Sécurité des réseaux 5G/6G par l'IA : Prédiction de la durée des attaques pour une mitigation intelligente

par

Mohamed Anis SAKKA

MÉMOIRE PRÉSENTÉ À L'ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
COMME EXIGENCE PARTIELLE À L'OBTENTION DE LA MAÎTRISE
AVEC MÉMOIRE EN GÉNIE DES TECHNOLOGIES DE L'INFORMATION
M. Sc. A.

MONTRÉAL, LE "15 MAI 2026"

ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
UNIVERSITÉ DU QUÉBEC

PRÉSENTATION DU JURY

CE MÉMOIRE A ÉTÉ ÉVALUÉ

PAR UN JURY COMPOSÉ DE:

M. Rami Langar, directeur de mémoire
Département de génie logiciel et TI à l'École de technologie supérieure

M. Wael Jaafar, codirecteur
Département de génie logiciel et TI à l'École de technologie supérieure

Mme. Bassant Selim, présidente du jury
Département de génie des systèmes à l'École de technologie supérieure

M. Kaiwen Zhang, membre du jury
Département de génie logiciel et TI à l'École de technologie supérieure

IL A FAIT L'OBJET D'UNE SOUTENANCE DEVANT JURY ET PUBLIC

LE "13 MAI 2026"

À L'ÉCOLE DE TECHNOLOGIE SUPÉRIEURE

REMERCIEMENTS

Je tiens à exprimer ma sincère gratitude à mon directeur de recherche, le professeur Rami Langar, pour son encadrement exceptionnel, son soutien constant et ses précieux conseils tout au long de ce projet de maîtrise. Sa vision stratégique, son expertise approfondie et sa disponibilité ont été déterminantes dans la réalisation de cette thèse. Ses orientations éclairées et son accompagnement personnalisé ont grandement contribué à la qualité de ce travail.

Je souhaite adresser mes remerciements les plus chaleureux à mon co-directeur, le professeur Wael Jaafar, pour son engagement remarquable, ses retours constructifs et son soutien indéfectible. Sa rigueur scientifique, ses suggestions pertinentes et son mentorat ont significativement enrichi mon parcours de recherche et ont été essentiels à l'avancement de ces travaux.

Je remercie chaleureusement l'ensemble des membres du laboratoire LASI pour avoir créé un environnement de recherche stimulant et collaboratif. L'esprit d'équipe et l'atmosphère bienveillante qui règnent au sein du laboratoire ont rendu cette expérience enrichissante.

Je suis profondément reconnaissant envers ma famille pour son amour inconditionnel et son soutien constant. À mes parents, merci pour vos sacrifices, votre confiance absolue et votre encouragement sans faille tout au long de ce parcours. À mes frères et sœurs, votre présence réconfortante et vos encouragements ont été une source de motivation permanente.

À ma fiancée, merci pour ton soutien indéfectible, ta patience et ta présence à mes côtés dans les moments tant difficiles que joyeux. Ta compréhension et ton encouragement ont été des piliers essentiels dans l'aboutissement de ce projet.

Enfin, je tiens à remercier tous mes amis, tant ceux restés dans mon pays d'origine qui m'ont soutenu à distance, que ceux rencontrés ici au Canada, dont l'amitié et le soutien ont illuminé ce parcours.

Cette recherche a été réalisée grâce au soutien financier du programme IDEaS du Ministère de la Défense nationale du Canada.

Sécurité des réseaux 5G/6G par l'IA : Prédiction de la durée des attaques pour une mitigation intelligente

Mohamed Anis SAKKA

RÉSUMÉ

La prolifération rapide des architectures réseau 5G et O-RAN a considérablement élargi le paysage des cybermenaces, créant des défis de sécurité critiques qui exigent des solutions innovantes dépassant les systèmes de détection d'intrusion conventionnels. Cette recherche introduit un cadre innovant pour la prédiction de la durée des cyberattaques qui transforme la cybersécurité d'une détection réactive en une gestion proactive des menaces, adressant ainsi le manque fondamental en intelligence temporelle des menaces qui limite actuellement les stratégies de mitigation efficaces dans les réseaux de nouvelle génération. Le cadre repose sur trois contributions principales : une architecture neuronale avancée basée sur les Transformers spécialement conçue pour capturer les motifs temporels complexes dans les séquences d'attaques, démontrant des améliorations significatives de performance par rapport aux modèles récurrents traditionnels ; un paradigme d'apprentissage fédéré préservant la confidentialité avec sélection adaptive des clients et optimisation dynamique de l'entraînement, permettant une allocation intelligente des ressources à travers les nœuds réseau distribués tout en maintenant la confidentialité des données et l'efficacité computationnelle ; et le nouveau cadre hybride FML-AD qui combine de manière synergique l'apprentissage fédéré avec des techniques de méta-apprentissage pour surmonter les barrières d'adaptation au domaine entre les environnements d'entraînement et le déploiement opérationnel réel. Le mécanisme de sélection adaptive des clients priorise stratégiquement les nœuds sous-performants basé sur des métriques de validation en temps réel, tandis que l'optimisation dynamique de l'entraînement ajuste l'allocation des époques et le traitement par lots selon les capacités des clients et les distributions de données. La validation expérimentale sur un testbed 5G O-RAN a confirmé la haute précision du cadre FML-AD, avec une erreur moyenne de seulement 2.0 secondes dans la prédiction de la durée résiduelle des attaques. Cette performance permet d'optimiser l'allocation des ressources de sécurité et de planifier les stratégies de mitigation en temps réel, tout en maintenant une qualité de service optimale pour les utilisateurs légitimes et en préservant la confidentialité des données across les infrastructures réseau hétérogènes. Ces résultats établissent les bases d'une nouvelle génération de systèmes de cybersécurité adaptatifs, capables d'anticiper proactivement les menaces et d'ajuster dynamiquement les mécanismes de défense. Les perspectives de recherche incluent la gestion d'attaques multiples, l'optimisation énergétique des modèles et le développement de moteurs d'inférence temps réel pour environnements industriels, renforçant ainsi la résilience des infrastructures critiques face aux cybermenaces évolutives dans l'écosystème 5G et au-delà.

Mots-clés: Prédiction de Durée des Cyberattaques, Apprentissage Fédéré, Réseaux de Transformers, Sélection Adaptive des Clients, Sécurité 5G, Architecture O-RAN, Méta-Apprentissage, Sécurité Réseau, Mitigation d'Attaques, IA Préservant la Confidentialité

AI-Driven Security for 5G/6G Networks : Predicting Attack Duration for Intelligent Mitigation

Mohamed Anis SAKKA

ABSTRACT

The rapid proliferation of 5G and O-RAN network architectures has dramatically expanded the cyber threat landscape, creating critical security challenges that demand innovative solutions beyond conventional intrusion detection systems. This research introduces an innovative framework for cyberattack duration prediction that transforms cybersecurity from reactive detection to proactive threat management, addressing the fundamental gap in temporal threat intelligence that currently limits effective mitigation strategies in next-generation networks. The framework is built upon three key contributions : an advanced Transformer-based neural architecture specifically engineered for capturing complex temporal patterns in attack sequences, demonstrating significant performance improvements over traditional recurrent models ; a privacy-preserving federated learning paradigm with adaptive client selection and dynamic training optimization that enables intelligent resource allocation across distributed network nodes while maintaining data confidentiality and computational efficiency ; and the novel FML-AD hybrid framework that synergistically combines federated learning with meta-learning techniques to overcome domain adaptation barriers between training environments and real-world operational deployment. The adaptive client selection mechanism strategically prioritizes underperforming nodes based on real-time validation metrics, while the dynamic training optimization adjusts epoch allocation and batch processing according to client capabilities and data distributions. Extensive experimental validation conducted on multiple benchmark datasets and a sophisticated 5G O-RAN testbed demonstrates exceptional prediction accuracy for remaining attack duration, facilitating optimized resource allocation, intelligent mitigation planning, and guaranteed quality of service for legitimate users while maintaining robust privacy safeguards across heterogeneous network infrastructures. This research establishes the foundational principles for next-generation adaptive cybersecurity systems capable of proactive threat anticipation and dynamic defense optimization, with promising future research directions encompassing multi-attack scenario handling, energy-aware model optimization, and real-time inference engines for large-scale industrial deployments, ultimately contributing to enhanced resilience of critical digital infrastructures against evolving cyber threats in the 5G era and beyond.

Keywords: Cyberattack Duration Prediction, Federated Learning, Transformer Networks, Adaptive Client Selection, 5G Security, O-RAN Architecture, Meta-Learning, Network Security, Attack Mitigation, Privacy-Preserving AI

TABLE DES MATIÈRES

Page

INTRODUCTION 1

CHAPITRE 1 CONTEXTE GÉNÉRAL 3
1.1 Introduction 3
1.2 Formulation du Problème 5
1.3 Le Rôle de la Prédiction de la Durée des Attaques dans l'Atténuation Proactive 7
1.3.1 Mécanisme des Attaques DDoS en 5G O-RAN 7
1.3.2 Limites des Stratégies d'Atténuation Actuelles 8
1.3.3 Atténuation Adaptative par la Prédiction de la Durée des Attaques 8
1.4 Solution Proposée et Contributions 9
1.5 Travaux Connexes 12
1.5.1 Prédiction de la Durée des Attaques et Approches Centralisées 12
1.5.2 Apprentissage Distribué et Respectueux de la Vie Privée en 5G/O-RAN .. 15
1.5.3 Lacunes de Recherche et Positionnement 16
1.6 Conclusion 18

CHAPITRE 2 PREDICTING CYBERATTACK DURATION IN NEXT GENERATION NETWORKS : A NOVEL TRANSFORMER-BASED APPROACH 19
2.1 Methodology 19
2.1.1 Data Preprocessing : Attack_id and Remaining_time features creation 20
2.1.2 Feature Selection and Data Normalization 21
2.1.3 Data Splitting and Sequence Creation 22
2.1.4 Proposed Transformer-based Model architecture 24
2.2 Simulation Results 24
2.2.1 Hyperparameter Auto-Tuning 24
2.2.2 Models' Performance Evaluation 26
2.2.3 Testing results on DoS attacks scenario 28
2.3 Conclusion 29

CHAPITRE 3 FML-AD : A FEDERATED META-LEARNING FRAMEWORK FOR CYBERATTACK DURATION PREDICTION IN 5G O-RAN ENVIRONMENTS 31
3.1 Proposed Attack Duration Prediction Method : FML-AD 31
3.1.1 Phase 1 : FLAD Offline Training 33
3.1.1.1 Data Preprocessing and Temporal Modeling Pipeline 33
3.1.1.2 Data Partitioning for FL Scenarios 34
3.1.1.3 FLAD Framework Implementation 36
3.1.2 Phase 2 : 5G O-RAN Dataset Generation and Initial FLAD Prediction ... 38

3.1.3 Phase 3 : Improving Attack Duration Prediction using a Novel Meta-Model 38
3.1.4 Phase 4 : FML-AD Online Prediction 39
3.2 Experimental Results 40
3.3 Conclusion 45

CONCLUSION ET RECOMMANDATIONS 47

BIBLIOGRAPHIE 49

LISTE DES TABLEAUX

Page

Tableau 2.1 Unique Attack ID Counts and Maximum Remaining Time (Total Attack Duration) per Attack Category 20

Tableau 2.2 Assumed Attack Duration Thresholds 21

Tableau 2.3 Hyperparameters for RNN, LSTM, TRANSFORMER, and XGBOOST Models 25

Tableau 2.4 Mean Absolute Error (MAE) for Train and Validation Sets 26

Tableau 3.1 Attack flows data distribution 35

Tableau 3.2 Hyperparameters configuration 37

Tableau 3.3 Mean Absolute Error (MAE) and Execution Time for Training and Validation 39

Tableau 3.4 Performance Comparison of Federated Learning Strategies (Global Model Performance) 40

Tableau 3.5 Meta-Model Performance Comparison on Processing 1st Stage Predictions 43

LISTE DES FIGURES

Page

Figure 1.1 Comparaison entre les stratégies d'atténuation statiques (V1) et adaptatives (V2) basées sur la prédiction de la durée des attaques 9

Figure 2.1 Evolution of Error during Training for the studied DL/ML models 27

Figure 2.2 Evolution of Error during Validation for the studied DL/ML models 27

Figure 2.3 Transformer-based Model Performance : Actual Duration vs. Predicted Duration 28

