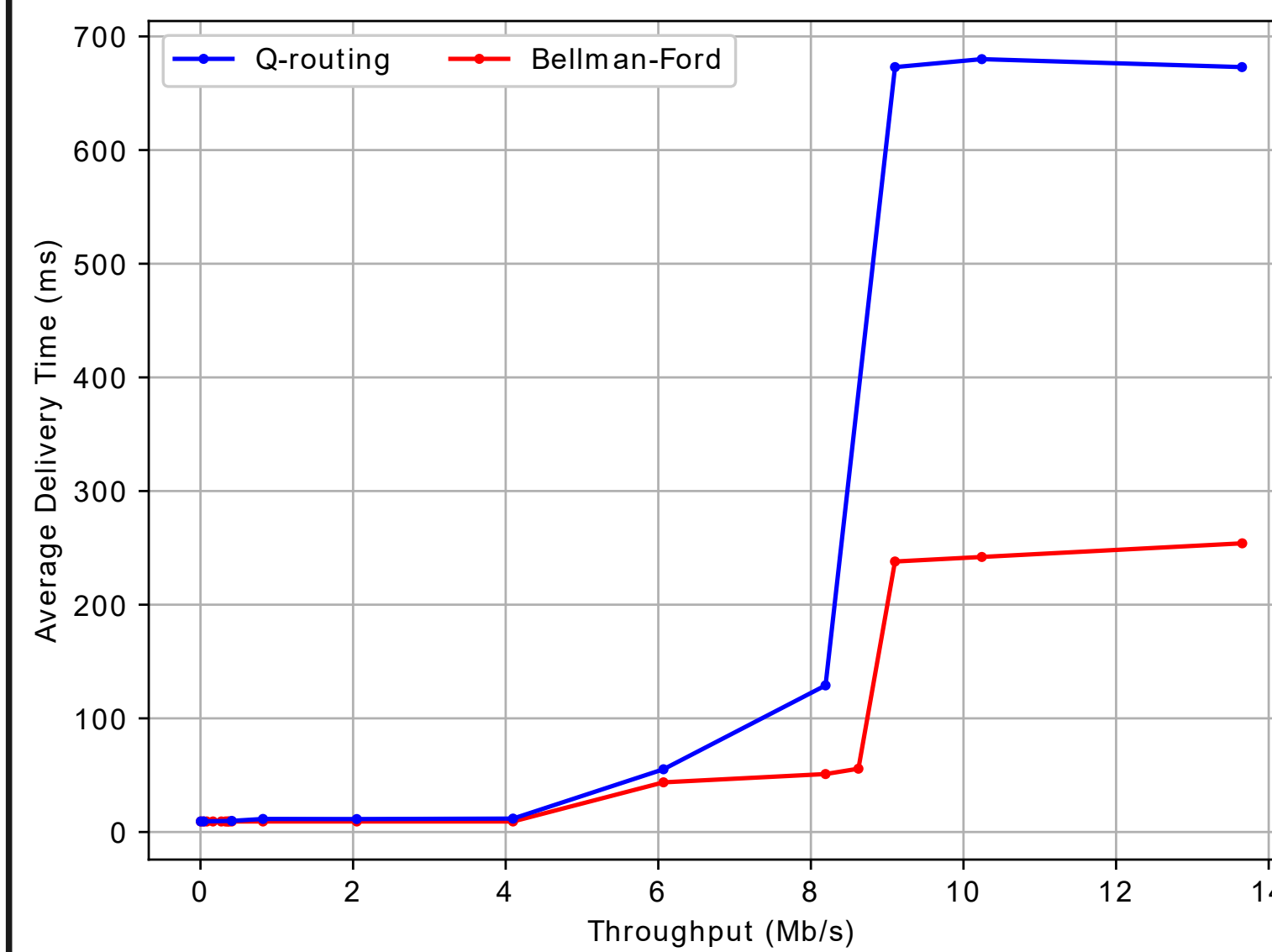


Figure 4 : latence moyenne en fonction du débit des flux CBR

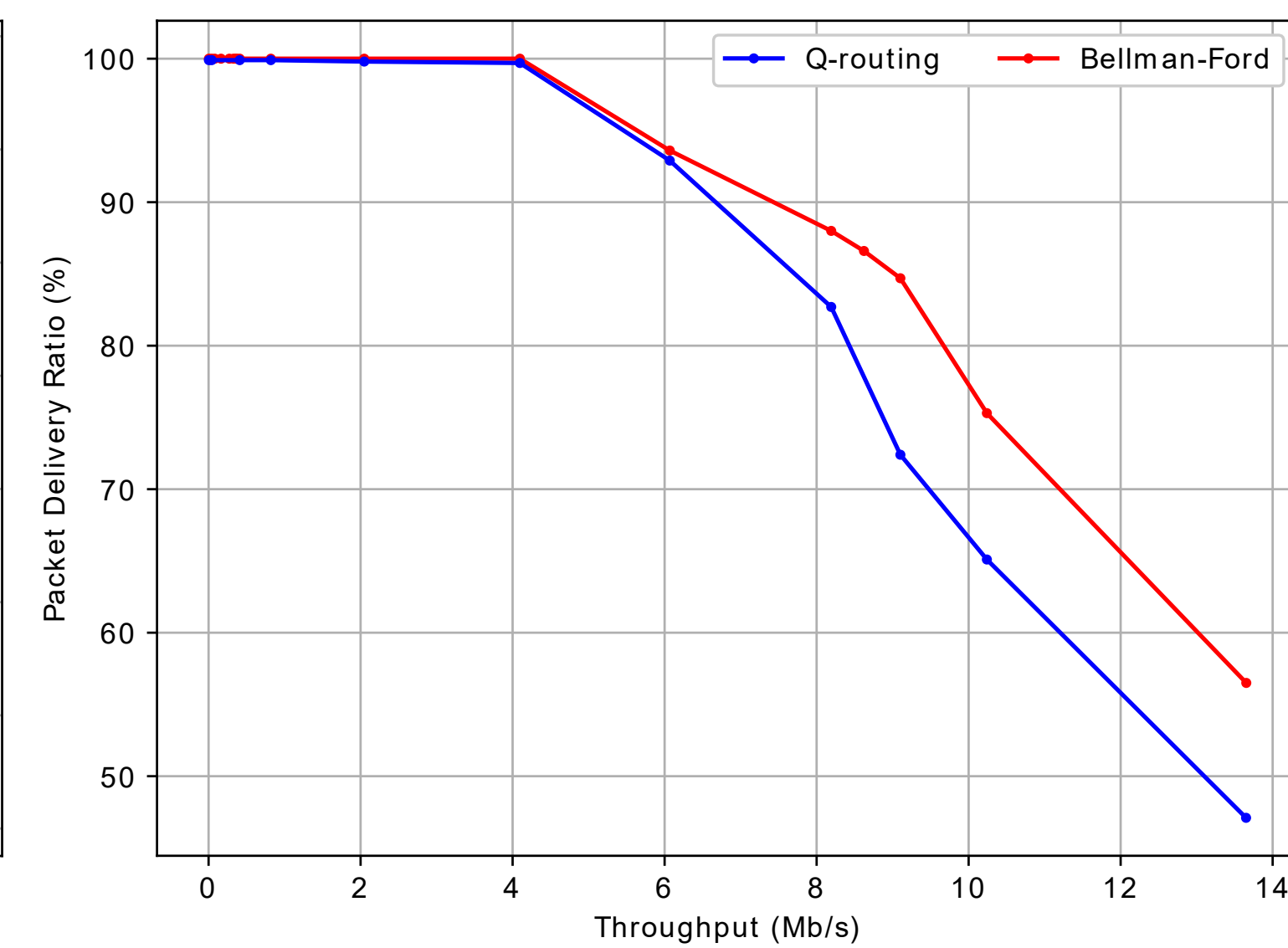


Figure 5 : taux de paquets reçus en fonction du débit des flux CBR

Avec une charge faible en arrière-plan : *Q-routing* ≈ Bellman-Ford

Avec une charge élevée en arrière-plan : *Q-routing* << Bellman-Ford

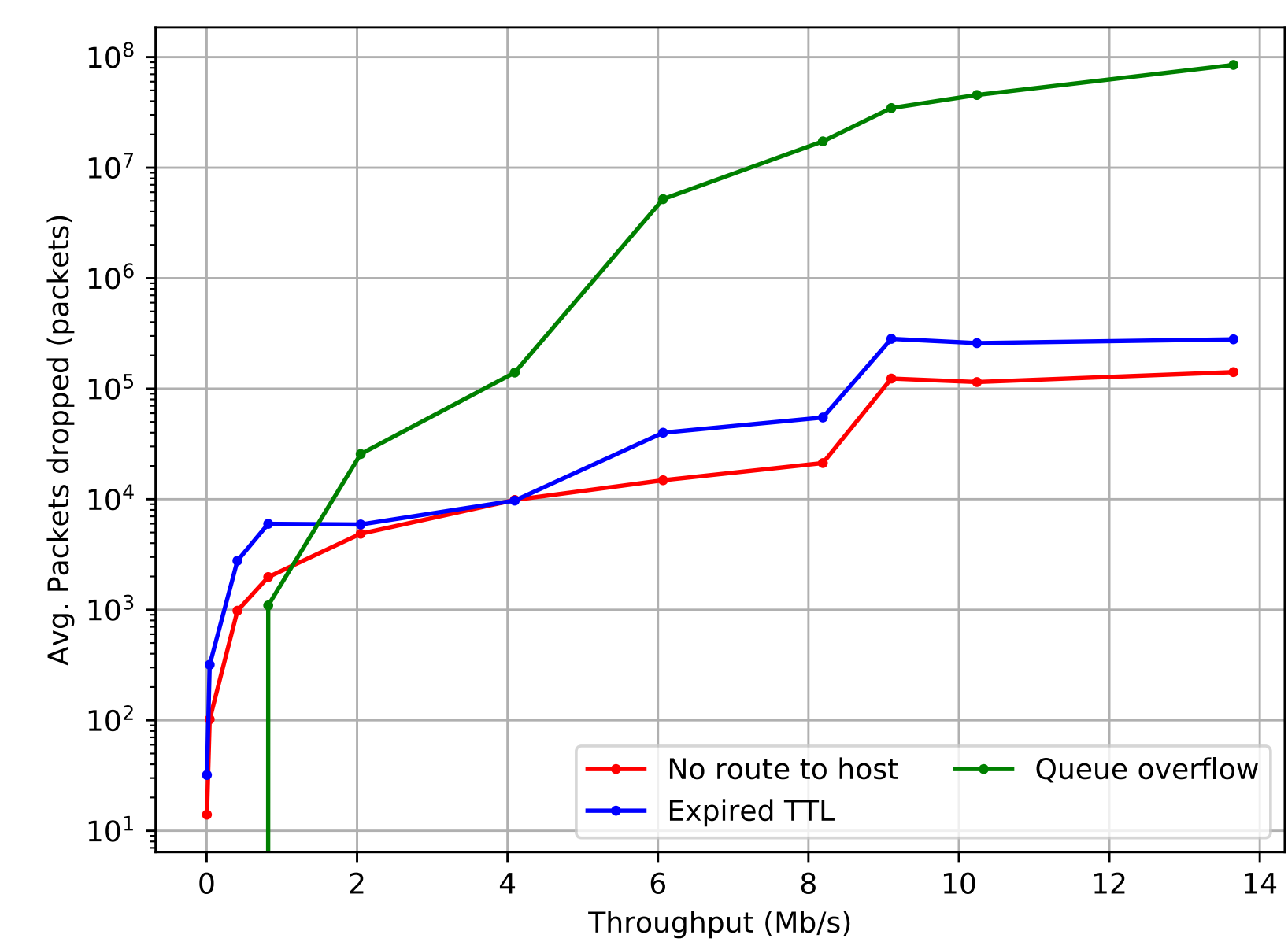


Figure 6 : Pertes des paquets en fonction du débit des flux CBR

Le dépassement de file est la première cause de perte de paquet (en moyenne 10^6 paquets). Il apparaît à des débits relativement faibles, le débit cumulé étant inférieur à la capacité des liens.
 ⇒ Une importante charge réseau est créée en plus des flux CBR. Les boucles de routage peuvent expliquer ce phénomène.

