

État de l'art des politiques de répartition des flux sur les différentes bandes WiFi

Oussama Benaziz - Jules Bonhotal

25 février 2024

1 Introduction

Notre projet de recherche s'est articulé autour de l'analyse et de la comparaison des résultats issus de 6 études, dans le but de fournir une évaluation comparative de différentes stratégies de répartition des flux. Ce travail a débuté par l'exploration d'un document fournis par nos professeurs référents.

Par la suite, nous avons étendu nos recherches en consultant diverses références, y compris celles proposées comme référence dans le document initial. L'utilisation de Google Scholar nous a permis d'identifier des études partageant des métriques et des stratégies similaires à celles préalablement sélectionnées. Notre recherche a été particulièrement orientée vers les travaux se focalisant sur les stratégies MLO.

Notre méthodologie a ensuite impliqué une analyse des études sélectionnées, visant à extraire les données relatives aux performances des différentes stratégies. Nous avons aussi opéré une comparaison des diverses stratégies, en nous appuyant sur les métriques communes identifiées au cours de nos investigations. En parallèle, nous avons exploré comment ces solutions ont été évaluées, mettant en lumière les méthodologies d'évaluation, les critères de performance, et les contextes spécifiques associés à chaque étude. Cette démarche nous a permis de rédiger des comparaison entre les différentes méthodes de repartions des flux.

Vous pouvez retrouver ce rapport et les autres ressources que nous avons produites dans le cadre de cette UE sur le Git à l'adresse suivante : <https://forge.univ-lyon1.fr/p2007990/ouverture-a-la-recherche>.

1.1 Présentation du MLO

Dans le contexte des réseaux sans fil, notamment dans le cadre du Wi-Fi, une interface désigne le Point de connexion par lequel un périphérique communique avec le réseau. Chaque interface correspond à une bande de fréquences spécifique,

par exemple, 2.4 GHz, 5 GHz ou 6 GHz dans le cas du Wi-Fi 6E. Les interfaces permettent aux appareils de se connecter et de transmettre des données sur des fréquences dédiées, offrant ainsi une capacité de communication sans fil.

Chaque interface est divisée en différentes bandes de fréquences appelées canaux. Chaque interface dispose de plusieurs canaux, et chaque canal a une largeur de bande spécifique sur laquelle les données peuvent être transmises. Les canaux sont utilisés pour segmenter le spectre disponible et éviter les interférences entre les périphériques voisins. En sélectionnant un canal approprié, les périphériques peuvent optimiser leurs performances en minimisant les interférences et en maximisant la qualité de la connexion sans fil.

Avec le Wi-Fi sans la fonctionnalité MLO ou Multi-Link Operation (Wi-Fi 6 et inférieurs), les appareils Wi-Fi utilisent une seule interface radio à la fois pour les transmissions de données. Cependant, ils peuvent utiliser plusieurs canaux et peuvent correspondre aux figures [1a] et [1b]. Avec l'introduction du MLO dans le Wi-Fi 7, un dispositif peut désormais exploiter simultanément plusieurs interfaces radio pour envoyer et recevoir des données sur différentes interfaces, comme présenté dans la figure [1c].

En exploitant plusieurs interfaces, le MLO vise à améliorer les débits de données et à réduire les retards de latence. Le but de ce rapport est d'évaluer si la technologie MLO offre une amélioration par rapport aux versions antérieures, mais surtout de fournir une comparaison entre les différentes méthodes de distribution des données parmi les différentes interfaces et au sein de leurs canaux.

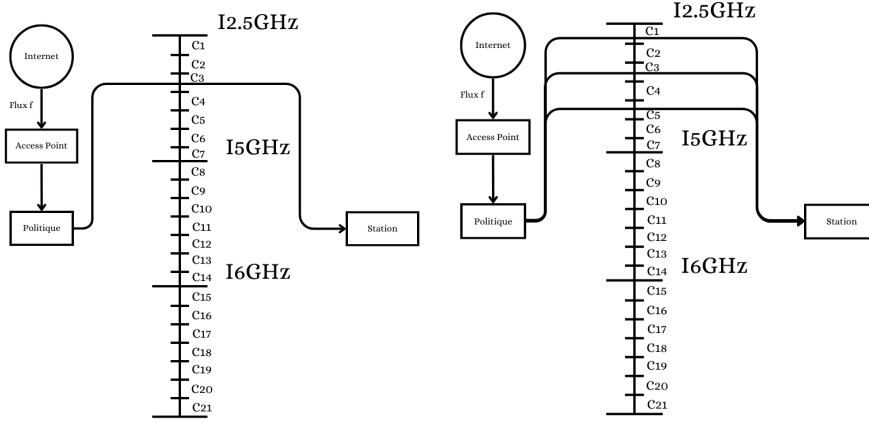
1.2 Présentation des configurations de distribution des flux

Cette section présente les différentes configurations possibles de connexions en fonction des différentes stratégies de répartition des flux qui seront présentées dans ce document.

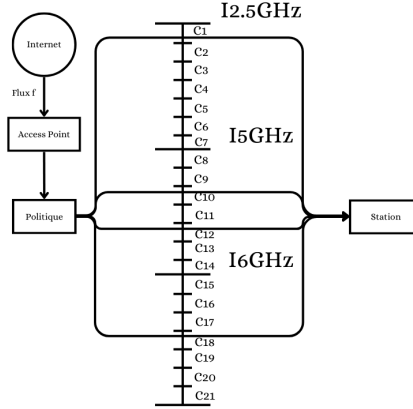
Une interface représente une bande de fréquence spécifique à laquelle un dispositif, tel qu'une station ou un point d'accès (AP), peut émettre ou recevoir des signaux. Chaque interface est associée à une technologie ou une fréquence particulière, par exemple, 2.4GHz, 5GHz, ou 6GHz.

Un canal, d'autre part, désigne une subdivision de la bande de fréquences attribuée à une interface. Chaque interface peut avoir plusieurs canaux, et le choix du canal détermine la fréquence exacte sur laquelle les communications auront lieu. La sélection du canal est un aspect crucial de la gestion du trafic dans le réseau, influençant la qualité et la performance des transmissions.

Dans la figure [1a], on voit une station et une AP, chacune ayant trois interfaces (*12.4GHz*, *5GHz* et *6GHz*). Chacune de ces interfaces est composée de canaux qui peuvent être utilisés pour faire transiter un flux donné.



(a) Mono-Interface Mono-Canal. (b) Mono-Interface Multi-Canal.



(c) Multi-Interface Multi-Canal.

FIGURE 1 – Comparaison des différentes Configurations.

Dans la figure [1a], le flux f arrive à une AP. L'AP décide du lien entre la station et l'AP, qui dans cette figure est le canal ($C3$) de l'interface $12.4GHz$, en fonction de la stratégie qui lui a été attribuée préalablement. L'AP et la station réservent le canal ($C3$) et le flux peut être transféré depuis l'AP vers la station. Cette configuration est mono-interface mono-canal.

