



# Modèles de Prédiction basés sur des Données pour la Configuration des Réseaux Sans-fil

Samir Si-Mohammed, Fabrice Théoleyre

## ► To cite this version:

Samir Si-Mohammed, Fabrice Théoleyre. Modèles de Prédiction basés sur des Données pour la Configuration des Réseaux Sans-fil. CORES 2025 - 10èmes Rencontres Francophones sur la Conception de Protocoles, l'Evaluation de Performances et l'Expérimentation des Réseaux de Communication, Jun 2025, Saint Valery-sur-Somme, France. hal-05028694

**HAL Id: hal-05028694**

**<https://hal.science/hal-05028694v1>**

Submitted on 10 Apr 2025

**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.

# *Modèles de Prédiction basés sur des Données pour la Configuration des Réseaux Sans-fil*

Samir Si-Mohammed<sup>1</sup> et Fabrice Théoleyre<sup>1</sup>

<sup>1</sup> CNRS/Université de Strasbourg, France

---

Les réseaux sans-fil sont largement utilisés pour un nombre considérables d'applications allant des bâtiments intelligents aux chaînes de production. Toutefois, les performances de ces réseaux dépendent fortement de leur configuration. De plus, les réseaux sans-fil sont bien souvent hétérogènes : appliquer la même configuration à tous les clients et points d'accès est dans un tel cas une approche sous-optimale. Les jumeaux numériques représentent selon nous un outil efficace pour permettre des réseaux autonomes pouvant se reconfigurer automatiquement en fonction de l'environnement radio. À cette fin, les jumeaux numériques des réseaux doivent être capables d'intégrer cette hétérogénéité dans leurs modèles et de capturer l'impact d'une configuration sur la performance d'une liaison radio donnée. Nous proposons donc ici un modèle de prédiction par lien, capable de prédire le taux de réception de paquets attendu pour une configuration MAC donnée. Notre évaluation expérimentale sur un réseau 802.15.4 démontre la pertinence d'une méthode de prédiction basée sur les données pour capturer les spécificités des liens.

**Mots-clefs :** Internet des Objets, Expérimentation, Apprentissage Automatique, Évaluation de Performance.

---

## 1 Introduction

Les réseaux sans-fil jouent un rôle clé dans des domaines comme la santé ou les bâtiments intelligents. Dans l'Industrie 4.0, ils facilitent le déploiement de l'IoT et l'analyse des données en temps réel, permettant par exemple la maintenance prédictive. Face à la complexité croissante des réseaux, des outils et connaissances fines sont nécessaires à leur déploiement, et surtout à leur configuration. Les Jumeaux Numériques (JN) émergent comme une solution clé, offrant une réplique virtuelle des réseaux qui évolue en fonction des données en temps réel. Ils permettent d'explorer virtuellement diverses configurations sans les tester physiquement pour ne pas perturber le fonctionnement du réseau. Un défi majeur est de prédire le comportement d'une configuration inconnue à partir de données observées, par exemple pour les réseaux IEEE 802.15.4 où les paramètres MAC influencent fortement les performances. La généralisation des JNs consiste donc à prédire les valeurs des indicateurs de performance (KPIs) sans tester préalablement les configurations correspondantes.

De nombreux travaux ont analysé les performances réseau via des testbeds et optimisé les paramètres MAC avec des techniques d'apprentissage automatique pour améliorer le débit, la latence et l'efficacité énergétique. Contrairement aux approches précédentes, nous proposons ici un modèle prédictif distinct pour chaque lien du réseau sans-fil, tenant compte des spécificités de chaque lien. Cette granularité améliore significativement la précision des prédictions. Nos contributions sont d'une part une analyse approfondie d'un déploiement expérimental, mettant en évidence : (i) L'hétérogénéité des liens sans fil, (ii) l'impact des configurations sur les performances et (iii) l'importance de considérer chaque lien individuellement. D'autre part, nous proposons une méthode de prédiction peu coûteuse basée sur les données pour estimer l'impact de configurations inexplorées, démontrant qu'un modèle spécifique par lien réduit fortement les erreurs de prédiction.