Figure 3.1 FML-AD workflow illustrating the four-phase methodology : (1) FLAD Offline Training, (2) 5G-O-RAN Dataset Generation and Initial FLAD Prediction, (3) Improving Attack Duration Prediction using a Novel Meta-Model (Offline), (4) FML-AD Online Prediction 32

Figure 3.2 Attack Duration Distribution on CICIoT Dataset 35

Figure 3.3 Convergence of centralized methods (Training and validation) 40

Figure 3.4 Convergence of FL methods (Different data distribution scenarios) 41

Figure 3.5 Performance comparison between FLAD and FML-AD predictions across temporal thresholds during 120-second DDoS TCP attack on 5G O-RAN testbed 44

LISTE DES ABRÉVIATIONS, SIGLES ET ACRONYMES

5G	Fifth Generation Mobile Networks
AI	Artificial Intelligence
ANOVA	Analysis of Variance
ASC	Agence Spatiale Canadienne
CICIoT	Canadian Institute for Cybersecurity Internet of Things
CNN	Convolutional Neural Network
CN	Core Network
CU	Central Unit
DDoS	Distributed Denial-of-Service
DL	Deep Learning
DND	Department of National Defence
DoS	Denial-of-Service
DU	Distributed Unit
ETS	École de Technologie Supérieure
FedAvg	Federated Averaging
FedProx	Federated Proximal Term Regularization
FFN	Feed-Forward Network
FL	Federated Learning
FLAD	Federated Learning with Adaptive Client Selection
FML-AD	Federated Meta-Learning for Attack Duration
gNB	gNodeB (5G Base Station)
ICMP	Internet Control Message Protocol
IDM	Intrusion Detection Model
IDS	Intrusion Detection Systems
IDEaS	Innovation for Defence Excellence and Security

IG	Information Gain
IID	Independent and Identically Distributed
IoT	Internet of Things
LASI	Laboratory of Computer Architectures and Systems
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
MEC	Multi-access Edge Computing
ML	Machine Learning
MSE	Mean Squared Error
Near-RT RIC	Near-Real-Time RAN Intelligent Controller
non-IID	Non Independent and Identically Distributed
O-RAN	Open Radio Access Network
QoS	Quality of Service
RAN	Radio Access Network
ReLU	Rectified Linear Unit
RIC	RAN Intelligent Controller
RNN	Recurrent Neural Networks
SA	Standalone
SOC	Security Operations Centers
TCP	Transmission Control Protocol
UDP	User Datagram Protocol
UE	User Equipment
UNSW-NB15	University of New South Wales Network Binary 15
V2X	Vehicle-to-Everything
VoIP	Voice over Internet Protocol
XGBoost	eXtreme Gradient Boosting
xApp	RAN Application

INTRODUCTION

Les réseaux mobiles sont devenus une composante cruciale des communications modernes, fournissant des services essentiels tant pour les particuliers que pour les entreprises. Cependant, avec l'expansion rapide des infrastructures réseau, notamment la 5G, l'exposition aux cybermenaces s'est significativement accrue. Ces réseaux, de par leur nature omniprésente et les services critiques qu'ils supportent, constituent des cibles privilégiées pour diverses cyberattaques pouvant entraîner des interruptions de service, des fuites de données et des pertes financières substantielles pour les opérateurs et utilisateurs.

Alors que la communauté cybersécurité a réalisé des progrès significatifs en détection et classification d'attaques, un aspect crucial de la gestion des menaces reste largement inexploré : prédire combien de temps une attaque va persister dans le réseau. Cette lacune est particulièrement préoccupante car l'estimation précise de la durée des attaques fournit aux administrateurs réseau une intelligence cruciale pour mettre en œuvre des stratégies de mitigation rapides et efficaces. Plus spécifiquement, ces prédictions permettent une meilleure allocation des ressources, minimisent l'impact sur les utilisateurs légitimes et maintiennent la Qualité de Service pendant les incidents de sécurité.

Les approches traditionnelles de cybersécurité échouent souvent à prédire le comportement dynamique des attaques, qui varient considérablement en type et durée. Certaines attaques, comme le Deni de Service (DoS), peuvent épuiser les ressources en peu de temps, tandis que d'autres, telles que les attaques de Reconnaissance ou d'Exploitation, peuvent s'étendre sur des périodes plus longues. La nature diverse de ces attaques rend difficile l'adaptation en temps réel des systèmes statiques ou basés sur des règles prédéfinies.

Ces dernières années, l'Intelligence Artificielle (IA) a émergé comme un outil puissant pour les applications de cybersécurité. En exploitant les techniques d'apprentissage automatique (ML) et d'apprentissage profond (DL), l'IA a transformé les mesures de cybersécurité traditionnelles,

permettant des défenses plus proactives et dynamiques. Les systèmes basés sur l'IA sont capables de détecter des schémas d'attaque complexes, de prédire les menaces futures et d'automatiser les stratégies de mitigation en temps réel. Ces capacités sont particulièrement cruciales pour la défense des réseaux mobiles à grande échelle, hautement susceptibles à un large éventail de cybermenaces.

Malgré ces avancées, la majorité de la recherche et de l'implémentation en cybersécurité basée sur l'IA se concentre principalement sur la détection et la classification d'attaques. Bien que ces tâches soient critiques, un aspect tout aussi important mais souvent négligé est la prédiction de la durée des attaques. Comprendre combien de temps une attaque va durer permet aux administrateurs réseau d'appliquer des stratégies de mitigation plus efficacement. Par exemple, prédire avec précision le temps restant de différentes catégories d'attaques peut aider à allouer les ressources efficacement, garantissant que les utilisateurs légitimes subissent des perturbations minimales, tout en empêchant le réseau d'être submergé.

Cette thèse aborde cette lacune fondamentale en proposant une approche novatrice pour la prédiction de la durée des cyberattaques dans les environnements 5G et O-RAN. Notre travail explore comment l'intelligence temporelle peut transformer les stratégies de sécurité réseau, en se focalisant sur le développement de modèles avancés de prédiction, leur adaptation aux environnements distribués, et leur intégration dans des frameworks de mitigation adaptative.

À travers cette recherche, nous visons à démontrer que la prédiction de durée d'attaque représente un changement de paradigme essentiel : passer d'une sécurité réactive à une sécurité proactive et contextuelle, capable non seulement de détecter les menaces mais aussi d'anticiper leur comportement temporel pour une protection plus efficace et efficiente des infrastructures réseau modernes.

CHAPITRE 1

CONTEXTE GÉNÉRAL

Ce chapitre présente le contexte général de la thèse en introduisant les fondements technologiques et scientifiques qui motivent ce travail. Il examine l'évolution des réseaux mobiles de prochaine génération, en particulier les architectures 5G et O-RAN, et met en lumière les défis émergents en matière de cybersécurité associés à leur ouverture croissante, leur logicialisation et leur nature distribuée.

S'appuyant sur ce contexte, le chapitre formule le problème de prédiction de la durée des cyberattaques et discute de son rôle critique dans l'activation de stratégies d'atténuation plus efficaces et adaptatives. Il présente ensuite les principales orientations de recherche et les contributions de cette thèse, suivies d'une revue exhaustive de la littérature pertinente. Enfin, le chapitre identifie les lacunes de recherche clés qui motivent l'approche proposée.

1.1 Introduction

Les réseaux mobiles sont devenus une composante fondamentale de l'infrastructure numérique moderne, supportant un large spectre de services pour les individus comme pour les organisations. Leur importance s'est considérablement accrue avec le déploiement des réseaux de cinquième génération (5G), qui permettent des cas d'usage avancés tels que le transport autonome, l'automatisation industrielle, la santé intelligente, la connectivité massive de l'Internet des objets (IoT) et d'autres applications sensibles à la latence. Parallèlement, l'émergence des architectures de réseau d'accès radio ouvert (O-RAN) a introduit un nouveau niveau d'ouverture, de flexibilité et de programmabilité dans les écosystèmes de réseaux mobiles. En s'appuyant sur la logicialisation des réseaux, la virtualisation et des interfaces ouvertes désagrégées, les systèmes 5G et O-RAN s'éloignent des infrastructures monolithiques traditionnelles pour évoluer vers des environnements hautement dynamiques, distribués et pilotés par les logiciels Mavoungou, Kaddoum, Taha & Matar (2016).

Bien que ces transformations créent d'importantes opportunités en termes d'interopérabilité, de scalabilité et d'innovation de services, elles introduisent également des défis majeurs en matière de sécurité. L'ouverture des interfaces O-RAN, la virtualisation des fonctions réseau et la nature distribuée des infrastructures 5G élargissent considérablement la surface d'attaque. Par conséquent, les réseaux mobiles de nouvelle génération sont devenus des cibles attractives pour un large éventail de cyberattaques, notamment les attaques par déni de service, les campagnes de reconnaissance, les tentatives d'exploitation, ainsi que les attaques ciblant les composants virtualisés ou définis par logiciel. Ces menaces peuvent entraîner de graves perturbations de service, des violations de données et des pertes opérationnelles et financières substantielles, compte tenu notamment de la nature critique des services concernés.

Dans ce contexte, la cybersécurité est devenue une exigence centrale pour garantir la fiabilité, la résilience et la fiabilité des réseaux mobiles modernes. Des progrès considérables ont été réalisés ces dernières années en matière de détection et de classification des attaques, notamment grâce à l'adoption de techniques d'intelligence artificielle (IA), d'apprentissage automatique (ML) et d'apprentissage profond (DL) Khan & Ghafoor (2024); Amrollahi *et al.* (2020); Banerjee *et al.* (2019). Les systèmes de cybersécurité basés sur l'IA sont de plus en plus capables d'identifier des schémas d'attaque complexes, de réduire les fausses alarmes et de permettre une réponse aux incidents plus rapide. Ils ont amélioré les performances des systèmes de détection d'intrusion (IDS), des plateformes de surveillance de la sécurité et des pipelines d'atténuation automatisée, en faisant des composants essentiels des architectures de sécurité de prochaine génération.

Cependant, malgré ces avancées, la recherche en cybersécurité existante reste largement centrée sur les questions de savoir si une attaque est en cours et de quel type elle est. Une question tout aussi importante reste insuffisamment traitée : combien de temps durera l'attaque ? Dans les contextes opérationnels, la capacité à estimer la durée restante d'une attaque en cours peut fournir une intelligence temporelle précieuse pour la prise de décision en matière de sécurité. Cette information peut contribuer à déterminer combien de temps les mesures d'atténuation doivent rester actives, comment les ressources doivent être allouées pendant une attaque, et à quel moment les conditions normales de service peuvent être rétablies en toute sécurité. Par

conséquent, la prédiction de la durée des cyberattaques représente une étape importante vers des mécanismes de cybersécurité plus proactifs et adaptatifs.

L'importance de ce problème est particulièrement évidente dans les environnements 5G et O-RAN, où les décisions de sécurité affectent directement à la fois la protection du réseau et la qualité de service (QoS). Par exemple, l'isolation du trafic suspect ou compromis dans des tranches dédiées ou des environnements de type sinkhole peut aider à contenir une attaque, mais le maintien de cette isolation plus longtemps que nécessaire peut pénaliser inutilement les utilisateurs légitimes et dégrader les performances du service. Inversement, relâcher le trafic isolé trop tôt peut exposer le réseau à une activité malveillante continue. Ces compromis soulignent la nécessité de modèles prédictifs capables d'estimer l'évolution temporelle des cyberattaques plutôt que de s'appuyer uniquement sur des règles réactives.

Cette thèse est motivée par l'observation que la prédiction de la durée des cyberattaques reste peu explorée, en particulier dans les environnements 5G et O-RAN distribués. Elle étudie comment les techniques avancées de modélisation temporelle et d'apprentissage distribué peuvent être combinées pour fournir une prédiction précise, scalable et respectueuse de la vie privée de la durée des attaques. Plus précisément, ce travail construit un cadre de recherche progressif qui commence par l'apprentissage profond centralisé pour la prédiction temporelle et s'étend vers des approches d'apprentissage fédéré et méta-apprentissage distribués adaptées aux scénarios de déploiement réels en O-RAN.