Cette figure représente une configuration non-MLO car le flux transite via une seule interface ($12.4GHz$) et utilise un seul canal ($C3$).

La figure [1b] présente une situation similaire à la figure [1a] mais avec une politique multi-canaux.

Le flux f arrive à une AP. L'AP décide du lien entre la station et l'AP, qui dans cette figure est composé des canaux $C1$, $C3$ et $C5$ de l'interface $12.4GHz$, en

fonction de sa stratégie. L'AP et la station réservent les canaux sélectionnés et le flux peut être transféré depuis l'AP vers la station via ces canaux simultanément.

Cette figure représente une configuration non-MLO et multi-canal, car le flux transite en passant par plusieurs canaux.

La figure [1c] présente une configuration Multi-Link Operation (MLO) où une station communique avec une AP via plusieurs interfaces.

Dans cette configuration MLO, le flux f arrive à l'AP. L'AP, en utilisant sa stratégie MLO, décide de répartir le trafic entre les interfaces disponibles et choisit des canaux au sein de ces interfaces. Dans la figure, les canaux $C1$ de $12.4GHz$, $C10$ et $C11$ de $5GHz$, et $C18$ de $6GHz$ sont sélectionnés pour établir le lien entre l'AP et la station. Ces canaux sont réservés par l'AP et la station, permettant ainsi au flux d'être transféré simultanément sur ces différentes interfaces.

Cette représentation illustre une configuration MLO, car le flux transite via plusieurs interfaces ($12.4GHz$, $5GHz$, et $6GHz$) et utilise plusieurs canaux ($C1$, $C10$, $C11$, $C18$).

2 Présentation des Méthodes de Répartitions

Dans le cadre des réseaux Wi-Fi, une méthode de répartition se réfère aux algorithmes et stratégies qui dirigent la distribution des flux de données entre les AP et les canaux disponibles. Ces méthodes sont essentielles pour garantir une utilisation optimale des ressources et minimiser la congestion du réseau. Elles impliquent des décisions telles que l'affectation des flux à des canaux spécifiques. Cette section examinera en détail diverses méthodes de répartition, analysant leurs mécanismes et évaluant leur efficacité dans des conditions variables.

2.1 SLCI

La stratégie Single Link Less Congested Interface (SLCI) est appliquée lors de l'arrivée d'un flux. Elle sélectionne l'interface présentant le moins de congestion, puis alloue le nouveau flux entrant à cette interface.

Un nouveau flux de données arrive dans un AP utilisant le SLCI et est destinée à une station d'un utilisateur.

L'AP utilisant le SLCI analyse les niveaux de congestion sur chaque interface en prenant en compte la charge spécifique de chaque canal. En fonction de ces informations, le SLCI attribue le flux de données à l'interface spécifique et aux canaux présentant la congestion la moins élevée. Pendant la durée de vie du flux, le SLCI surveille dynamiquement la congestion sur l'interface et les canaux attribués.

En cas de changement significatif de la congestion, il peut réévaluer la situation et, si nécessaire, ajuster l'allocation du flux vers une autre interface et d'autres canaux moins encombrés.

Lorsque le flux de données se termine, le SLCI libère l'interface et les canaux qui étaient alloués pour d'autres utilisations.

Dans la référence [3], le gestionnaire de trafic (Traffic Manager) commence par recueillir des informations sur la congestion actuelle des interfaces des autres AP disponibles, puis les transmet à l'AP concernée pour qu'elle puisse les prendre en compte.

Cette stratégie est décrite dans les références suivantes : [2], [4], [5]

2.2 MCAA

La stratégie Multi Link Congestion-aware Load Balancing at flow arrivals (MCAA) s'applique dès l'arrivée d'un flux, répartissant le trafic du nouveau flux entrant en fonction de l'occupation observée des canaux au niveau de l'AP. Cette répartition tient compte des interfaces activées¹ à la station de réception. Plus précisément, soit ρ_i le pourcentage de temps de canal disponible (libre) à l'interface i . Ensuite, la fraction du trafic du flux allouée à l'interface i est donnée par $\ell_i \in J = \frac{\ell \rho_i}{\sum_{j \in J} \rho_j}$, où ℓ représente la charge de trafic et J l'ensemble des interfaces activées à la station cible. Si d'autres flux actifs sont présents à l'AP, leur allocation de trafic reste inchangée.

Les AP utilisant le MCAA observent constamment le trafic sur les interfaces et les canaux, même lorsqu'aucun flux n'est attribué. Lorsqu'un nouveau flux de données arrive dans un AP utilisant le MCAA et doit être distribué à une station, il prend en compte les niveaux moyens de trafic sur chaque interface de la station, en prenant en compte la charge spécifique de chaque canal. En fonction de ces informations, le MCAA attribue le flux de données à l'interface spécifique et aux canaux présentant le trafic moyen le moins élevé. Pendant la durée de vie du flux, l'AP utilisant le MCAA surveille dynamiquement le trafic sur l'interface et les canaux attribués. En cas de changement significatif du trafic, il peut réévaluer la situation et, si nécessaire, ajuster l'allocation du flux vers une autre interface et d'autres canaux moins encombrés.

Lorsque le flux de données se termine, le MCAA libère l'interface et les canaux qui étaient alloués pour d'autres utilisations.

1. Une station est considérée comme activée si au moins une liaison sans fil a été établie avec l'AP en question.

Cette stratégie est décrite dans les références suivantes : [2], [4], [5]

2.3 MCAB

La stratégie Multi Link Congestion-aware Load Balancing (MCAB) opère en collectant les valeurs d'occupation des canaux lors de l'arrivée d'un flux ou à intervalles réguliers. Elle classe ensuite tous les flux, y compris le flux entrant, en ordre croissant en fonction du nombre d'interfaces activées à la station de destination. Les flux avec un nombre inférieur d'interfaces activées sont traités en priorité. En cas d'égalité du nombre d'interfaces activées entre plusieurs flux à la station de destination, la priorité est déterminée par l'heure d'arrivée. Par la suite, la stratégie procède à la reallocation du trafic des flux en suivant la même procédure que celle utilisée dans la stratégie MCAA (Multi Link Congestion-aware Load Balancing at flow arrivals).

Les APs utilisant le MCAB observent constamment la congestion des canaux, même lorsqu'aucun flux n'est à attribuer.

Lorsque des flux de données doivent être envoyés à des stations via l'AP, celui-ci classe tous les flux en ordre croissant en fonction du nombre d'interfaces activées à la station de destination. Les flux avec un nombre moindre d'interfaces activées sont privilégiés. En cas d'égalité quant au nombre d'interfaces activées pour plusieurs flux à la station de destination, la priorité est attribuée en fonction de leur moment d'arrivée. Les flux ayant été reçus antérieurement bénéficient d'une priorité accrue.

L'AP procède ensuite à l'allocation du trafic. Cette allocation suit la même procédure que celle utilisée dans la stratégie MCAA, décrite ci-dessus.