## 2 État de l'art

Plusieurs travaux se sont intéressés à l'évaluation des performances des réseaux IoT en utilisant différentes approches, notamment à travers des plateformes expérimentales. Par exemple, [Nin21] étudie RPL

sur IEEE 802.15.4 en analysant la latence et le débit en fonction des tailles de messages et des fréquences de transmission. L'optimisation des paramètres MAC quant à elle a toujours suscité un vif intérêt. Les auteurs de [Mou20] utilisent des perceptrons multicouches, des forêts aléatoires, k-NN et des arbres de décision pour identifier les valeurs optimales des paramètres MAC, visant à minimiser le délai de bout en bout. De même, dans [Bil17], les auteurs proposent une approche basée sur les réseaux de neurones artificiels pour estimer les paramètres MAC optimaux de l'IEEE 802.15.4. Leur modèle prédit, par simulation, les configurations minimisant le délai de transmission de bout en bout. Les méthodes basées sur les données et les jumeaux numériques suscitent un intérêt croissant pour leur capacité à modéliser le comportement des réseaux et à explorer des scénarios "what-if". Par exemple, [Sam24] emploie le concept de jumeaux numériques afin de modéliser de façon plus précise la consommation énergétique des réseaux IoT à l'aide de simples régressions linéaires. Ces travaux illustrent le potentiel de l'IA et des méthodes statistiques pour fournir des prédictions fiables et exploitables au niveau d'un réseau. En particulier, notre travail s'intéresse à la prédiction de l'impact de configurations MAC, et se distingue par une approche plus granulaire : au lieu d'un modèle prédictif unique pour l'ensemble du réseau, nous proposons un modèle spécifique à chaque lien, permettant d'intégrer leurs caractéristiques propres.

### 3 Exemple de Déploiement

Nous avons mené une expérimentation de 20 heures sur le site de Grenoble de SLICES-FR, en adoptant une topologie en étoile où une passerelle collecte les paquets de 8 nœuds M3. Ce scénario reproduit un environnement de bâtiment intelligent, avec des nœuds répartis le long d'un couloir à différentes distances de la passerelle. Il convient de noter que le bâtiment héberge également d'autres réseaux (e.g., Wi-Fi) pouvant engendrer des interférences. L'expérimentation repose sur le système d'exploitation Contiki-NG exécutant la pile protocolaire IEEE 802.15.4. Chaque nœud génère un trafic à débit constant (CBR) vers la passerelle à raison d'un paquet par seconde. La configuration MAC est mise à jour toutes les 100 secondes pour l'ensemble des nœuds, et chaque nœud peut adopter une configuration MAC différente.

Nous nous intéressons aux paramètres suivant dans le standard :

- `macMinBE` et `macMaxBE` contrôlent le backoff, influençant le temps d'attente avant retransmission ;
- `macMaxCSMABackoffs` limite le nombre de tentatives de transmissions avant un abandon ;
- `aMaxFrameRetries` fixe le nombre maximal de retransmissions en cas d'échec.

Comme indicateur clé de performance, nous nous focalisons sur le nombre de (re)transmissions nécessaires pour qu'un paquet soit délivré d'un nœud à la passerelle. Plus précisément, pour chaque lien composant le réseau (*nœud*  $\rightarrow$  *gateway*), nous collectons deux séries temporelles :

1. **Configurations** :  $C_l = [C_{l,1}, \dots, C_{l,k}]$ , où  $C_{l,i}$  représente la configuration utilisée par le lien  $l$  durant le  $i^{me}$  intervalle de temps. L'ensemble des configurations possibles est noté  $\mathcal{C}$ .

2. **Taux de Réception des Paquets (PRR)** :  $PRR_l = [PRR_{l,1}, \dots, PRR_{l,k}]$ , où  $PRR_{l,i}$  correspond au taux moyen de réception des paquets pour le lien  $l$  durant l'intervalle de temps  $i$ . Le PRR est défini comme le rapport entre le nombre de paquets reçus par la passerelle et le nombre de paquets transmis sur le lien  $l$ .

La Figure 1 illustre le taux de réception obtenu (PRR) par chaque nœud pour différentes configurations communes à plusieurs nœuds. Pour simplifier, chaque lien est désigné par son émetteur (i.e., nœud) dans la suite. On peut remarquer à travers la table qu'une même configuration (e.g., (6, 9, 3, 3)) peut être optimale pour certains liens (e.g., m3-133) et inefficace pour d'autres (e.g., m3-166). Ce phénomène s'explique par les erreurs au niveau PHY, influençant la fiabilité du lien. De plus, le PRR évolue de manière non uniforme. Par exemple, m3-159 affiche un PRR élevé pour (6, 7, 6, 6), mais un PRR faible pour une configuration voisine, phénomène inverse pour m3-163. Ces observations soulignent la nécessité d'une stratégie spécifique à chaque lien : un modèle doit associer le PRR attendu à chaque lien et configuration, ce que nous abordons dans la section suivante via une approche basée sur les données.