Conclusions et perspectives

Problèmes identifiés :

- la stratégie gloutonne de *Q-routing* montre ses limites lorsque la localisation et le niveau de charge varient beaucoup;
- la mise à jour fréquente de la latence induit un surcoût en termes de charge réseau et de calcul.

Néanmoins,

- notre implémentation est entièrement décentralisée et distribuée;
- *Q-routing* permet le contournement de chemin saturé.

Perspectives :

- implémenter *Q-routing* sur une base OLSR (au lieu d'une base Bellman-Ford);
- réduire le nombre de mise à jour (compromis entre la fraîcheur de l'information et la quantité de mise à jour);
- modifier la récompense pour qu'elle prenne en compte d'autres paramètres que la latence.

Références :

- [1] J. A. Boyan and M. L. Littman, "Packet routing in dynamically changing networks: A reinforcement learning approach," in Advances in neural information processing systems, 1994, pp. 671–678.
- [2] P. Wang and T. Wang, "Adaptive Routing for Sensor Networks using Reinforcement Learning," in The Sixth IEEE International Conference on Computer and Information Technology (CIT'06), 2006, pp. 219–219, doi: 10.1109/CIT.2006.34.
- [3] R. Schoonderwoerd, O. Holland, J. Bruten, and L. Rosenkrantz, "Ants for load balancing in telecommunication networks," HP Labs, Bristol, Tech. Report 96–35, 1996.
- [4] C. J. C. H. Watkins and P. Dayan, "Q-learning," Mach Learn, vol. 8, no. 3, pp. 279–292, May 1992, doi: 10.1007/BF00992698.