Cette stratégie est décrite dans la référence suivante : [4]

2.4 VDS

La stratégie Séparation Vidéo et Données (VDS) distingue les flux de données des flux vidéo en allouant les flux de données aux interfaces 2,4 GHz ou 5 GHz, tandis que les flux vidéo sont attribués à la interface 6 GHz. Contrairement à d'autres stratégies, le VDS n'alloue pas les flux de données sur plusieurs interfaces, mais assigne l'ensemble du trafic à une seule interface, choisissant toujours la moins encombrée.

La stratégie VDS n'est pas décrite très en détail dans la référence [5], donc il est compliqué de la décrire dans un exemple.

Cette stratégie est décrite dans la référence suivante : [5]

2.5 MH-RSAC

Le Multi-Headed Recurrent Soft-Actor Critic (MH-RSAC) est un algorithme de renforcement adapté au contexte du réseau IEEE 802.11be. Il utilise une version discrète du Soft-Actor Critic (SAC) pour optimiser la distribution du trafic dans un environnement avec des AP et des stations. L'approche MH-RSAC intègre une technique utilisant le machine learning pour traiter les cas où une station a deux ou trois interfaces disponibles. Le processus d'apprentissage de MH-RSAC repose sur des itérations impliquant la sélection d'actions, l'exécution dans l'environnement, le stockage des expériences, l'apprentissage périodique, et la mise à jour des réseaux cibles pour stabiliser l'apprentissage. Cette approche vise à optimiser la distribution du trafic dans des scénarios MLO complexes.

Cette stratégie suit le pseudopode trouvée dans la figure [2] qui viens de [2] a l'adresse <https://arxiv.org/pdf/2303.08959.pdf>.

Algorithm 1: Multi-Headed Recurrent Soft-Actor Critic

```

1 Initialize actor's policy  $\phi$ , critic's Q-function parameters  $\theta_1, \theta_2$  and
  experience replay buffer  $\mathcal{D}$ . Set critic's Q-target function equal to main
  parameters  $\hat{\theta}_1 \leftarrow \theta_1$  and  $\hat{\theta}_2 \leftarrow \theta_2$ ,  $\text{op} \in \{\text{min}, \text{avg}\}$  corresponds to
  the operator,  $L_{\text{per}}$  the update periodicity,  $n_{\text{up}}$  number of updates per
  learning step and  $\rho \in [0, 1]$  is the polyak factor.
2 for environment step  $t \leftarrow 1$  to  $T$  do
3   Observe state  $s = [C_1^t, C_2^t, C_3^t, O_f, T_{id}]$  and select action for each
     head  $a^{h_{1,2}} \sim \pi_{\phi}^{h_{1,2}}(\cdot|s)$ 
4   Execute  $a$  in the environment
5   Observe next state  $s'$  and reward  $r$ 
6   Store  $(s, a, r, s')$  in experience replay buffer  $\mathcal{D}$ 
7   if  $t \bmod L_{\text{per}} = 0$  then
8     for  $n \leftarrow 1$  to  $n_{\text{up}}$  do
9       Randomly sample a batch of transitions with size  $B$  from  $\mathcal{D}$ 
10      Compute targets for the Q-functions (Critic) for each head
11       $h_{1,2}$ :
12      where  $\hat{a}^{h_{1,2}} \sim \pi_{\phi}^{h_{1,2}}(\cdot|s')$ 
13      Update Q-functions (Critic) considering each head  $h_{1,2}$ :
14       $L_1^{h_{1,2}}(\theta_1, \mathcal{D}) = \mathbb{E}_{(s,a,r,s') \in B} [(Q_{\theta_1}^{h_{1,2}}(s, a^{h_{1,2}}) - y_1(r, s'))^2]$ 
15       $L_2^{h_{1,2}}(\theta_2, \mathcal{D}) = \mathbb{E}_{(s,a,r,s') \in B} [(Q_{\theta_2}^{h_{1,2}}(s, a^{h_{1,2}}) - y_2(r, s'))^2]$ 
16       $L(\phi^{h_{1,2}}, \mathcal{D}) = L_1^{h_{1,2}}(\theta_1, \mathcal{D}) + L_2^{h_{1,2}}(\theta_2, \mathcal{D})$ 
17      Update policy (Actor) considering each head  $h_{1,2}$ :
18       $L_1^{h_{1,2}}(\phi^{h_{1,2}}, \mathcal{D}) = \mathbb{E}_{(s) \in B} [\text{op}_{k=1,2} Q_{\theta_{1,k}}^{h_{1,2}}(s, \hat{a}^{h_{1,2}}) -$ 
19       $\alpha \log \pi_{\phi}^{h_{1,2}}(\hat{a}^{h_{1,2}}|s)]$ 
20       $L_2^{h_{1,2}}(\phi^{h_{2,2}}, \mathcal{D}) = \mathbb{E}_{(s) \in B} [\text{op}_{k=1,2} Q_{\theta_{1,k}}^{h_{1,2}}(s, \hat{a}^{h_{1,2}}) -$ 
21       $\alpha \log \pi_{\phi}^{h_{1,2}}(\hat{a}^{h_{1,2}}|s)]$ 
22       $L(\phi^{h_{1,2}}, \mathcal{D}) = L_1^{h_{1,2}}(\phi^{h_{1,2}}, \mathcal{D}) + L_2^{h_{1,2}}(\phi^{h_{2,2}}, \mathcal{D})$ 
23      Finally, update Q-functions target networks with Polyak
24      averaging:
25       $\hat{\theta}_i \leftarrow \rho \hat{\theta}_i + (1 - \rho) \theta_i$  for  $i = 1, 2$ 
26    end
27  end

```

FIGURE 2 – Pseudopode de la stratégie MH-RSAC

Lorsqu'un nouveau flux arrive, l'algorithme doit décider de la manière dont le trafic associé sera distribué entre les différentes interfaces disponibles.

MH-RSAC sélectionne la tête appropriée en fonction du nombre d'interfaces de la station. Dans ce contexte, une "tête" représente une entité responsable de la prise

de décision pour un certain nombre d'interfaces spécifiques de la station. Chaque tête est responsable de la prise de décision pour un certain nombre d'interfaces (par exemple, une tête pour deux interfaces et une autre pour trois interfaces). Une tête prend en compte ses choix pour passer à ses nouvelles décisions. La tête sélectionnée choisit si le flux doit être affecté à une ou plusieurs interfaces et lesquelles. Il n'est pas précisé comment le ou les canaux sont sélectionnés.

Les résultats de l'exécution, y compris les récompenses ou les pénalités associées à la distribution du trafic, sont enregistrés en tant qu'expériences dans la mémoire de l'algorithme. Périodiquement, l'algorithme utilise les expériences stockées pour mettre à jour ses connaissances et améliorer sa capacité à prendre des décisions plus efficaces dans des situations similaires à l'avenir.

Le processus se répète à chaque arrivée de nouveau flux, permettant à MH-RSAC de continuer à affiner ses politiques de distribution du trafic en fonction des conditions changeantes du réseau.