### 4 Prédiction basée sur les données

L'objectif de l'approche proposée est de concevoir un modèle prédictif capable d'estimer l'impact d'une configuration ( $C_{l,i}$ ) inconnue sur la performance d'un nœud  $PRR_{l,i}$ , en se basant sur des observations passées.

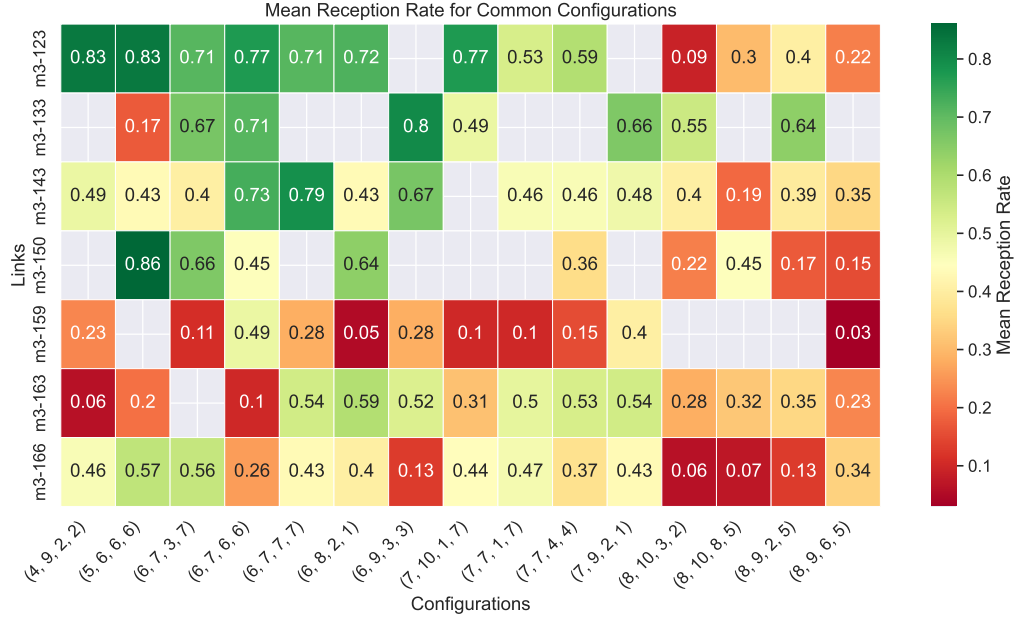


FIGURE 1 : Taux de réception moyen pour des configurations communes sur différentes liaisons.

#### 4.1 Préparation des données

Pour chaque lien et pour chaque configuration, nous calculons le **Taux de Réception Moyen** ( $\overline{PRR}_{l,c}$ ), défini comme le taux moyen de réception des paquets pour le lien  $l$  sous la configuration  $c$ , calculé comme suit :

$$\overline{PRR}_{l,c} = \frac{1}{|\{i \in \mathbb{N} \mid C_{l,i} = c\}|} \sum_{i \in \mathbb{N} \mid C_{l,i} = c} PRR_{l,i} \quad (1)$$

Ensuite, les entrées du modèle de prédictions sont les paramètres de configuration ( $c \in C$ ) pour chaque lien  $l$ , tandis que la sortie du modèle de prédiction est le taux de réception moyen correspondant ( $\overline{PRR}_{l,c}$ ).

#### 4.2 Entraînement

Nous privilégions des modèles de prédiction légers afin de limiter la complexité computationnelle, essentielle pour un jumeau numérique efficace. Chaque lien nécessitant une exécution continue du modèle, nous optons pour des méthodes de régression classiques relativement simples (arbres de décision, SVR, etc.) pour estimer le taux de réception moyen d'une configuration donnée.

Nous explorons trois approches pour entraîner nos modèles de prédiction : (i) la **régression globale** agrège les données de tous les liens pour construire un modèle unique généralisant les performances du réseau. Cette approche capture des tendances générales mais peut être inefficace dans des environnements hétérogènes, (ii) La **régression spécifique à un lien** entraîne un modèle distinct pour chaque lien, offrant des prédictions plus précises mais nécessitant davantage de données et de ressources, et (iii) l'**approche k-NN spécifique à un lien** prédit les performances en s'appuyant sur les configurations les plus proches via un algorithme de clustering. Chaque méthode présente un compromis entre généralisation et précision, que nous souhaitons quantifier.