1.2 Formulation du Problème

Malgré la disponibilité de nombreuses techniques de détection et d'atténuation des attaques, les stratégies de cybersécurité existantes restent principalement réactives. La plupart des mécanismes de défense conventionnels sont déclenchés après qu'une attaque a déjà été identifiée, et ils opèrent généralement sans connaissance explicite de la durée attendue de l'attaque. Les approches courantes telles que le blocage d'adresses IP, la limitation du débit de trafic, la redirection de

flux ou l'isolation temporaire sont donc appliquées sur la base d'observations instantanées plutôt que sur des prévisions temporelles éclairées.

Cette limitation devient particulièrement problématique dans des environnements réseau dynamiques et hétérogènes tels que la 5G et l'O-RAN. Dans ces contextes, les attaques peuvent évoluer rapidement et présenter des comportements temporels très variables selon leur type, leur intensité, leur origine et leur cible. Certaines attaques, comme les rafales de déni de service, peuvent être courtes mais très perturbatrices, tandis que d'autres, comme les campagnes de reconnaissance ou d'exploitation, peuvent persister sur de longues périodes et se dérouler de manière plus progressive. En raison de cette diversité, les règles d'atténuation statiques ou les stratégies basées sur des délais d'expiration fixes sont souvent inadéquates.

Un défi opérationnel clé réside dans la détermination de la durée appropriée des actions d'atténuation. Si les mesures défensives sont levées trop tôt, le réseau peut rester exposé à l'attaque en cours, compromettant ainsi la sécurité et la continuité du service. D'autre part, si l'atténuation est maintenue plus longtemps que nécessaire, les utilisateurs légitimes peuvent subir une dégradation inutile du service, une latence accrue, un accès restreint ou une redirection excessive du trafic. Cela entraîne un compromis inhérent entre la protection de la sécurité et la préservation de la QoS.

Dans cette perspective, le problème central abordé dans cette thèse peut être formulé comme suit : *comment la durée restante d'une cyberattaque en cours peut-elle être prédite avec précision afin de soutenir une atténuation proactive et adaptative dans les environnements 5G et O-RAN ?*

Résoudre ce problème est difficile pour plusieurs raisons. Premièrement, la durée d'une cyberattaque n'est pas une propriété statique mais un phénomène temporel qui dépend de l'évolution séquentielle des schémas de trafic et du comportement de l'attaque. Deuxièmement, les données de trafic réseau sont de haute dimension, hétérogènes et souvent distribuées à travers plusieurs entités, rendant l'analyse centralisée difficile dans les environnements opérationnels. Troisièmement, les modèles entraînés sur des jeux de données de référence peuvent ne pas bien se généraliser aux déploiements réels en raison de la dérive de domaine, où les distributions de

données et les caractéristiques des attaques diffèrent significativement entre les jeux de données de laboratoire et les infrastructures réelles.

Par conséquent, résoudre le problème de prédiction de la durée des cyberattaques requiert bien plus que des techniques standard de classification d'attaques. Cela nécessite des modèles capables d'apprendre des dépendances temporelles à partir de données de trafic séquentielles, de s'adapter à des environnements hétérogènes et d'opérer au sein d'architectures distribuées respectueuses de la vie privée. Cette combinaison d'exigences motive l'utilisation de techniques avancées d'apprentissage profond, d'apprentissage fédéré et de méta-apprentissage dans le présent travail.

1.3 Le Rôle de la Prédiction de la Durée des Attaques dans l'Atténuation Proactive

Cette section examine la problématique des attaques DDoS dans les environnements 5G O-RAN et établit la motivation fondatrice de notre travail. Plus précisément, nous analysons les limites des approches actuelles d'atténuation des attaques DDoS et démontrons comment la prédiction de la durée des attaques permet un glissement fondamental d'une défense réactive vers une gestion de la sécurité proactive et adaptative.

1.3.1 Mécanisme des Attaques DDoS en 5G O-RAN

Les attaques DDoS perturbent la disponibilité des services en submergeant les cibles d'un trafic malveillant provenant de botnets coordonnés. Dans les environnements 5G O-RAN, la logicialisation accrue et les interfaces ouvertes exacerbent cette menace. En effet, les attaquants peuvent utiliser l'usurpation d'adresse IP pour dissimuler leur identité, rendant l'identification des sources extrêmement difficile Mirkovic & Reiher (2004). L'utilisation de botnets peut comprendre des appareils divers, allant des gadgets IoT aux terminaux industriels, qui participent aux attaques à l'insu des utilisateurs légitimes Abubakar *et al.* (2020). Une fois identifiées comme victimes ou sources d'attaque, des stratégies de détection et d'atténuation seront sollicitées pour contenir la menace. Cette combinaison de sources distribuées, d'usurpation d'IP et de

coordination sophistiquée rend les mesures de sécurité réactives traditionnelles insuffisantes, nécessitant des capacités prédictives avancées.

1.3.2 Limites des Stratégies d'Atténuation Actuelles

Les stratégies d'atténuation DDoS existantes sont caractérisées par des compromis fondamentaux entre l'efficacité sécuritaire et la qualité du service, conduisant souvent à des résultats sous-optimaux dans les environnements 5G dynamiques. Le blocage d'adresses IP, bien que simple à mettre en œuvre, est de plus en plus inefficace face aux botnets modernes qui font tourner dynamiquement les adresses IP pour échapper à la détection et aux règles de blocage Mirkovic & Reiher (2004). De plus, les techniques de redirection de trafic, telles que la redirection des flux malveillants vers des tranches sinkhole, peuvent préserver la disponibilité des services centraux Bousalem, Silva, Langar & Cherrier (2022a), mais introduisent une latence significative, augmentent la complexité du routage et imposent une surcharge substantielle à la gestion du réseau. De même, les approches de limitation de débit peuvent contribuer à maintenir la stabilité globale du réseau, mais dégradent souvent indistinctement la qualité de service pour les utilisateurs légitimes dont les schémas de trafic sont capturés par des seuils d'atténuation larges Chang (2002).

Une limitation commune et critique, omniprésente dans toutes ces stratégies, est leur nature intrinsèquement réactive. Elles opèrent sur la base des caractéristiques instantanées de l'attaque sans aucune intelligence temporelle concernant la durée attendue de l'attaque. Par conséquent, les opérateurs réseau se trouvent face à un dilemme : relâcher prématurément les défenses et risquer une nouvelle exposition à une attaque en cours, ou prolonger inutilement des mesures restrictives qui dégradent la QoS.

1.3.3 Atténuation Adaptative par la Prédiction de la Durée des Attaques

La limitation centrale des approches d'atténuation actuelles est leur incapacité à anticiper la durée des attaques, conduisant soit à une sur-isolation, soit à une ré-exposition prématurée.

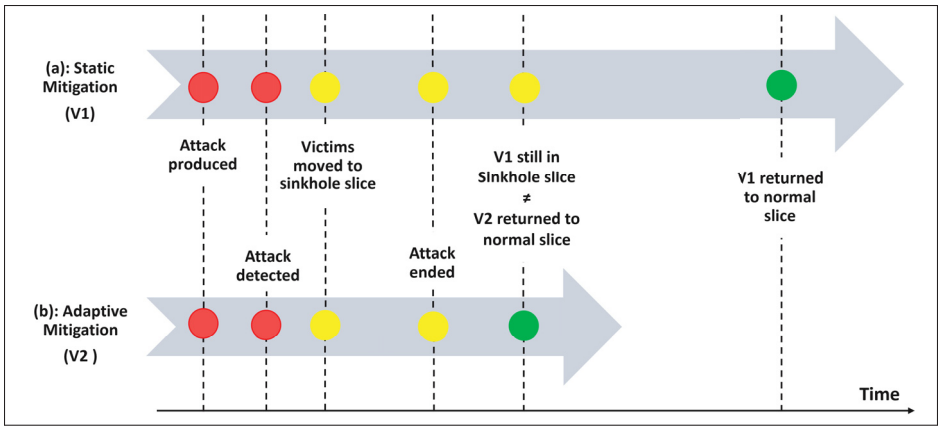


Figure 1.1 Comparaison entre les stratégies d'atténuation statiques (V1) et adaptatives (V2) basées sur la prédiction de la durée des attaques

Comme illustré à la Fig. 1.1, les stratégies statiques comme l'isolation sinkhole maintiennent la Victime 1 (V1) confinée plus longtemps que nécessaire, causant une dégradation prolongée du service. En revanche, une prédiction précise de la durée permet un glissement paradigmatique vers une gestion proactive de la sécurité, comme illustré pour la Victime 2 (V2). Plus précisément, une prévision précise de la durée permet à V2 d'être restaurée en service normal au moment optimal, équilibrant sécurité et QoS. Cette intelligence temporelle transforme l'atténuation selon trois aspects clés : (1) une allocation dynamique des ressources adaptée à la durée attendue de l'attaque, (2) un timing optimal pour activer/désactiver des contre-mesures telles que les transitions sinkhole, et (3) une réduction des dommages collatéraux en minimisant les fenêtres d'atténuation à large spectre.

1.4 Solution Proposée et Contributions

Pour répondre aux limites des stratégies d'atténuation réactives existantes, cette thèse propose un cadre complet de prédiction de la durée des cyberattaques dans les environnements 5G et O-RAN. L'idée principale est d'enrichir les systèmes de cybersécurité d'une intelligence temporelle, leur permettant non seulement de détecter les attaques, mais aussi d'estimer combien de temps elles devraient persister. Ce faisant, les mécanismes d'atténuation peuvent être appliqués plus

efficacement, en équilibrant les objectifs de sécurité avec la continuité du service et l'efficacité des ressources.

La solution proposée suit une évolution de recherche progressive. La première étape se concentre sur un contexte d'apprentissage centralisé, où le problème de prédiction de la durée des cyberattaques est formulé comme une tâche de prévision de séries temporelles. Cette formulation permet l'exploitation des dépendances séquentielles dans le trafic réseau et ouvre la voie à l'utilisation d'architectures d'apprentissage profond spécifiquement conçues pour la modélisation temporelle. Dans ce contexte, un modèle basé sur le Transformer est développé pour capturer les dépendances à court et long terme dans l'évolution des attaques. Grâce à son mécanisme d'auto-attention, l'architecture Transformer offre un moyen puissant de modéliser des schémas temporels complexes à travers plusieurs catégories d'attaques.

Bien que cette première étape démontre la faisabilité et l'efficacité de la modélisation temporelle profonde pour la prédiction de la durée des attaques, elle reste limitée par les contraintes de l'entraînement centralisé. Dans les infrastructures 5G et O-RAN réelles, les données de trafic sont distribuées à travers plusieurs nœuds et domaines administratifs, rendant l'agrégation brute des données indésirable tant du point de vue de la confidentialité que de la perspective opérationnelle. De plus, les modèles entraînés dans des contextes de référence centralisés peuvent voir leurs performances se dégrader lorsqu'ils sont déployés dans des environnements réels en raison de la dérive de domaine.

Pour surmonter ces limitations, la deuxième étape de cette thèse étend le problème aux contextes d'apprentissage distribué à travers la conception d'un cadre de Méta-Apprentissage Fédéré pour la Prédiction de la Durée des Attaques (FML-AD). Ce cadre combine l'apprentissage fédéré avec la sélection adaptative des clients et le raffinement basé sur le méta-apprentissage afin de répondre simultanément à trois défis majeurs : la préservation de la confidentialité, l'hétérogénéité statistique entre les clients distribués et l'adaptation de domaine aux conditions réelles du réseau. Ce faisant, le cadre proposé va au-delà de la prédiction précise sur des jeux

de données hors ligne et cible le déploiement pratique dans des environnements 5G O-RAN opérationnels.