Cette stratégie est décrite dans la référence suivante : [2]

2.6 SL et MB-SL

Le Single Link (SL) est une approche où les appareils communiquent via une seule interface sans fil dans une bande de fréquence spécifique, telle que 2,4 GHz ou 5 GHz (la manière dont le choix est fait n'est pas détaillée). Chaque station est connectée à son AP respectif via une seule interface et un seul canal. Le Multi-Band Single Link (MB-SL) introduit la capacité d'utiliser plusieurs canaux au sein d'une même interface parmi les suivantes : 2,4 GHz, 5 GHz et 6 GHz. Cela permet aux appareils de répartir leur communication sur différents canaux, réduisant potentiellement la congestion et améliorant les performances globales du réseau. La méthode de choix des canaux ou de l'interface n'est pas détaillée.

Ces approches n'utilisent pas le MLO et servent à mesurer l'amélioration apportée par le MLO.

Ces deux stratégies ne sont pas décrites en détail dans la référence [5], car elles ne sont là que pour représenter les anciennes méthodes qui n'utilisent pas le MLO. Il est donc compliqué de fournir un exemple de leur fonctionnement.

Cette stratégie est décrite dans la référence suivante : [5]

2.7 MABs

L'article discute d'une stratégie d'allocation dynamique basée sur l'utilisation d'algorithmes d'apprentissage automatique, en particulier les Bandits Multibras (MABs), pour optimiser la gestion d'un Réseau Local sans Fil (WLAN). Cette approche novatrice repose sur deux composantes clés : l'Allocation Dynamique de Canal (DCA) et la Sélection Dynamique d'AP (DAPS). Le DCA permet aux AP de choisir dynamiquement les canaux radio afin d'optimiser l'utilisation des fréquences disponibles. Simultanément, le DAPS autorise les stations WLAN à sélectionner de manière adaptative l'AP le plus approprié en fonction de la qualité de la liaison et de la charge du réseau. Ces mécanismes sont conçus pour améliorer l'efficacité du réseau en réduisant les délais d'accès, en optimisant l'utilisation des canaux et en ajustant les récompenses en fonction des performances du réseau. L'objectif global est de créer un réseau WLAN adaptatif, où les actions des agents sont guidées par des récompenses et des retours d'expérience, permettant ainsi une amélioration continue des performances. Par exemple, une réduction significative des délais d'accès et une augmentation de l'utilisation des canaux peuvent être observées, contribuant ainsi à une expérience utilisateur améliorée. Ce paradigme d'allocation dynamique offre une solution proactive et autonome pour l'optimisation des réseaux WLAN, contribuant à une meilleure qualité de service et à une utilisation plus efficace des ressources.

Cette stratégie est décrite dans la référence suivante : [6]

2.8 Synthèses sur les solutions décrites

Dans cette section, nous présentons une comparaison des différentes stratégies de répartition des flux que nous avons étudiées.

Dans la tables 1 et 2 on retrouve les information suivantes :

- **MLO** : Indique si la stratégie intègre des techniques d'apprentissage automatique (MLO) pour prendre des décisions adaptatives.
- **Multy-Canaux** : Précise si la stratégie permet l'utilisation simultanée de plusieurs canaux pour distribuer les flux de données.
- **Critères de sélection des liens** : Décrit les métriques ou critères spécifiques utilisés par chaque stratégie pour prendre des décisions sur la répartition des flux.
- **Références** : Fournit les références bibliographiques des études qui ont été analysées pour chaque stratégie.
- **Sections** : Indique les parties spécifiques du rapport où chaque stratégie est examinée et discutée en détail.
- **Type d'algorithme** : Présente le type d'algorithme utilisé par chaque stratégie pour prendre des décisions de répartition des flux.

- **Basée sur AP/station** : Indique si la stratégie est exécutée sur l’AP ou la station pour prendre des décisions.
- **Granularité du choix** : Décrit le niveau de détail auquel la stratégie effectue des choix, que ce soit au niveau de l’interface, du canal, etc.

Ainsi, ces éléments offrent une vue d’ensemble complète des caractéristiques étudiées pour chaque stratégie, facilitant ainsi la compréhension et la comparaison des différentes approches.

Stratégie	MLO	Multy-Canaux	Critère de sélection du/des liens	Références	Section
SLCI	Non	Oui	Taux d’occupation des Interfaces	[2], [4], [5]	3.1, 3.4, 3.5
MCAA	Oui	Oui	Taux d’occupation des Interfaces et des canaux	[2], [4], [5]	3.1, 3.4, 3.5
MCAB	Oui	Oui	Taux d’occupation des Interfaces et des canaux	[4]	3.4
VDS	Non	Non	Type de données et Taux d’occupation des Interfaces	[5]	3.1
MH-RSAC	Oui	Oui	Nombre d’interfaces disponibles	[2]	3.5
SL	Non	Non	Taux d’occupation des Interfaces	[5]	3.1
MB-SL	Non	Oui	Taux d’occupation des Interfaces	[4], [5]	3.1, 3.4
MABs	Oui	Oui	Délais d’accès, Utilisation des canaux	[3], [6]	3.6

TABLE 1 – Comparaison des stratégies partie 1.

Stratégie	Type d’algorithme	Basée sur AP/station	Granularité du choix
SLCI	Observation continue des liens potentiels	AP (+ contrôleur externe pour [4])	Interface et canal
MCAA	Observation continue des liens potentiels	AP (+ contrôleur externe pour [4])	Interface et canal
MCAB	Observation continue des liens potentiels	AP (+ contrôleur externe pour [4])	Interface et canal
VDS	Séparation par type de données	AP	Interface
MH-RSAC	Machine Learning	AP	Interface et canal
SL	Non MLO	AP	Interface
MB-SL	Non MLO	AP	Interface
MABs	Machine Learning	AP et Station	Interface et canal

TABLE 2 – Comparaison des stratégies partie 2.

Enfin, la Table 3 présente les critères de comparaison utilisés dans différentes références pour évaluer les performances des stratégies.

3 Analyse de performances des solutions étudiées

Dans cette section, présenterons notre évaluation des performances des stratégies de répartition des flux que nous avons étudiées. Après avoir présenté les caractéristiques et les méthodes de répartition dans la section précédente, nous nous penchons désormais sur les résultats concrets obtenus par chaque stratégie. Nous examinons les performances à travers divers critères et métriques évalués dans les études que nous avons analysées. L’objectif est de dégager des tendances, de comparer les résultats et de discuter de l’efficacité relative de chaque approche dans des contextes spécifiques.

	[2]	[4]	[5]	[6]
SLCI	Perte de Débit	Satisfaction des requêtes + taux de congestion des ca- naux	perte de débit moyen	
MCAA	Perte de Débit	Satisfaction des requêtes + taux de congestion des ca- naux	perte de débit moyen	
MCAB		Satisfaction des requêtes + taux de congestion des ca- naux		
VDS			perte de débit moyen	
MB-SL			perte de débit moyen	
MH-RSAC	Perte de Débit			
Multi-Armed Bandits				convergence du système, ré- duction de la variabilité des performances

TABLE 3 – Critères de comparaison des Stratégies par référence

3.1 Multi-link Operation in IEEE 802.11be WLANs [5]

3.1.1 Méthodologie

L'étude évalue les performances d'une stratégie d'allocation de trafic MLO. Elle utilise des simulations sur 500 déploiements avec différentes configurations.