### 5 Application

Nous comparons les trois approches décrites précédemment en réalisant six fois des expériences de 20 heures pour renforcer la robustesse de l'évaluation. La suite des configurations reste inchangée, mais des expérimentations concurrentes ou du trafic Wi-Fi peuvent introduire des variations. L'exactitude des prédictions est mesurée via l'erreur quadratique moyenne (MSE) :

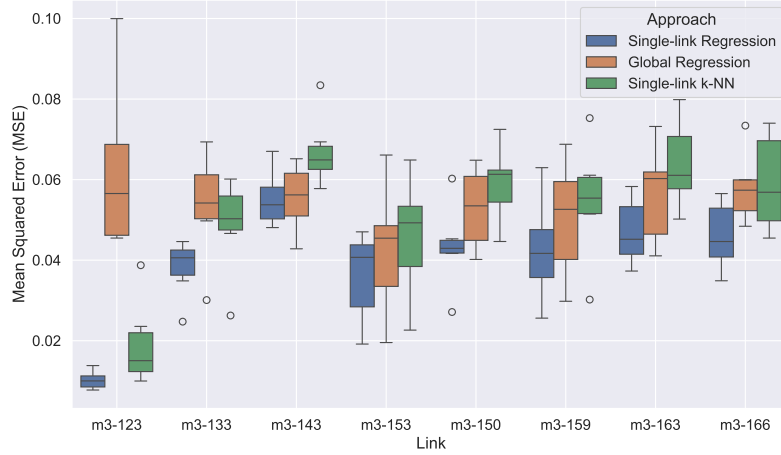


FIGURE 2 : Comparaison des trois approches sur le déploiement donné.

$$MSE(l) = \frac{1}{n} \sum_{i=1}^n \left( PRR_{l,i} - \widehat{PRR}_{l,i} \right)^2 \quad (2)$$

où  $PRR_{l,i}$  est la valeur observée du taux de réception des paquets pour le lien  $l$  et l'intervalle  $i$ , tandis que  $\widehat{PRR}_{l,i}$  est la valeur prédite.

La Figure 2 montre l'erreur MSE sur les données de test (70 % des données utilisées pour l'entraînement, 30 % pour le test). Le modèle de *régression par lien* donne les meilleures performances car il exploite uniquement les mesures de chaque lien, évitant la perte d'informations spécifiques. En revanche, la *régression globale* agrège des données hétérogènes, ce qui dégrade la précision, notamment pour des liens atypiques comme *m3-123*. L'approche *k-NN par lien* n'arrive pas à s'adapter efficacement au lien : *m3-143* est par exemple mal prédit.

Enfin, les valeurs de MSE obtenues restent faibles ( $\leq 0.08$ ), prouvant que la méthode proposée permet de prédire efficacement le taux de réception moyen d'une configuration inconnue.

## 6 Conclusion

Dans ce travail, nous avons développé des modèles prédictifs pour évaluer l'impact des configurations MAC inexplorées dans les réseaux IEEE 802.15.4, en analysant leur effet sur le nombre de (re)transmissions. Trois approches ont été comparées : régression globale, régression par lien unique et k-NN par lien unique. Les résultats montrent que le modèle de régression par lien unique offre la meilleure précision pour prédire la qualité du lien pour des configurations inconnues. Ces modèles peuvent être intégrés dans un jumeau numérique pour réseaux sans fil. Dans les travaux futurs, nous comptons ajouter d'autres métriques de performance et étudier comment ces modèles peuvent être intégrés à des jumeaux numériques de réseaux.

**Partage de code :** Afin de permettre la reproductibilité, nous fournissons le code source/ensemble de données dans le lien suivant : <https://github.com/SamirSim/AINA-2025-Config-Pred>

## Références

- [Bil17] Bilal Al-Kaseem et al. A new intelligent approach for optimizing 6lowpan mac layer parameters. *IEEE Access*, 2017.
- [Mou20] Moussa Aboubakar et al. An efficient and adaptive configuration of ieee 802.15.4 mac for communication delay optimisation. In *IEEE NoF*, 2020.
- [Nin21] Nina Santi et al. Automated and reproducible application traces generation for iot applications. In *Proceedings of the 17th ACM Symposium on QoS and Security for Wireless and Mobile Networks*, 2021.
- [Sam24] Samir Si-Mohammed et al. Ns+ ndt : Smart integration of network simulation in network digital twin, application to iot networks. *Future Generation Computer Systems*, 2024.