Les principales contributions de cette thèse peuvent être résumées comme suit :

- **Formulation de la prédiction de la durée des cyberattaques comme un problème de prévision de séries temporelles :** Cette thèse formalise l'estimation de la durée restante des cyberattaques en cours comme une tâche de prédiction temporelle. Cette formulation introduit une intelligence prédictive dans les processus d'atténuation et fournit une base pour passer d'une défense réactive à une gestion de la sécurité proactive.
- **Conception d'un modèle basé sur le Transformer pour la prédiction centralisée de la durée des attaques :** Une architecture d'apprentissage profond basée sur le Transformer est développée pour modéliser l'évolution temporelle des cyberattaques à partir de données de trafic séquentielles. Le modèle proposé exploite l'auto-attention pour capturer des dépendances temporelles complexes et fournit une prédiction précise de la durée à travers plusieurs catégories d'attaques.
- **Développement d'un cadre distribué de Méta-Apprentissage Fédéré (FML-AD) :** Pour soutenir le déploiement dans des environnements 5G O-RAN distribués, cette thèse introduit FML-AD, un nouveau cadre qui intègre l'apprentissage fédéré et le méta-apprentissage pour la prédiction de la durée des attaques. Ce cadre permet un apprentissage collaboratif sans partage des données brutes et améliore l'adaptabilité dans des environnements hétérogènes.
- **Intégration de mécanismes de sélection adaptative des clients et d'adaptation de domaine :** Le cadre distribué proposé renforce la robustesse grâce à la priorisation adaptative des clients lors de l'entraînement fédéré et au raffinement du méta-modèle pour corriger les erreurs de prédiction en présence de dérive de domaine. Ces mécanismes améliorent l'efficacité de convergence et soutiennent une meilleure généralisation dans différentes conditions de déploiement.
- **Validation par évaluation hors ligne et déploiement réel en 5G O-RAN :** Les approches proposées sont évaluées à la fois par des expériences contrôlées hors ligne et par un déploiement dans un banc d'essai 5G O-RAN réel. Cette validation démontre la faisabilité

pratique du cadre et son efficacité en termes de précision de prédiction, d'efficacité d'entraînement et de performance opérationnelle en temps réel.

Dans l'ensemble, cette thèse contribue à l'avancement de la cybersécurité dans les réseaux mobiles de prochaine génération en introduisant des mécanismes prédictifs et adaptatifs pour l'estimation de la durée des attaques. En combinant la modélisation temporelle, l'intelligence distribuée et la validation en monde réel, elle fournit un cadre cohérent pour améliorer les décisions d'atténuation dans des environnements 5G et O-RAN complexes.

1.5 Travaux Connexes

Cette section passe en revue la littérature la plus pertinente liée à la prédiction de la durée des cyberattaques et à l'apprentissage distribué dans les environnements 5G et O-RAN. Elle met en évidence l'évolution des approches traditionnelles basées sur la détection vers des modèles de prédiction temporelle et l'intelligence distribuée, et discute des limitations qui motivent les contributions de cette thèse.

1.5.1 Prédiction de la Durée des Attaques et Approches Centralisées

La majorité des recherches existantes en cybersécurité s'est principalement concentrée sur la détection et la classification des attaques, avec une attention comparativement limitée accordée à la caractérisation temporelle des cyberattaques. En particulier, la prédiction de la durée de persistance d'une attaque dans le réseau reste un problème peu exploré, malgré son importance pour permettre des stratégies d'atténuation plus efficaces et adaptatives.

Les premiers efforts en sécurité réseau s'appuyaient sur des systèmes à base de règles et des analyses statistiques pour détecter les anomalies dans les schémas de trafic. Bien qu'efficaces pour les signatures d'attaques connues, ces approches manquent d'adaptabilité et sont incapables de faire face à la nature dynamique des cybermenaces modernes. L'émergence des techniques d'apprentissage automatique (ML) a considérablement amélioré les capacités de détection en

permettant aux modèles d'apprendre des schémas complexes à partir des données. Cependant, la plupart des solutions basées sur le ML sont conçues pour des tâches de classification et ne traitent pas explicitement la prédiction temporelle.

Dans un effort pour aller au-delà des stratégies d'atténuation statiques, l'apprentissage par renforcement (RL) a été exploré comme une approche prometteuse pour la prise de décision adaptative en cybersécurité. En particulier, des méthodes basées sur le RL ont été proposées pour estimer la durée des attaques par déni de service distribué (DDoS) Bousalem *et al.* (2023a). Dans ces approches, le processus d'atténuation est modélisé comme un problème de prise de décision séquentielle, où un agent apprend des politiques optimales basées sur les états du réseau observés et des signaux de récompense. Cela permet un ajustement dynamique des actions d'atténuation, comme le maintien ou la libération du trafic isolé, afin d'équilibrer sécurité et qualité de service (QoS).

Malgré leur adaptabilité, les approches basées sur le RL présentent plusieurs limites. Elles nécessitent généralement un entraînement extensif et une ingénierie soigneuse des récompenses, les rendant sensibles au réglage des hyperparamètres et difficiles à stabiliser. De plus, leur capacité à capturer les dépendances temporelles à long terme est limitée, en particulier dans les scénarios impliquant des attaques de longue durée ou complexes à plusieurs étapes. Ces défis réduisent leur efficacité dans des environnements à grande échelle et hétérogènes tels que les réseaux 5G.

Des travaux plus récents ont exploré les techniques de modélisation de séries temporelles pour l'estimation de la durée des cyberattaques. Par exemple, Sun et al. Sun, Lu, Ho & Li (2023) ont proposé des modèles analytiques pour l'estimation en temps réel de la durée des attaques par déni de service (DoS). Bien que ces méthodes fournissent des insights théoriques sur le comportement temporel des attaques, elles reposent sur des formulations mathématiques complexes et sont généralement restreintes à des types d'attaques spécifiques, limitant leur applicabilité dans des environnements réels.

En parallèle, les architectures d'apprentissage profond (DL) ont été de plus en plus adoptées pour la modélisation de données séquentielles en cybersécurité. Les réseaux de neurones récurrents (RNN) et les modèles à mémoire à court et long terme (LSTM) ont été largement utilisés pour capturer les dépendances temporelles dans le trafic réseau. Cependant, ces architectures souffrent de limitations telles que le problème du gradient qui disparaît et une capacité restreinte à modéliser des dépendances à longue portée.

Pour surmonter ces problèmes, les architectures basées sur les Transformers ont récemment émergé comme une alternative puissante pour la modélisation de séquences. En exploitant les mécanismes d'auto-attention, les Transformers peuvent capturer à la fois les dépendances à court et à long terme plus efficacement que les modèles récurrents. Plusieurs études ont appliqué les Transformers à des tâches de détection d'intrusion en sécurité réseau. Par exemple, Aloufi et al. Hamhoum, Lakhdar & Cherkaoui (2024) ont proposé un système de détection d'intrusion basé sur les Transformers au sein d'une architecture Zero Trust pour les réseaux 5G, atteignant des performances de détection élevées. De même, Zhou et al. Wang, Jian, Tan, Wu & Huang (2023b) ont introduit un modèle de détection d'intrusion basé sur les Transformers avec auto-supervision, capable de s'adapter à des schémas d'attaques diversifiés. Khan et al. Sun *et al.* (2022) ont en outre combiné des modèles Transformer avec l'apprentissage fédéré pour la sécurité des réseaux intelligents, démontrant une forte précision de classification.

Malgré ces avancées, les approches existantes basées sur les Transformers restent principalement axées sur la détection des attaques plutôt que sur la prédiction temporelle. De plus, beaucoup de ces modèles sont entraînés sur des jeux de données de référence tels que UNSW-NB15 qui, bien que complets, sont principalement conçus pour des tâches de classification. Les études exploitant ce jeu de données, notamment les systèmes de détection d'intrusion basés sur CNN et DL Vibhute *et al.* (2024); Aleesa, Thanoun, Mohammed & Sahar (2021), soulignent l'efficacité de l'apprentissage profond pour la détection d'anomalies mais ne traitent pas le problème de prédiction de la durée des attaques.

Dans l'ensemble, les approches centralisées ont considérablement amélioré les capacités de détection des cyberattaques, mais elles restent limitées dans leur capacité à modéliser la dynamique temporelle des attaques et à soutenir des stratégies d'atténuation prédictives dans des environnements réels.

1.5.2 Apprentissage Distribué et Respectueux de la Vie Privée en 5G/O-RAN

La nature distribuée et hétérogène des réseaux 5G et O-RAN a suscité un intérêt croissant pour les paradigmes d'apprentissage décentralisé, en particulier l'apprentissage fédéré (FL). Dans de tels environnements, les données réseau sont générées à travers plusieurs nœuds, rendant l'agrégation centralisée des données peu pratique et soulevant d'importantes préoccupations en matière de confidentialité.

L'apprentissage fédéré répond à ces défis en permettant un entraînement collaboratif des modèles entre des clients distribués sans partage des données brutes. Le travail fondateur de McMahan et al. McMahan, Moore, Ramage, Hampson & y Arcas (2017) a introduit l'optimisation fédérée efficace en communication, qui a depuis été étendue à diverses applications dans la sécurité des réseaux. Li et al. Li, Huang, Yang, Wang & Zhang (2020b) ont démontré l'efficacité du FL dans des scénarios de sécurité IoT, soulignant sa capacité à préserver la confidentialité des données tout en maintenant des performances compétitives.

Cependant, les systèmes FL font face à des défis significatifs en raison de l'hétérogénéité statistique, où les distributions de données diffèrent entre les clients. Ce problème peut conduire à une convergence lente et à une performance dégradée du modèle. Pour y remédier, plusieurs approches ont été proposées, notamment l'algorithme FedProx Li *et al.* (2020a), qui introduit un terme proximal pour améliorer la stabilité de l'entraînement dans des conditions non-IID. Cette approche a été appliquée avec succès dans la sécurité du découpage réseau 5G Wang, Zhang, Liu & Chen (2023a) et dans des environnements d'informatique mobile en périphérie Chen, Wang, Li & Zhou (2022).

Des travaux plus récents se sont concentrés sur l'amélioration des stratégies d'agrégation et des mécanismes de sélection des clients dans des contextes fédérés. Bouzinis et al. Bouzinis *et al.* (2025) ont proposé StatAvg, une méthode d'agrégation statistique conçue pour renforcer la robustesse face à des distributions de données très hétérogènes. Doriguzzi-Corin et Siracusa Doriguzzi-Corin & Siracusa (2024) ont introduit le cadre FLAD, qui incorpore une sélection adaptative des clients pour prioriser les clients sous-performants et améliorer l'efficacité de convergence et les performances de détection.

En plus des avancées algorithmiques, le déploiement de modèles d'apprentissage automatique au sein des architectures O-RAN a également été exploré. Dzaferagic et al. Dzaferagic *et al.* (2025) ont démontré la faisabilité du déploiement de modèles ML en tant que xApps au sein du Contrôleur Intelligent du Réseau d'Accès Radio (RIC), soulignant l'intégration de l'intelligence pilotée par les données dans de vrais environnements réseau. De même, Phyu et al. Phyu, Stanica & Naboulsi (2023) ont proposé un cadre d'apprentissage fédéré pour la prédiction du trafic à travers des stations de base distribuées, mettant l'accent sur la scalabilité et la préservation de la confidentialité.

Malgré ces développements, la plupart des approches basées sur le FL existantes en sécurité réseau se concentrent principalement sur la détection des attaques et la classification des anomalies. Très peu d'attention a été accordée aux tâches de prédiction temporelle telles que l'estimation de la durée des cyberattaques. De plus, les défis liés à l'adaptation de domaine et au déploiement en monde réel restent largement non résolus, car les modèles entraînés sur des jeux de données de référence présentent souvent des performances dégradées lorsqu'ils sont appliqués à des environnements réseau opérationnels.

1.5.3 Lacunes de Recherche et Positionnement

La revue de la littérature met en évidence plusieurs lacunes importantes dans la recherche actuelle en cybersécurité pour les environnements 5G et O-RAN.

Premièrement, bien que des progrès significatifs aient été réalisés dans la détection et la classification des attaques à l'aide de techniques ML et DL, la dimension temporelle des cyberattaques reste largement peu explorée. En particulier, la prédiction de la durée des attaques en cours a reçu une attention limitée, malgré son importance critique pour permettre des stratégies d'atténuation proactives et adaptatives.

Deuxièmement, les approches existantes qui abordent la prédiction de la durée des attaques sont généralement restreintes à des types d'attaques spécifiques ou s'appuient sur des modèles analytiques qui manquent de scalabilité et de flexibilité dans des environnements réels. Les méthodes basées sur l'apprentissage par renforcement, bien qu'adaptatives, souffrent de défis liés à la stabilité, à la complexité de l'entraînement et à la capacité limitée à capturer les dépendances temporelles à long terme.

Troisièmement, la plupart des modèles d'apprentissage profond avancés, y compris les architectures basées sur les Transformers, sont développés dans des contextes centralisés à l'aide de jeux de données de référence. Bien que ces modèles démontrent de fortes performances dans des environnements contrôlés, ils ne parviennent pas à répondre aux défis clés liés à la confidentialité des données, à la propriété des données distribuées et à la dérive de domaine entre les données d'entraînement et les conditions réelles du réseau.