3.1.2 Métriques Mesurées

L'étude utilise le pourcentage de perte de vitesse de traitement (throughput) vis à vis de la vitesse de traitement optimal pour les différents point de passage du flux. Elle se sert de cet indicateur de performance pour évaluer différentes méthodes d'allocation de trafic, en se concentrant sur les flux vidéo et de données. L'objectif est de maintenir ces pertes de vitesse de traitement en dessous d'un seuil de 5%.

3.1.3 Simulation et Paramètres

La simulation est composée sur 500 déploiements aléatoires, évalués via la plateforme Neko. Les AP et stations sont configurés avec 3 interfaces sur différentes bandes de fréquences. Les politiques SLCI et MCAA sont appliquées de manière aléatoire au AP, sauf pour l'AP mesurée qui utilise SLCI, MCAA, VDS ou MSLA de manière prédéterminée. Seul le trafic descendant est considéré, avec des stations demandant des flux vidéo ou de données.

3.2 Neko

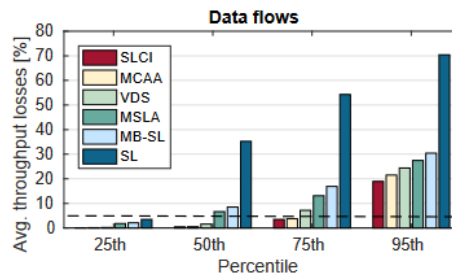
Les simulations au niveau des flux sont réalisées à l'aide de la plateforme de simulation Neko3, qui met en œuvre l'abstraction CSMA/CA. Cette abstraction

repose sur l'occupation du canal observée par chaque point d'accès (AP) pour calculer le temps d'antenne allouable à chaque flux, préservant ainsi la part équitable inhérente du Wi-Fi dans les ressources spectrales. Le simulateur Neko3 a été développé par l'Université Pompeu Fabra <https://www.upf.edu/web/wnrg/windmal>. Il pouvait être trouvé sur cette page Git <https://github.com/wn-upf>, mais tous les liens vers les simulateurs ou les pages les décrivant sont des liens morts et nous n'avons pas plus de connaissance à son sujet.

3.2.1 Présentation des Résultats de l'étude

3.3 Résultat des stratégie

- **Politique VDS (Virtual Distribution System) :**
 - Pertes moyennes de débit maintenues sous 5% dans seulement 50% des 500 scénarios évalués.
 - Dans les pires cas, les pertes de débit pour les flux vidéo atteignent près de 40%.
- **Politiques SLCI (Single Link Control Information) et MCAA (Multi-Channel Access with Aggregation) :**
 - Surmontent les problèmes de performance de VDS dans 75% des scénarios en équilibrant la charge de trafic.
- **Comparaison avec les approches héritées (MB-SL et SL) :**
 - MB-SL présente des avantages par rapport à SL en réduisant la congestion mais maintient les pertes de débit sous 5% dans seulement 25% des scénarios.
 - Les stratégies MLO maintient des pertes de débit acceptables dans 75% des scénarios, montrant son potentiel pour des scénarios actuellement problématiques avec SL et MB-SL.



(a) Pertes moyennes de débit pour les flux vidéo et de données.

FIGURE 3 – Figure pour les pertes moyennes de débit pour les flux vidéo et de données.

Selon les résultats présentés et comme on peut le voir dans la figure [3], les politiques SLCI et MCAA semblent donner les meilleurs résultats en termes de débit, mais SLCI a une moins bonne satisfaction moyenne des requêtes. Ces résultats indiquent que les stratégies MCAB et MCAA sont meilleures que les stratégies non-MLO VDS, MB-SL, SL et SLCI. Cependant, il est difficile de sélectionner la meilleure stratégie entre MCAB et MCAA.

3.4 Dynamic traffic allocation in ieee 802.11be multi-link wlans [4]

3.4.1 Méthodologie

Cette étude utilise la plateforme de simulation Neko pour évaluer les performances de différentes politiques de répartition. 500 simulations sont réalisées (dans la partie qui nous intéresse). Elles font varier les paramètres tels que la fraction de stations MLO, les politiques d'allocation de trafic, et la période d'adaptation. Il est important de noter que les simulations considèrent le trafic descendant uniquement, et les données transmises sont des vidéos.

3.4.2 Métriques Mesurées

Les métriques utilisées comprennent la satisfaction moyenne des flux, le taux de d'occupation des interfaces et l'efficacité spectrale. La satisfaction d'un flux est le rapport entre le temps d'antenne alloué par le point d'accès pendant la durée de vie du flux et la quantité totale de temps d'antenne requise. Le taux de d'occupation quantifie le niveau d'engorgement ou de saturation sur chaque interface, permettant d'identifier les éventuels points de congestion. Enfin, l'efficacité spectrale évalue la capacité du système à tirer le meilleur parti des ressources spectrales disponibles, crucial dans un contexte où plusieurs interfaces radio sont utilisées simultanément.

3.4.3 Paramètres de simulation

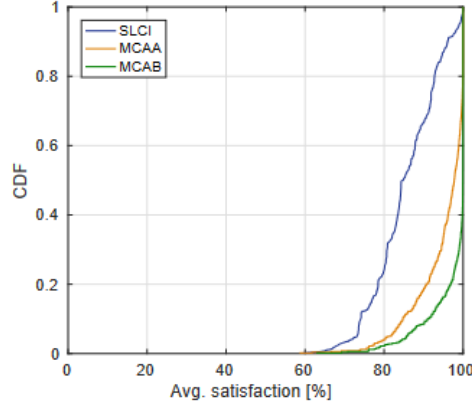
Les simulations de flux ont été effectuées avec la plateforme Neko. Les paramètres comprennent des éléments tels que la fréquence porteuse (2,437 GHz/5,230 GHz/6,295 GHz), la largeur de canal (20 MHz/40 MHz/80 MHz), la puissance de transmission des AP/stations (20/15 dBm), le seuil de détection (CCA, -82 dBm), la durée moyenne des flux ($T_{on} = 3$ s), la période d'adaptation MCAB ($\sigma = 1$ s), et le taux d'erreur des paquets (10%).

Chaque simulation dure 120 secondes, avec différentes configurations de fréquence et de canal, ainsi qu'une variabilité du nombre de simulations. Ces paramètres ont été choisis pour refléter des conditions réalistes et diverses dans un environnement Wi-Fi.