Enfin, bien que l'apprentissage fédéré ait émergé comme un paradigme prometteur pour l'entraînement de modèles distribués et respectueux de la vie privée, son application en cybersécurité a été largement limitée aux tâches de détection. L'intégration de la prédiction temporelle, de l'apprentissage distribué et de l'adaptation de domaine reste un défi de recherche ouvert.

Cette thèse répond à ces limitations en proposant un cadre unifié qui combine des techniques de modélisation temporelle, d'apprentissage fédéré et de méta-apprentissage pour la prédiction de la durée des cyberattaques. En comblant le fossé entre la modélisation centralisée et le déploiement distribué, l'approche proposée permet une prédiction précise, scalable et respectueuse de la vie privée de la durée des cyberattaques dans de vrais environnements 5G O-RAN.

1.6 Conclusion

Ce chapitre a établi le contexte général de la thèse en examinant l'évolution des réseaux 5G et O-RAN ainsi que les nouveaux défis en matière de cybersécurité introduits par leurs architectures distribuées, pilotées par les logiciels et ouvertes. Il a mis en évidence les limitations croissantes des stratégies d'atténuation traditionnelles, qui restent largement réactives et ne disposent pas de la capacité à intégrer l'intelligence temporelle dans les processus de prise de décision.

Dans ce contexte, la prédiction de la durée des cyberattaques a été identifiée comme un problème critique et peu exploré, avec des implications fortes pour améliorer l'efficacité de l'atténuation, optimiser l'allocation des ressources et préserver la qualité de service dans des environnements réseau dynamiques. Le chapitre a en outre démontré comment l'intégration de la prédiction temporelle permet un glissement vers des mécanismes de cybersécurité proactifs et adaptatifs.

Une revue exhaustive de la littérature a ensuite été menée, couvrant à la fois les approches centralisées pour la modélisation des cyberattaques et les paradigmes d'apprentissage distribué tels que l'apprentissage fédéré. Cette analyse a révélé plusieurs limitations clés, notamment le manque de focus sur la prédiction temporelle, les contraintes de l'entraînement centralisé et les défis liés à l'hétérogénéité des données et à l'adaptation de domaine dans les déploiements réels.

Sur la base de ces observations, la lacune de recherche a été clairement identifiée, motivant le développement d'un cadre unifié combinant modélisation temporelle, apprentissage distribué et mécanismes adaptatifs pour la prédiction de la durée des cyberattaques dans les environnements 5G O-RAN.

Les chapitres suivants s'appuient sur ces fondements. Le Chapitre 2 présente une approche basée sur le Transformer pour la prédiction centralisée de la durée des attaques. Le Chapitre 3 étend ce travail aux contextes distribués à travers le cadre de Méta-Apprentissage Fédéré proposé (FML-AD). Nous concluons ensuite la thèse avec des perspectives pour les recherches futures.

CHAPITRE 2

PREDICTING CYBERATTACK DURATION IN NEXT GENERATION NETWORKS : A NOVEL TRANSFORMER-BASED APPROACH

Mohamed Anis Sakka[1] , Wael Jaafar[1] , Rami Langar[12]

[1] Département de Génie Logiciel et des TI, École de Technologie Supérieure (ÉTS),
1100 R. Notre Dame O, Montréal, QC H3C 1K3
[2] University Gustave Eiffel, LIGM-CNRS UMR 8049, F-77454, Marne-la-Vallée, France

Paper presented at the 2025 IEEE International Conference on Communications (ICC), June 10, 2025, Montreal, QC, Canada.
DOI: 10.1109/ICC52391.2025.11161837.

This chapter presents the first contribution of this thesis, focusing on the problem of cyberattack duration prediction in a centralized setting. Building upon the limitations identified in Chapter 1, this work introduces a Transformer-based approach designed to capture the temporal dynamics of cyberattacks and provide accurate duration estimation.

The proposed model leverages attention mechanisms to model complex time dependencies in network traffic, enabling more effective forecasting compared to traditional machine learning and recurrent architectures. This chapter corresponds to the first published work and establishes the foundation for subsequent developments.

The remainder of this chapter presents the methodology, experimental setup, and performance evaluation of the proposed Transformer-based model.

2.1 Methodology

In this section, we present our methodology to predict the remaining time duration of cyberattacks. To do so, we start by modifying the *UNSW-NB15* dataset by including relevant features for our prediction task. Then, after feature selection and data normalization, we proceed to data splitting and sequence creation to build a new time-series dataset. Finally, we apply a Transformer-based model on the new obtained dataset to predict the remaining time of ongoing attacks. In what follows, we detail each of these steps.

Tableau 2.1 Unique Attack ID Counts and Maximum Remaining Time (Total Attack Duration) per Attack Category

Attack Category	Train Set		Validation Set		Test Set	
	Count	Max Duration	Count	Max Duration	Count	Max Duration
Analysis	6	1122.36	2	550.85	2	568.68
Backdoor	6	830.01	2	566.50	2	409.28
DoS	16	1799.34	5	1747.60	6	1707.68
Exploits	199	299.93	66	299.40	67	298.58
Fuzzers	13	3599.60	5	3504.43	5	3543.10
Generic	6	1589.02	2	1278.67	2	1050.12
Reconnaissance	6	2052.79	2	1329.32	2	1403.70
Shellcode	8	59.74	3	57.11	3	7.50
Worms	6	41.86	2	16.12	2	24.82

2.1.1 Data Preprocessing : Attack_id and Remaining_time features creation

To tackle the challenges of predicting cyberattack durations, we begin by enhancing the *UNSW-NB15* dataset, which includes diverse network traffic records essential for cybersecurity research. This dataset, with 49 features and over 2.5 million entries across nine attack types, served as the basis for our predictive model. After merging the dataset files, we removed rows with missing or irrelevant values to ensure data quality.

Our preprocessing focused on creating the *attack_id* feature, grouping flows by attack type, destination IP, and specific time thresholds (see Table 2.2) to distinguish between different attack phases. This process enabled logical segmentation of extended attacks. Once the *attack_id* was established, we calculated the *remaining_time* for each attack by deducting the cumulative duration from the total duration. This feature, central to our prediction task, transformed the dataset to align with our time-series regression approach.

To ensure temporal consistency, data was sorted by attack start time (*Stime*), allowing the model to capture sequential attack patterns accurately. Specific thresholds, as shown in Table 2.2, were also used to establish when new attack instances should begin, providing a framework for handling long-duration attacks within each category.

Tableau 2.2 Assumed Attack Duration Thresholds

Attack Type	**Threshold (seconds)**
Fuzzers	3600
Analysis	1800
Backdoors	3600
DoS	1800
Exploits	300
Generic	7200
Reconnaissance	1800
Shellcode	60
Worms	7200

2.1.2 Feature Selection and Data Normalization

Once the above two features are added to the original dataset, we start the preprocessing phase by handling missing data, where samples containing missing values were removed. Following this, Min-Max scaling was applied to standardize feature values within the range [0, 1], ensuring all features contributed equally during model training. Categorical variables were then converted to numerical values through label encoding to facilitate compatibility with different ML algorithms.

Feature selection is also critical in optimizing the model's accuracy and generalization capabilities. To identify the most relevant predictors for estimating the remaining attack duration, a dual approach combining **Information Gain (IG)** and **Pearson Correlation** was employed. Information Gain quantified the reduction in entropy when a feature was introduced, allowing the prioritization of features with strong discriminatory power related to attack duration Odhiambo

Omuya, Onyango Okeyo & Waema Kimwele (2021). On the other hand, Pearson Correlation is used to identify and eliminate highly correlated features to prevent multicollinearity, thus maintaining the model's performance integrity Deprez & Robinson (2024).

By integrating these techniques, the original set of 49 features was refined to a more effective set of 32 predictors. This careful selection process streamlined model training and contributed to the accuracy and robustness of the predictions, underscoring the importance of rigorous preprocessing in the methodology.

2.1.3 Data Splitting and Sequence Creation

In our study, predicting the remaining duration of ongoing cyberattacks is modeled as a time-series regression problem. To do so, the data preprocessing phase involves several additional critical steps : dataset splitting, sequence creation, and structuring for time-series forecasting.

In the first step, we organize the new obtained dataset by splitting it according to attack categories (e.g., DoS, Exploits, etc.) and unique attack identifiers (*attack_id*) added previously. This grouping allows the model to learn from sequences specific to each attack type, as well as variations within those types. Within each category, the data is further divided into **training (60%)**, **validation (20%)**, and **test (20%)** sets. This division ensures that each set contains distinct attack instances, preventing data leakage and enabling robust model evaluation. Table 2.1 outlines these metrics for each set, showing the distinct attack occurrences and their respective durations. This information is crucial as it directly impacts the model's ability to generalize across different attack types and durations.

Once the dataset is split, sequences are created using a sliding window approach. Each sequence consists of a fixed number of consecutive observations (e.g., 30 time steps), which collectively represent the progression of an attack over time. These sequences are generated within individual *attack_id* groups, maintaining a chronological order based on the *remaining_time* to capture the evolving nature of each attack. This method allows the model to focus on the dynamics specific to each instance.

Algorithme 2.1 Transformer Model Architecture with Attention Mechanism

Input : Input data with shape $(batch\_size, seq\_len, input\_dim)$
Output : Predicted remaining attack time

Step 1 : Input Layer :
- Input shape : $(batch\_size, seq\_len, input\_dim)$
- seq_len : Number of time steps (set to 30)
- $input\_dim$: Features per time step (set to 32)

Step 2 : Attention Mechanism :
- Apply *attention mechanism* to focus on relevant time steps in the sequence.
- Compute *Query (Q)*, *Key (K)*, and *Value (V)* using linear transformations.
- Compute attention scores :

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$

where d_k is the dimensionality of the keys. Apply softmax to obtain attention weights.

Step 3 : Flatten Layer :
- Flatten the output to a 2D tensor $(batch\_size, features)$ for further processing.

Step 4 : Fully Connected Layers :
- Apply fully connected layers with ReLU activation :
- **Layer 1 :** Input : $seq\_len \times input\_dim$, Output : 64 (ffdim).
- **Layer 2 :** Input : 64, Output : 32.
- **Layer 3 :** Input : 32, Output : 16.

Step 5 : Output Layer :
- Final output layer with one node to predict remaining attack time.

Step 6 : Training Details :
- **Loss Function :** Mean Absolute Error (MAE).
- **Optimizer :** Adam optimizer.
- **Batch Size :** 64.
- **Epochs :** 50.

By modeling our problem as a time-series regression one, the new obtained dataset is equipped to leverage temporal patterns in attack data. Indeed, each input sequence contains historical network activity, enabling any ML/DL model to forecast the remaining time of an attack based on its past behavior. This approach helps us in anticipating attack duration, which is crucial for

timely and efficient threat mitigation in network security contexts. The new obtained dataset can be found here [1].

2.1.4 Proposed Transformer-based Model architecture

To predict the remaining time of ongoing cyberattacks, we propose a Transformer-based model with self-attention. Indeed, Transformers are known for capturing long-range dependencies, which makes it well-suited for network security tasks Vaswani *et al.* (2023).

A key advantage of Transformers lies in their self-attention mechanism, which identifies relevant sequence parts, making the model adept at handling diverse and evolving attack patterns. This adaptability enables the model to deliver robust predictions, essential for proactive cybersecurity in dynamic threat landscapes.

Our proposed model uses an encoder-based architecture to convert input sequences into rich latent representations, which then pass through fully connected layers to predict attack duration, as described in Algorithm 1.

2.2 Simulation Results

In this section, we present the simulation results of various models evaluated in our study : RNN, LSTM, Transformers, and XGBoost. We compare them in terms of Mean Absolute Error (MAE) with the best hyperparameters combination obtained after auto-tuning, tested on the UNSW-NB15 dataset, and finally, evaluate the performance of the Transformer-based model on a DoS attacks scenario.

2.2.1 Hyperparameter Auto-Tuning

Hyperparameter tuning is a critical step for enhancing models' performance, yet traditional methods like grid search can be computationally expensive and time-consuming. To address

[1] https://github.com/sakkovic/attack-duration-dataset.git

Tableau 2.3 Hyperparameters for RNN, LSTM, TRANSFORMER, and XGBOOST Models

Model	Hyperparameter	Value
LSTM	Hidden Dimension	218
	Number of Layers	1
	Dropout	0.439
	Batch Size	64
	Optimizer	Adam
	Loss Function	MAE
	Learning Rate	0.00078
	Number of Epochs	100
RNN	Hidden Dimension	250
	Number of Layers	1
	Dropout	0.172
	Batch Size	64
	Optimizer	Adam
	Loss Function	MAE
	Learning Rate	0.00096
	Number of Epochs	100
XGBoost	Objective	MSE
	Loss Function	MAE
	Learning Rate	0.1
	Maximum Depth	6
	Random Seed	42
	Early Stopping Rounds	10
	Number of Boosting Rounds	1000
Transformer	Sequence Length	30
	Input Dim	32
	Feed-Forward Dim	64
	Learning Rate	0.00004
	Batch Size	64
	Optimizer	Adam
	Loss Function	MAE
	Epochs	50

this, we adopted Optuna Akiba, Sano, Yanase, Ohta & Koyama (2019), a state-of-the-art hyperparameter optimization framework that efficiently searches for optimal hyperparameters through dynamic exploration. Unlike manual tuning methods, Optuna employs advanced search techniques such as Bayesian optimization to intelligently navigate the search space, reducing the

time and computational overhead, while increasing the likelihood of discovering the best model configurations. This automated approach not only accelerated the tuning process but also led to improved prediction accuracy and model robustness. The selected hyperparameters for each model are summarized in Table 2.3.

2.2.2 Models' Performance Evaluation

Tableau 2.4 Mean Absolute Error (MAE) for Train and Validation Sets

Model	**Training MAE**	**Validation MAE**
XGBoost	2.0930	5.7838
RNN	14.1235	13.7692
LSTM	4.6206	4.8406
Transformer	**1.3717**	**1.9421**

Table 2.4 highlights the performance of the four studied models (XGBoost, RNN, LSTM, and Transformers) in terms of MAE on both training and validation datasets. MAE is crucial for regression tasks as it reflects the average error magnitude, indicating the model's prediction accuracy.

As seen in Table 2.4, the Transformer-based model outperforms other studied ML/DL models by achieving the lowest training MAE of 1.3717 and validation MAE of 1.9421, with nearly 60% reduction compared to the well-known LSTM model. Despite its ability to handle sequential data, LSTM, which appears to be the second best model, did not achieve the same level of generalization, particularly when dealing with complex attack types like DoS and Exploits.

RNN, although similar to LSTM in architecture, performs worse, with a validation MAE of 13.7692 (corresponding to an increase of 86% compared to the Transformer-based model), which indicates that RNN struggles to adapt to the diverse attack types in the dataset. This highlights its limitations in capturing long-term dependencies as effectively as the LSTM and Transformer models.

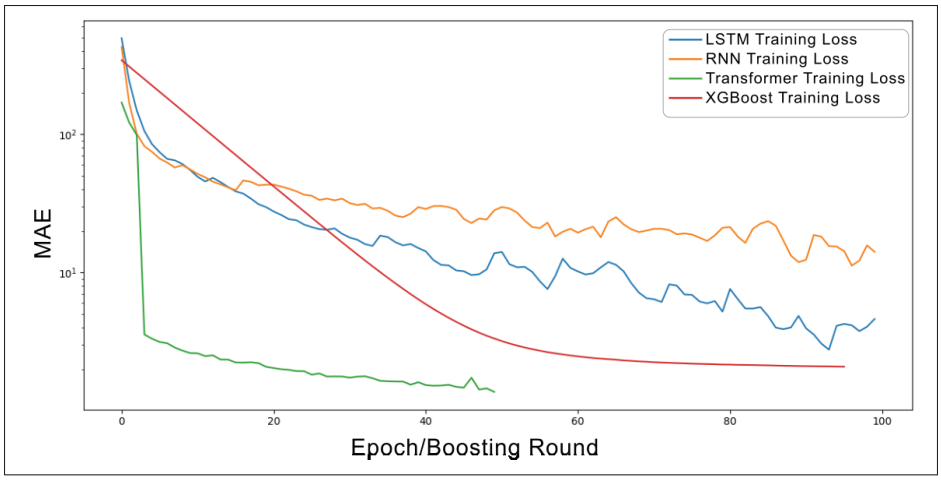


Figure 2.1 Evolution of Error during Training for the studied DL/ML models

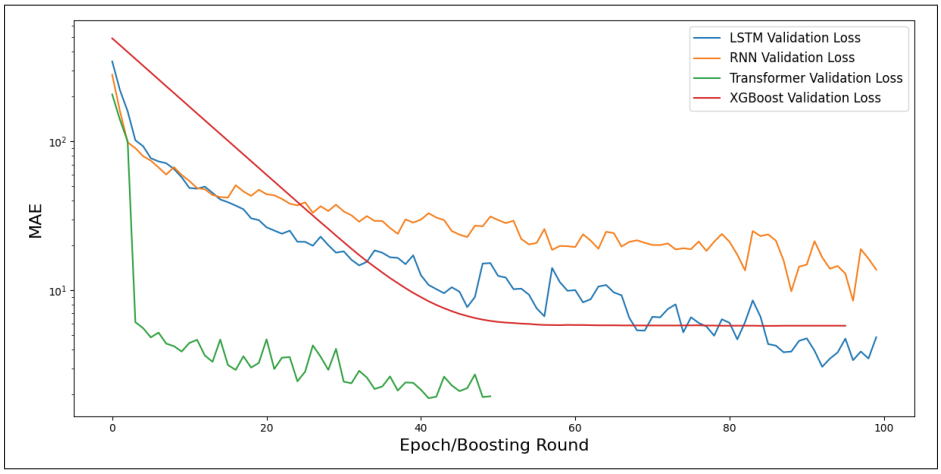


Figure 2.2 Evolution of Error during Validation for the studied DL/ML models

On the other hand, XGBoost, a tree-based model, shows a training MAE of 2.0930 and validation MAE of 5.7838 (corresponding to an increase of 66% compared to the Transformer-based model), performing moderately well but not as efficiently as the DL models in generalizing across varied attack categories.

Moreover, the convergence speed of the Transformer-based model is noticeably faster than the other models, as depicted in Figs. 2.1 and 2.2, where we show the training and validation loss

curves, respectively. In those figures, we can see that the Transformer-based model reached lower errors early in the process and maintained them consistently throughout training and validation sets.

These results highlight the ability of the Transformer model to capture underlying patterns and generalize to unseen data, making it highly effective for predicting the remaining duration of ongoing attacks across different attack types.

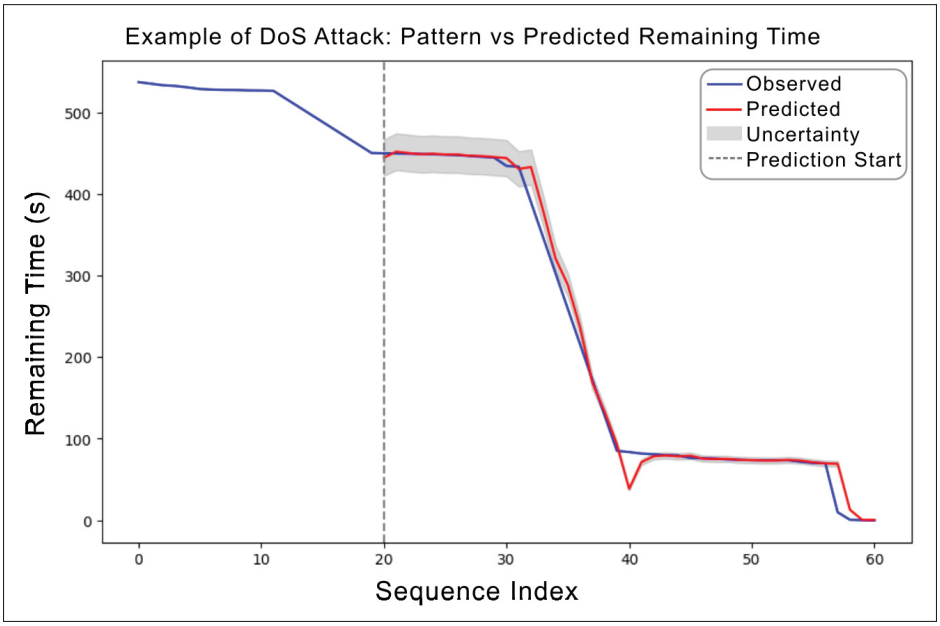


Figure 2.3 Transformer-based Model Performance : Actual Duration vs. Predicted Duration

2.2.3 Testing results on DoS attacks scenario

To further assess the performance of the Transformer-based model in real-time, we compare in Fig. 2.3 the predicted and observed values of the remaining time on DoS attacks scenario.

In this scenario, we predict the remaining time of the attack beginning from sequence 20 of our obtained dataset. As shown in Fig. 2.3, the predicted values (in red) closely align with the actual observed values (in blue), highlighting the model's robustness and accuracy in forecasting ongoing attack durations. The narrow gray shaded area represents uncertainty and

further underscores the model's confidence in its predictions. This consistent performance across various sequences illustrates the model's reliability in handling different attack types, sequence numbers, and durations. Such accurate predictions are essential for effectively applying the model in diverse cyberattack scenarios, providing timely insights into attack durations, which are crucial for enhanced mitigation and resource allocation strategies.

2.3 Conclusion

In this study, we presented a comprehensive approach for predicting the remaining duration of ongoing cyberattacks by proposing a Transformer-based model with an attention mechanism. Using the well-know UNSW-NB15 dataset, we first explain how to add two new features : attack_ID and remaining_time, and then reformatting it into sequences for time-series regression. These steps are crucial to capture temporal dynamics of the attacks and then predict attack duration with high prediction accuracy. Through simulations, we showed that our Transformer-based model outperforms traditional models like LSTM, XGBoost, and RNN in both training and validation with a validation MAE reduction up to 60%, 66%, and 86%, respectively. These results highlight its capacity to predict attack duration across various types of attacks accurately.

As future work, we plan to implement our model in a real 5G O-RAN platform to assess its efficacy in a live network environment, where accurately predicting the remaining duration of attacks will directly impact mitigation strategies. Furthermore, we aim to explore federated learning techniques to enable decentralized model training across multiple nodes, enhancing data privacy and security, while maintaining high prediction accuracy. This approach would further extend the model's utility in large-scale, distributed network systems where the protection of sensitive data and scalability are paramount.

CHAPITRE 3

FML-AD : A FEDERATED META-LEARNING FRAMEWORK FOR CYBERATTACK DURATION PREDICTION IN 5G O-RAN ENVIRONMENTS

Mohamed Anis Sakka[1] , Fahdah Alalyan[1] , Wael Jaafar[1] , Rami Langar[12]

[1] Département de Génie Logiciel et des TI, École de Technologie Supérieure (ÉTS), 1100 R. Notre Dame O, Montréal, QC H3C 1K3
[2] University Gustave Eiffel, LIGM-CNRS UMR 8049, F-77454, Marne-la-Vallée, France

Article submitted to IEEE Transactions on Network and Service Management (TNSM), March 2026

This chapter extends the previous contribution by addressing the limitations of centralized approaches in real-world 5G O-RAN environments. While the Transformer-based model presented in Chapter 2 demonstrated strong performance, its reliance on centralized data raises challenges related to privacy, scalability, and domain adaptation.

To overcome these limitations, this chapter introduces FML-AD, a federated meta-learning framework for cyberattack duration prediction. The proposed approach combines federated learning with adaptive client selection and meta-model refinement to enable distributed, privacy-preserving, and robust prediction across heterogeneous network environments.

This chapter corresponds to the second published work and represents the evolution of the proposed solution toward practical deployment in real 5G O-RAN systems.

3.1 Proposed Attack Duration Prediction Method : FML-AD

This section presents Federated Meta-Learning for Attack Duration Prediction (FML-AD), a novel decentralized framework designed to estimate the remaining duration of ongoing DDoS attacks across distributed and heterogeneous 5G O-RAN environments. FML-AD enables collaborative model training without raw data sharing, thus addressing concerns related to privacy, scalability, and operational efficiency.