3.4.4 Présentation des Résultats de l'étude

- **Politique SLCI (Single Link Less Congested Interface) :**
 - Satisfaction moyenne des flux : Résultats pire que MCAA et MCAB avec une chute du pourcentage de satisfaction de 8% avec CDF à 0.8. Amélioration de 6% par rapport à MCAA pour certains cas.
 - Congestion des interfaces : Problème d'engorgement sur l'interface 6 GHz, mais les autres interfaces restent la grande majorité du temps sous les 90%.
 - Efficacité spectrale : Utilisation inéquitable des interfaces.
- **Politique MCAA (Multi Link Congestion-aware Load balancing at flow arrivals) :**
 - Satisfaction moyenne des flux : Performances mitigées, en particulier pour les flux vidéo.
 - Congestion des interfaces : Les interfaces 2,4 GHz et 5 GHz présentent des niveaux de congestion important.
 - Efficacité spectrale : Utilisation sous-optimale des ressources spectrales.
- **Politique MCAB (Multi Link Congestion-aware Load balancing) :**
 - Satisfaction moyenne des flux : Amélioration significative, maintient une satisfaction de 95% pour la plupart des scénarios.
 - Congestion des interfaces : Équilibre plus uniforme entre les interfaces, en particulier pour le 6 GHz.
 - Efficacité spectrale : Maximisation de la livraison du trafic, réduction des épisodes de congestion.

On peut voir dans la figure [4] ainsi que dans les autres données de la référence [4] que les stratégies MCAB et MCAA présentent les meilleurs résultats, à la fois pour la satisfaction moyenne des requêtes et en proposant une meilleure répartition des flux afin d'éviter la congestion des canaux. MCAA tend à avoir une meilleure satisfaction dans un environnement comportant plus d'AP non MLO et répartit les flux de manière plus homogène sur les différentes interfaces. Cependant, MCAB a tendance à fournir de meilleurs résultats dans les environnements contenant une grande proportion de MLO.



(a) Satisfaction moyenne des Requêtes par Stratégie.

FIGURE 4 – Figure pour la satisfaction moyenne des requêtes par politique.

3.5 RL meets Multi-Link Operation in IEEE 802.11be :Multi-Headed Recurrent Soft-Actor Critic-based Traffic Allocation [2]

3.5.1 Méthodologie

La méthodologie adoptée dans cette recherche repose sur l'utilisation d'un réseau IEEE 802.11be avec un ensemble prédéfini de M AP et N stations attachées par AP. Tous les AP sont considérés comme ayant la capacité Multi-Link Operation (MLO) avec 3 interfaces disponibles. Les stations peuvent avoir une capacité MLO variable. Les positions des stations par rapport à leur AP sont définies, avec 80% des utilisateurs placés de manière aléatoire dans un rayon r allant de 1 à 8 m et le reste dans un rayon r allant de 1 à 3 m. Les AP et les stations prennent en charge jusqu'à 16 flux spatiaux MIMO simultanés. Le modèle de perte de propagation suit un modèle d'entreprise, et un contrôle adaptatif du taux basé sur le rapport signal sur bruit (SNR) avec un taux de modulation QAM maximal de 4096 est appliqué.

3.5.2 Métriques

La Satisfaction des Flux (FS) est une mesure qui évalue le degré de contentement des flux de trafic dans un réseau. Elle est calculée en comparant le temps d'air alloué par le point d'accès (AP) pendant la durée de vie d'un flux avec le temps d'air total nécessaire. Une FS élevée indique une utilisation efficace des ressources

réseau, garantissant une expérience satisfaisante pour les utilisateurs.

Le Taux de Chute de Débit (TDR) est une métrique qui évalue la performance d'un système en mesurant la proportion du trafic dont le débit est perdu après une décision d'allocation de trafic. Il est calculé en comparant le débit attendu avec le débit réel après la décision d'allocation. Un faible TDR indique une meilleure efficacité dans la distribution du trafic.

3.5.3 Simulateur

Le simulateur utilisé est Neko 802.11be, un simulateur de niveau de flux. Les simulations sont réalisées en utilisant ce simulateur conjointement avec des agents d'apprentissage par renforcement (RL) basés sur PyTorch. La communication entre le simulateur et les agents RL s'effectue via la bibliothèque ZMQ broker. Les paramètres de simulation incluent la largeur de bande du canal (20 MHz/40 MHz/80 MHz/160 MHz), la fréquence porteuse (2.437 GHz/5.230 GHz/6.295 GHz), 5 APs, 16 flux spatiaux max, et le modèle de perte de propagation. Les types de flux (WB, V4K, VR) ont des distributions spécifiques. Les paramètres d'apprentissage par renforcement comprennent des valeurs telles que le taux d'apprentissage de l'acteur ($1e-4$), du critique ($1e-3$), le taux de remise (0.99), les couches cachées (64, 64), la taille des lots (512), la périodicité des mises à jour (50 étapes), et le nombre de mises à jour par session (10).

3.5.4 Présentation des Résultats de l'étude

- **MH-RSAC (Multi-Headed Recurrent Soft-Actor Critic) :**
 - Amélioration du TDR : Jusqu'à 35.2% par rapport à SLCI et 6% par rapport à MCAA.
 - Augmentation de la FS : Jusqu'à 25.6% et 6% par rapport à SLCI et MCAA, respectivement.
- **MCAA (Multi-Link Congestion-aware Load balancing at flow arrivals) :**
 - Performance mitigée de la FS : Surtout pour les flux vidéo.
 - Niveaux de congestion élevés : En particulier sur les interfaces 2,4 GHz et 5 GHz.
- **SLCI (Single Link Less Congested Interface) :**
 - Chute du pourcentage de satisfaction : Résultats moins favorables que MCAA et MCAB avec une baisse de 8% à CDF 0.8.
 - Problème d'engorgement sur l'interface 6 GHz : Autres interfaces restent généralement en dessous de 90%.

Comme on peut le voir dans la figure [5] tirée de la référence [2], la meilleure stratégie dans toutes les situations est MH-RSAC. Cependant, si cette différence

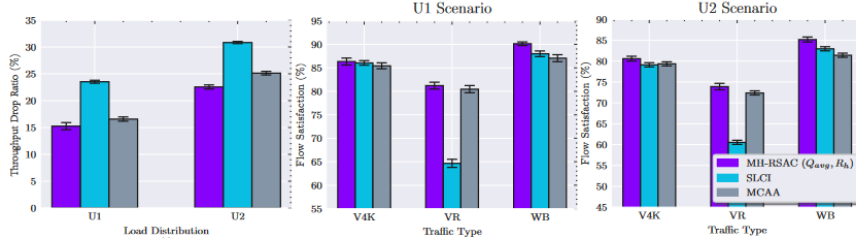


FIGURE 5 – Comparaison entre les résultats de la référence [2].

est très visible entre MH-RSAC et SLCI, elle est moins marquée avec MCAA, qui a des résultats similaires à MH-RSAC pour quasiment toutes les situations.