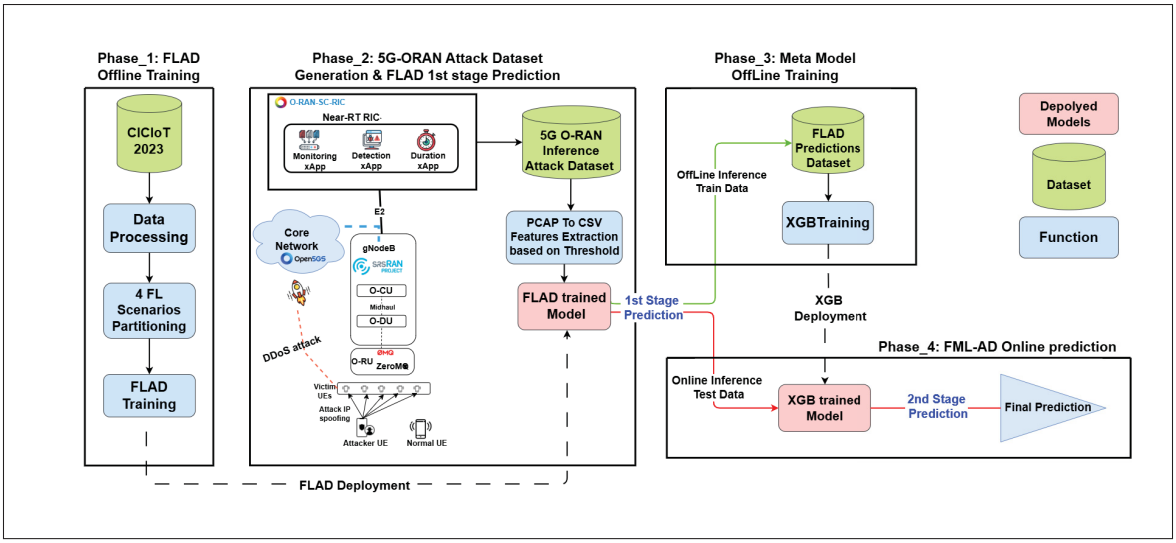


Figure 3.1 FML-AD workflow illustrating the four-phase methodology : (1) FLAD Offline Training, (2) 5G-O-RAN Dataset Generation and Initial FLAD Prediction, (3) Improving Attack Duration Prediction using a Novel Meta-Model (Offline), (4) FML-AD Online Prediction

Our methodology builds on our previous work in Sakka, Jaafar & Langar (2025), which established a robust foundation for predicting cyberattack durations using a Transformer-based architecture. The original model effectively captured temporal patterns within network flow sequences through its self-attention mechanism. However, the centralized learning architecture presented critical limitations for real-world deployment, including privacy concerns from raw data aggregation and operational barriers for multi-organizational deployment. The transition to FL and opting for the CICIoT2023 dataset Neto *et al.* (2023) addresses both privacy limitations and data volume constraints, thus creating a solid foundation for realistic network simulation and effective attack duration prediction.

As illustrated in Fig. 3.1, our framework integrates adaptive FLAD and meta-learning refinement through a structured four-phase approach. Both the FLAD-trained Transformer model and the XGBoost meta-model undergo comprehensive offline training and are subsequently deployed as a unified pipeline within a dedicated duration prediction xApp on the RIC. This architecture orchestrates sequential processing for real-time inference while maintaining the accuracy benefits of thorough model training. The four phases are detailed as follows :

3.1.1 Phase 1 : FLAD Offline Training

The proposed framework starts with comprehensive data preparation and federated model training using the CICIoT2023 dataset Neto *et al.* (2023). This dataset provides extensive IoT network traffic data with 49 features covering statistical, time-based, and protocol-specific attributes, serving to develop accurate duration prediction capabilities. We focus specifically on DDoS attacks and their subcategories within CICIoT2023, which offer substantial attack samples suitable for robust FL model development. Compared to our previous work with UNSW-NB15, CICIoT2023 has a richer DDoS data representation.

3.1.1.1 Data Preprocessing and Temporal Modeling Pipeline

Our data transformation pipeline of phase 1, depicted in Fig. 3.1, begins with temporal segmentation. Then, we introduce an `attack_id` feature by grouping flows based on three key attributes : attack label (encompassing DDoS subcategories including UDP flood, ICMP flood, TCP flood, SYN flood, and SynonymousIP flood), protocol type, and destination IP address. This grouping enables logical segmentation of overlapping attack phases into coherent sessions, forming the basis for temporal analysis. Moreover, we define the target variable `remaining_time` for each flow k, such as :

$$\texttt{remaining\_time}_k = T_{\text{total}} - \sum_{i=0}^{k-1} \texttt{flow\_duration}_i \tag{3.1}$$

where T_{total} is the total attack duration calculated as the sum of all flow durations within the attack session, and the summation term represents the cumulative duration of all previous flows in the chronological sequence. This transformation reframes the problem as a time-series regression task, where each sample represents the time remaining until the end of the attack. To preserve temporal dependencies, all flow records are sorted by start time (`Stime`) within each attack session, hence ensuring chronological ordering to capture sequential patterns and understand attack evolution dynamics.

Subsequently, we realize constant feature removal to eliminate zero-variance columns, followed by low variance filtering using Scikit-learn's Variance Threshold with a 0.001 cutoff to remove near-constant features Pedregosa *et al.* (2011). Then, we execute multicollinearity reduction through Pearson correlation analysis to eliminate features with correlation coefficients exceeding 0.9, while preserving relevant features including `flow_duration`, `remaining_time`, `label`, and `attack_id`. Finally, statistical feature selection via F-test (ANOVA) with p-value < 0.05 is applied to ensure that only features demonstrating significant relationships with the target variable are retained Fisher (1925). This rigorous procedure reduces the feature set from 49 to 21 relevant features.

To enable time-aware learning, we transform flow records into fixed-length sequences using a sliding window technique. Specifically, within each attack session, flows are sorted by the descending order of `remaining_time`. A window of 30 consecutive flows forms an input sequence, with the target being the `remaining_time` of the subsequent flow. This procedure allows Transformer models to capture both short-term and long-term dependencies in attack evolution, and thus prepares the data for the FL process.

We categorize attacks into three duration-based classes to support robust model learning and stratified partitioning : Short attacks (less than 600 seconds), Medium attacks (600-1800 seconds), and Long attacks (greater than 1800 seconds). As shown in Fig. 3.2, the empirical distribution reveals a long-tailed pattern with most attacks being short, while a smaller subset exhibits prolonged characteristics.

3.1.1.2 Data Partitioning for FL Scenarios

To ensure model generalization across diverse attack patterns, we implement dual stratification considering both duration category and attack type. The dataset is subdivided into a training subset (274 unique attack flows representing 70% of the dataset), a validation subset (78 attack flows for 20% of the dataset), and the testing subset (39 attack flows for 10% of the dataset), all with balanced representation across duration categories, as detailed in Table 3.1.

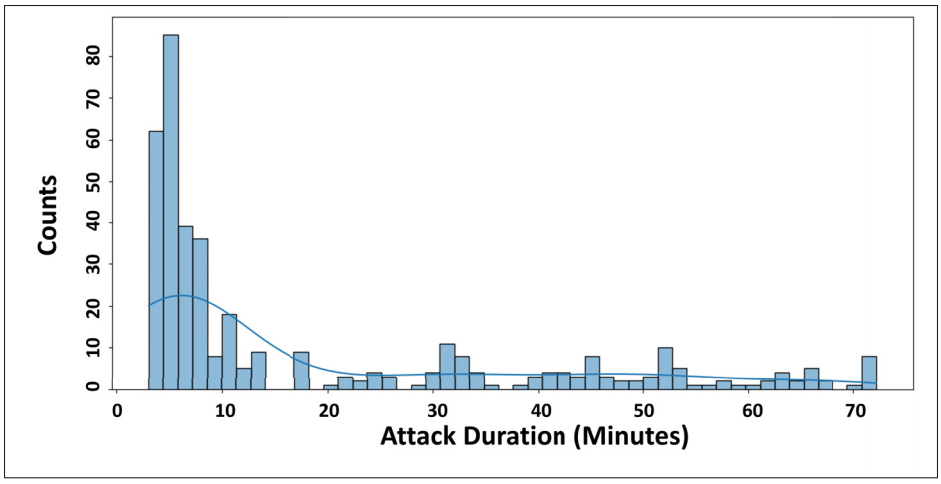


Figure 3.2 Attack Duration Distribution on CICIoT Dataset

Tableau 3.1 Attack flows data distribution

Duration Category	**Train**	**Validation**	**Test**
Short	161	47	23
Medium	41	11	5
Long	72	20	11
Total Unique Attacks	274	78	39

We emulate realistic FL environments through four distinct data partitioning strategies among the clients for training. The Independent and Identically Distributed (IID) strategy provides a balanced distribution of attack types and duration categories across all clients, serving as a baseline. The `Non-IID by Label` strategy involves clients receiving skewed distributions of attack types, simulating specialized network nodes that encounter specific attack categories. The `Non-IID by Duration` strategy provides imbalanced distributions of attack durations, reflecting scenarios where network components are exposed to only certain temporal attack patterns. Finally, the `Fully Non-IID` represents maximum heterogeneity with extreme skewing across both attack types and duration categories, representing the most challenging and realistic deployment scenario. The validation subset maintains an IID distribution across all scenarios to ensure consistent evaluation, while the testing subset is held out globally for fair performance comparison.

Algorithme 3.1 FLAD – Adaptive Client Selection

Input : Current global model $w^{(t)}$, validation losses $\{L_k\}_{k=1}^{K}$, parameters $e_{\min}$, $e_{\max}$, $s_{\min}$, $s_{\max}$
Output : Selected clients $\mathcal{S}_t$ and their training configurations $\{(e_k, s_k)\}$

Compute global avg. validation loss $\bar{L} = \frac{1}{K}\sum_{k=1}^{K} L_k$;
$\mathcal{S}_t \leftarrow \emptyset$;
foreach *client* $k \in \{1, \ldots, K\}$ **do**
    if $L_k \geq \bar{L}$ **then**
        Compute adaptive scale σ_k (eq.(3.2));
        Set epochs $e_k = e_{\min} + \sigma_k \cdot (e_{\max} - e_{\min})$;
        Set batch steps $s_k = s_{\min} + \sigma_k \cdot (s_{\max} - s_{\min})$;
        Add client k to $\mathcal{S}_t$ with config (e_k, s_k);
    end if
end foreach
return $\mathcal{S}_t$ *with* (e_k, s_k)

3.1.1.3 FLAD Framework Implementation

We implement the FLAD framework Doriguzzi-Corin & Siracusa (2024) with the optimized parameters described in Table 3.2, which corresponds to the "FLAD Training" block in Fig. 3.1. It operates through coordinated algorithms where the main training orchestrator manages the global FL workflow across multiple communication rounds. The adaptive client selection mechanism, detailed in Algo. 3.1, dynamically prioritizes underperforming clients based on validation metrics, using the deficiency score, expressed as follows :

$$\sigma_k = \frac{L_{\max} - L_k}{L_{\max} - L_{\min} + \epsilon}, \tag{3.2}$$

where L_k is the validation loss of client k, $L_{\max}$ and $L_{\min}$ are, respectively, the maximum and minimum validation losses across all clients, ϵ is a small constant for numerical stability, e_k represents the allocated training epochs, and s_k denotes the batch steps.

The output of this phase is the "FLAD-trained model" shown in Fig. 3.1 as the input of phase 2. Specifically, this global model is deployed within the duration prediction xApp on our 5G O-RAN testbed to generate initial predictions in live attack scenarios.

Tableau 3.2 Hyperparameters configuration

Parameter	**Value**
FL Parameters	
Total number of clients K	5
Number of communication rounds T	20
Early stopping patience P	5
Batch size	512
FLAD Architecture	
Input dimension	21
Sequence length	30
Hidden layer dimension	512
Number of attention heads	8
Feed Forward Network (FFN) dimension	2048
FLAD Hyperparameters	
Minimum epochs `e_min`	4
Maximum epochs `e_max`	10
Minimum steps `s_min`	2
Maximum steps `s_max`	10
Learning rate η	0.00344
Stability constant ϵ	10^{-8}
XGBoost Hyperparameters	
Learning rate	0.1
Maximum depth	5
Number of estimators `n_estimators`	100

3.1.2 Phase 2 : 5G O-RAN Dataset Generation and Initial FLAD Prediction

Phase 2 bridges offline training with real-world validation by generating a realistic inference dataset using our custom-built 5G O-RAN testbed. Our testbed emulates an O-RAN compliant architecture ORA (2023) with containerized Open5GS core network Ope (2024), modular gNodeBs implemented via the srsRAN Project SRS (2024), and virtual User Equipments (UEs) managed through ZeroMQ Zer (2024). A dedicated Monitoring xApp deployed within the Near-Real-Time RIC (Near-RT RIC) O-R (2024) monitors network traffic, performs feature extraction, and streams data for model inference.