3.6 Concurrent Decentralized Channel Allocation and Access Point Selection using Multi-Armed Bandits in multi BSS WLANs [6]

3.6.1 Méthodologie

Le texte aborde la problématique de l’optimisation des réseaux locaux sans fil d’entreprise (WLAN) en adoptant une approche méthodologique basée sur l’utilisation de bandits multi-bras (MABs). Il explore la faisabilité de résoudre les problèmes de sélection de canal décentralisée et de sélection d’AP dans des environnements WLAN d’entreprise. La méthodologie adoptée consiste à conférer aux AP et aux stations des agents capables d’apprendre et d’explorer la meilleure configuration de canal et d’association d’AP.

3.6.2 Métriques Mesurées

Pour évaluer l’efficacité de l’approche proposée, le texte utilise diverses métriques de performance. Ces mesures incluent la convergence du système vers des solutions optimales, la réduction de la variabilité des performances entre différents scénarios, et l’évaluation de la performance du système dans des conditions de réseau variées (différents trafics, nombre de stations et APs). Les métriques sont utilisées pour quantifier l’impact du cadre adaptatif basé sur les MABs par rapport à l’approche statique.

3.6.3 Présentation des Résultats de l’étude

Les résultats obtenus révèlent que le cadre adaptatif utilisant les MABs surpasse l’approche statique, indépendamment de la densité du réseau et des exigences de trafic. Ils montrent que le système converge vers des solutions plus performantes,

réduisant la variabilité des performances entre différents scénarios. De plus, les résultats mettent en lumière que la même performance (ou meilleure) peut être atteinte avec moins d'APs pour un nombre donné de stations. Ils soulignent également l'efficacité de l'algorithme de Thompson sampling dans des scénarios adversaires, où les actions d'un agent influent sur les récompenses observées par les autres agents.

Ces résultats démontrent la pertinence de l'approche basée sur les MABs pour la résolution des problèmes de configuration de réseau dans les WLANs d'entreprise, mettant en avant son adaptabilité et ses performances supérieures par rapport à des approches statiques.

3.7 OnlinePrimary Channel Selection for Dynamic Channel Bonding in High-Density WLANs [1]

3.7.1 Méthodologie

Dans le cadre de ce papier, plusieurs méthodes de répartition des canaux ont été explorées pour optimiser les performances des réseaux locaux sans fil (WLANs). Tout d'abord, l'allocation de canaux (CA) a été étudiée, permettant d'attribuer des canaux spécifiques à différentes stations ou réseaux, soit de manière statique, soit de manière dynamique en fonction des besoins. Ensuite, le bonding de canaux dynamique (DCB) a été utilisé pour combiner plusieurs canaux adjacents et augmenter la bande passante disponible. En combinant CA et DCB, les chercheurs ont cherché à maximiser l'utilisation du spectre et les performances globales du réseau.

3.7.2 Métriques Mesurées

Pour évaluer l'efficacité des méthodes de répartition des canaux, plusieurs métriques importantes ont été utilisées. Celles-ci comprennent le débit de données, qui mesure la quantité d'informations transférées par unité de temps, la latence, qui représente le temps écoulé entre l'envoi et la réception des données, le taux d'erreur, qui indique la probabilité de transmission incorrecte des données, et l'efficacité spectrale, qui mesure la capacité du système à utiliser efficacement le spectre disponible.

3.7.3 Présentation des Résultats de l'étude

Dans cette étude, quatre méthodes de sélection de canal primaire ont été évaluées : Fixe Primaire (FP), Dynamique-Aléatoire (DR), Dynamique-Libre (DF), et Dynamic-wise (DyWi). La méthode FP maintient le canal primaire constant en

toutes circonstances, ce qui peut limiter son adaptabilité aux variations du trafic. En revanche, DR sélectionne le canal primaire de manière aléatoire à chaque itération, introduisant ainsi de la variabilité dans le réseau. DF, quant à elle, choisit le canal primaire en fonction de sa disponibilité et de sa charge de trafic, ce qui peut favoriser une utilisation plus efficace des ressources. Enfin, DyWi vise à maximiser le débit en sélectionnant le canal primaire en fonction du débit moyen prévu pour chaque itération, ce qui lui permet d’optimiser les performances du réseau en tenant compte des conditions de trafic et des capacités du canal. Les résultats ont montré que DyWi surpassait les autres méthodes en termes de débit et de retard moyen, offrant ainsi une solution efficace pour améliorer les performances des réseaux WLANs.

4 Comparaison des politiques

La politique de répartition de trafic qui semble actuellement la plus performante est la politique MCAB (Multi Link Congestion-aware Load balancing).

4.1 Comparaison MLO et non-MLO

Dans de nombreux articles comparant les stratégies MLO et non-MLO, les résultats penchent en faveur des stratégies MLO. Par exemple, dans l’étude de López-Raventós et al. (2022) [4], la figure [4a] compare la stratégie non-MLO SLCI aux stratégies MLO MCAA et MCAB en se basant sur le pourcentage de satisfaction des requêtes des stations. Les résultats montrent clairement que la stratégie non-MLO SLCI présente des performances moins bonnes que les stratégies MLO. De même, dans la figure [3a] issue de la référence [5], qui compare la perte de quantité de données transférées, on observe les mêmes résultats.

Cependant, il est important de noter que les stratégies non-MLO étaient souvent des versions simplifiées utilisées uniquement à des fins de comparaison. De plus, les stratégies non-MLO étaient souvent mises en œuvre pour mesurer leur impact sur les performances des stratégies MLO dans un réseau, comme la stratégie VDS illustrée dans [5]. Les stratégies MLO étaient quant à elles souvent choisies avec un certain degré de sélection et d’optimisation pour l’environnement donné, introduisant ainsi un certain biais en faveur des stratégies MLO. Malgré ces biais, il semble relativement évident au vu des résultats des documents analysés que les stratégies MLO sont meilleures que les stratégies non-MLO.

4.2 Débit

Comme on peut le voir dans la figure [3], la référence [5] trouve que parmi les stratégies qu'elle compare, MCAA est la meilleure en termes de débit. Cependant, dans la référence [2], notamment dans la figure [5], on peut voir que la stratégie MH-RSAC est meilleure que MCAA, ce qui semble indiquer que la meilleure stratégie parmi celles analysées est MH-RSAC.

Il est important de noter que la différence de performance entre MCAA et MH-RSAC est relativement faible. De plus, cette comparaison est faite dans le cadre assez spécifique de la communication vers un casque VR, et il est difficile de dire si ces résultats seraient les mêmes dans un environnement plus conventionnel.

Nous n'avons pas non plus de comparaisons directes entre MCAB et MH-RSAC. Or, MCAB avait des résultats comparables à MCAA, et il se pourrait que dans ce contexte, il ait de meilleurs résultats.

Malgré tout, on peut dire avec une grande certitude que sur la mesure du débit, les stratégies MLO sont bien meilleures que les stratégies non-MLO, notamment SLCI, qui avait des résultats comparables à MCAA dans la référence [5], alors que dans la référence [2], cette différence était bien plus claire.

4.3 Analyse de stratégies individuelles

4.3.1 MCAA

MCAA montre de bons résultats dans toutes les références où il est étudié. On notera qu'il est notamment typiquement plus apte que MCAB et SLCI pour répartir équitablement les flux sur les différentes interfaces de l'AP. Ce qui peut être intéressant pour réduire le taux de congestion mais qui diminue un peu sa capacité à tirer profit des interfaces plus rapides.

4.3.2 MCAB

MCAB fournit aussi de bons résultats dans les références étudiées. Elle est notamment meilleure que MCAA dans les environnements qui contiennent plus de MLO. Donc, dans la mesure où la technologie MLO est probablement amenée à se démocratiser dans le futur, elle semble assez intéressante même si elle fournirait dans l'immédiat des résultats moins bons que MCAA.

4.3.3 MH-RSAC

Cette stratégie a été analysée dans le contexte assez spécifique de la transmission de données dans le cadre de la VR. Il est donc difficile de donner une opinion plus générale sur sa capacité à fonctionner dans des cadres plus communs.

On peut noter cependant qu'elle fournissait de meilleurs résultats que toutes les autres stratégies auxquelles elle a été comparée.

4.3.4 MABs

Les résultats de la stratégie MABs dépasse de loin l'approche statique, indépendamment de la densité du réseau ou des exigences de trafic. Cette approche conduit à des solutions plus performantes et moins variables entre différents scénarios. De plus, elle permet d'atteindre des performances équivalentes, voire meilleures, avec moins d'APs pour un nombre donné de stations. L'algorithme de Thompson sampling se révèle efficace même dans des scénarios adverses, renforçant ainsi la pertinence de l'approche basée sur les MABs pour résoudre les problèmes de configuration de réseau dans les WLANs d'entreprise, grâce à son adaptabilité et ses performances supérieures par rapport aux approches statiques.

5 Retour sur notre expérience

5.1 Ce qu'on a appris

Lors de ce projet, nous avons eu l'opportunité de nous familiariser davantage avec le monde de la recherche, notamment via les différentes études que nous avons analysées et grâce aux retours de nos professeurs référents. Nous avons appris à mieux lire et comprendre des articles scientifiques, ainsi qu'à mener des recherches comparatives sur un sujet donné en utilisant ces articles.

Nous avons aussi appris beaucoup de choses dans le domaine des réseaux, notamment sur les différentes échelles de séparation des flux de données telles que les interfaces et les canaux, ainsi que sur la manière dont elles étaient utilisées pour mieux organiser le réseau.

Enfin, nous avons pu développer nos compétences de synthèse et de rédaction en cherchant à expliquer clairement les résultats de notre recherche. Une fois de plus, les retours de nos professeurs référents ont été très instructifs et nous ont permis de prendre du recul sur notre travail et de trouver les points sur lesquels notre synthèse pouvait être améliorée.

Ce projet de recherche nous a permis d'explorer une part importante du travail d'informaticien : la recherche des solutions existantes dans le monde pour résoudre des problèmes spécifiques. Nous avons rarement l'occasion de mener un travail de recherche approfondi sur une technologie dans nos autres projets, qui

sont généralement axés sur l’implémentation plutôt que sur l’acquisition approfondie des connaissances dans un domaine particulier. Il était intéressant de pouvoir expérimenter cette facette de notre travail futur, en bénéficiant des retours de professionnels pour nous guider.

5.2 Les points négatifs

Nos compétences limitées dans le domaine des réseaux ont compliqué la tâche, surtout au début du projet. Il était difficile pour nous de bien comprendre les aspects plus techniques de certains articles dans les premières étapes du projet. Tout au long de notre travail, alors que nous pensions avoir une meilleure compréhension du sujet, nous avons dû revenir à plusieurs reprises sur des parties d’analyses déjà rédigées, car nous nous sommes rendu compte que nous avions mal interprété certaines parties des documents, ce qui a pu être frustrant à certains moments. Cependant, c’est grâce à ces erreurs et corrections que nous avons réussi à mieux comprendre les systèmes décrits.

Toutes les études que nous avons évaluées étaient entièrement en anglais, ce qui n’a pas facilité les problèmes de compréhension, d’autant plus que le langage utilisé était plutôt technique. Cela a rendu notre travail plus difficile et plus lent. Cependant, ces difficultés nous ont permis de développer notre vocabulaire technique.

6 Conclusion

En conclusion, ce projet d’ouverture à la recherche nous a permis d’explorer en les différentes stratégies de répartition des flux dans les réseaux sans fil, avec un accent particulier sur les approches basées sur le MLO. À travers l’analyse de différentes études et l’examen des performances de chaque stratégie, fournir une analyse des avantages et des limitations de chaque approche.

Nous avons constaté que les stratégies basées sur l’apprentissage machine, telles que MABs et MH-RSAC, offrent des perspectives prometteuses pour améliorer l’efficacité des réseaux sans fil. Ces approches adaptatives démontrent une capacité à ajuster dynamiquement leurs politiques de répartition en fonction des conditions du réseau, contribuant ainsi à une utilisation plus efficace des ressources disponibles.

En conclusion, ce projet nous a permis d’approfondir nos connaissances en matière de réseaux sans fil et d’explorer les possibilités offertes par les approches basées sur le MLO. Grâce à cet exercice de recherche et en analyse critique, nous sommes mieux préparés à appréhender les défis évolutifs et les opportunités naissantes au sein de ce secteur en constante transformation.

Références

- [1] Sergio Barrachina-Muñoz, Francesc Wilhelmi, and Boris Bellalta. Online primary channel selection for dynamic channel bonding in high-density wlans. *IEEE Wireless Communications Letters*, 2019. https://repositori.upf.edu/bitstream/handle/10230/43898/barrachina_ieee_onlin.pdf.
- [2] Pedro Enrique Iturria-Rivera, Marcel Chenier, Bernard Herscovici, Burak Kantarci, and Melike Erol-Kantarci. Rl meets multi-link operation in ieee 802.11be : Multi-headed recurrent soft-actor critic-based traffic allocation. March 2023. <https://arxiv.org/pdf/2303.08959.pdf>.
- [3] Álvaro López-Raventós and Boris Bellalta. Concurrent decentralized channel allocation and access point selection using multi-armed bandits in multi bss wlans. *Computer Networks*, 180, 2020. <https://www.sciencedirect.com/science/article/abs/pii/S1389128619314823>.
- [4] Álvaro López-Raventós and Boris Bellalta. Dynamic traffic allocation in ieee 802.11be multi-link wlans. February 2022. <https://arxiv.org/pdf/2202.12614.pdf>.
- [5] Álvaro López-Raventós and Boris Bellalta. Multi-link operation in ieee 802.11be wlans. *IEEE Wireless Communications Letters*, April 2022. <https://arxiv.org/pdf/2201.07499.pdf>.
- [6] Álvaro López-Raventós and Boris Bellalta. Concurrent decentralized channel allocation and access point selection using multi-armed bandits in multi bss wlans. *Wireless Networking (WN) Research Group*, 2023. <https://arxiv.org/pdf/2006.03350.pdf>.