Using our testbed, we first simulate network traffic and DDoS attacks to collect real network data, then perform initial predictions using the FLAD-trained model obtained in Phase 1. In this step, we assume that an attack is already detected by a dedicated *Attack Detection* xApp, allowing us to focus specifically on the attack duration prediction task. In particular, we simulate 10 distinct DDoS attack scenarios (TCP SYN and UDP floods) with durations ranging from 80 to 120 seconds. The raw packet capture (PCAP) data is processed through our PCAP-to-CSV conversion pipeline to extract flow-level features compatible with our model input requirements, following the feature engineering step. The trained FLAD model is then used to generate initial duration predictions on this testbed-collected dataset. The preprocessed datasets corresponding to Phase 1 and Phase 2 are publicly available at Sakka, Alalyan, Jaafar & Langar (2026).

3.1.3 Phase 3 : Improving Attack Duration Prediction using a Novel Meta-Model

Given that FLAD provides only preliminary predictions, we enhance the prediction quality in Phase 3 by implementing a meta-learning approach, where an XGBoost meta-model improves the initial FLAD predictions through error correction and domain adaptation. This phase focuses on bridging the domain shift between the CICIoT training environment and real 5G O-RAN deployments. Specifically, the meta-model is trained on the dataset generated in Phase 2, which contains FLAD predictions, ground truth durations, and temporal context features from the 5G O-RAN testbed. The XGBoost model, optimized with hyperparameters as shown in Table 3.2,

learns to map these suboptimal FLAD predictions and contextual features to refined duration estimates by capturing complex residual errors and non-linear relationships that the base model missed due to domain shift. The output of Phase 3 is the trained XGBoost meta-model, which is integrated with the FLAD model within the same duration prediction xApp on the RIC. This co-deployment creates a unified two-stage prediction pipeline where FLAD generates initial estimates, and XGBoost provides refinement, operating sequentially within the xApp runtime.

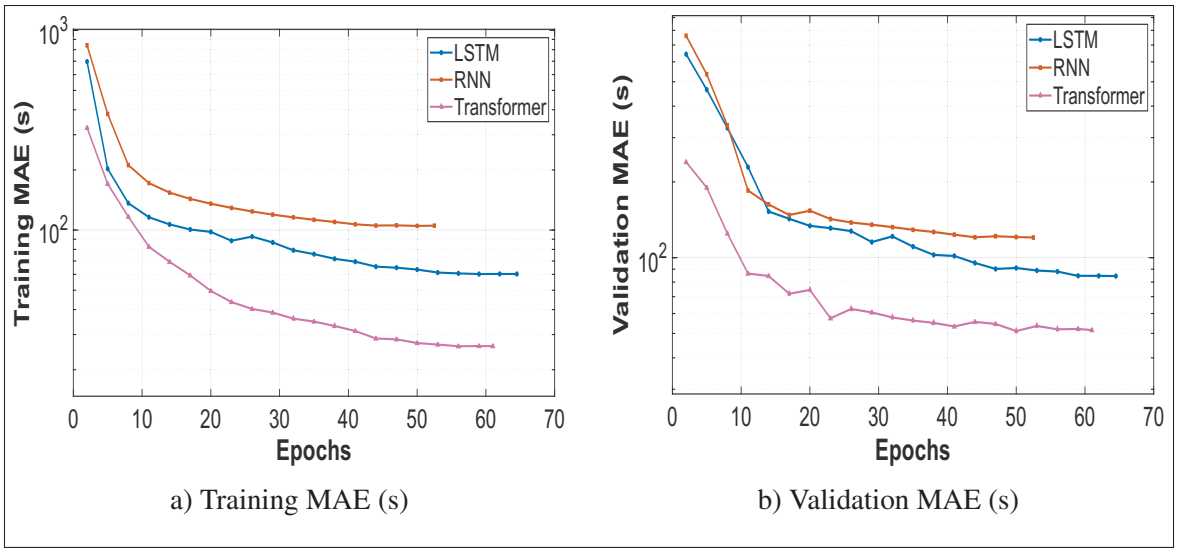
3.1.4 Phase 4 : FML-AD Online Prediction

The final phase operationalizes the complete FML-AD pipeline online, with both FLAD and XGBoost models deployed within the duration prediction xApp on the RIC. The workflow leverages three specialized xApps : (1) a monitoring xApp for feature extraction, (2) a detection xApp for identifying attacks, and (3) a duration prediction xApp that executes the two-stage forecasting.

Once an attack is detected, the duration prediction xApp processes the features through a dual-model approach : FLAD generates an initial estimate, which is then refined by an XGBoost meta-model. By operating across multiple temporal thresholds, the system dynamically updates its predictions as the attack evolves, combining FLAD's federated temporal insights with XGBoost's environment-specific error correction.

Tableau 3.3 Mean Absolute Error (MAE) and Execution Time for Training and Validation

Model	**Training MAE**	**Validation MAE**	**Training Time**
RNN	105.23	120.23	10h43min
LSTM	60.02	84.65	9h03min
Transformer	**26.24**	**51.57**	**5h54min**



a) Training MAE (s) b) Validation MAE (s)

Figure 3.3 Convergence of centralized methods (Training and validation)

Tableau 3.4 Performance Comparison of Federated Learning Strategies (Global Model Performance)

Scenario	**FedAvg**		**FedProx**		**FLAD**	
	Val MAE	**Test MAE**	**Val MAE**	**Test MAE**	**Val MAE**	**Test MAE**
IID	20.8	24.4	**20.5**	**22.1**	22.2	22.3
Non-IID (Label-based)	27.6	31.1	25.8	29.3	**21.7**	**23.4**
Non-IID (Duration-based)	26.9	31.6	25.2	31.3	**23.4**	**24.4**
Fully Non-IID	27.3	32.3	26.3	31.7	**22.3**	**24.1**
Training Time	5h53min		5h49min		**1h44min**	

3.2 Experimental Results

This section presents a comprehensive performance evaluation of our FML-AD framework, systematically progressing from centralized model evaluation through FL analysis to real-world validation on a 5G O-RAN testbed.

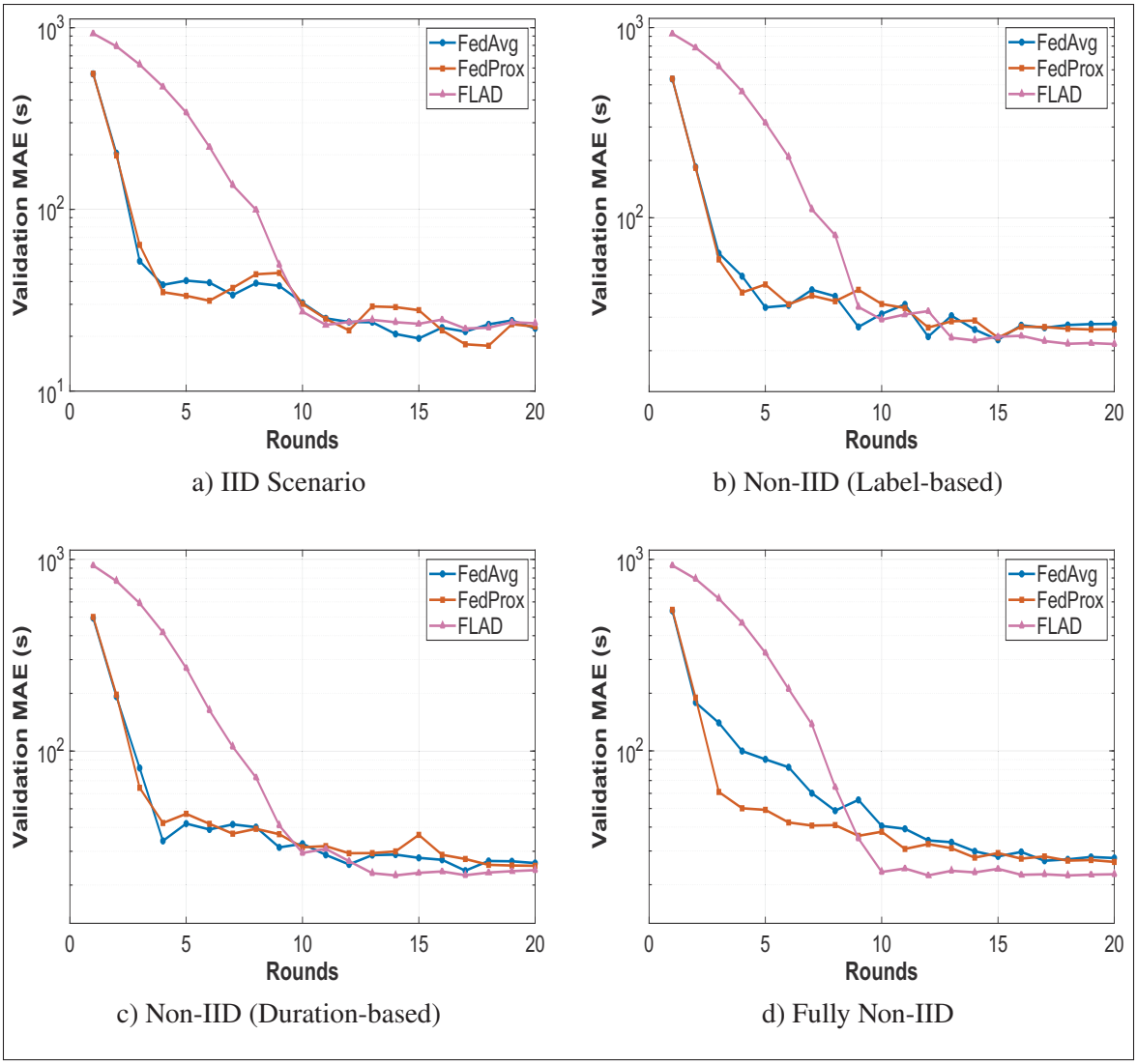


Figure 3.4 Convergence of FL methods (Different data distribution scenarios)

Our evaluation begins with centralized model evaluation to establish performance baselines and validate architectural choices for subsequent FL deployment. Specifically, we compare the proposed Transformer model to LSTM and recurrent neural network (RNN)-based benchmarks in Table 3.3. As shown in this table, the Transformer architecture outperforms the baselines in terms of MAE and training time. For instance, it achieves a validation MAE of 51.57 seconds, representing an improvement of 57.1% over the RNN model and a 39.1% improvement over LSTM while training 44.9% faster than the RNN model and 34.8% faster than the LSTM.

Fig. 3.3 corroborates these results through convergence analysis. Indeed, we can see in Fig. 3.3 that the Transformer reaches a lower MAE value faster than the LSTM and RNN baselines in both training and validation. This result is expected since the Transformer's self-attention mechanism enables a superior capture of long-range temporal dependencies than the other ones Sakka *et al.* (2025). These findings justify the subsequent selection of the Transformer as our foundational base model for implementation across FL agents.

Building on this, we evaluate the performances of three FL strategies in Table 3.4, namely FedAvg Li *et al.* (2020b), FedProx Wang *et al.* (2023a), and FLAD, with their parameters defined in Table 3.2. For this assessment, we opted for four data distribution scenarios, described as follows : *IID*, where data is distributed in a balanced manner between FL agents in terms of detection classes (attack or legitimate) and attack duration-based classes (short/medium/long), *Non-IID (Label-based)*, where data is distributed in an unbalanced manner according to the type of attacks per FL agent only, *Non-IID (Duration-based)*, where data distributed in an unbalanced manner according to their duration-based classes only, and *Full Non-IID*, where data is distributed in an unbalanced manner according to both the attack types and duration-based classes. From Table 3.4, we can see that FLAD converges the fastest during training, requiring only 1 hour 44 minutes on average, around 70% reduction compared to the centralized training approach (cf. the Transformer model in Table III) and both FedAvg and FedProx baselines. Moreover, FLAD demonstrates low MAE values, between 22.3 and 24.4 seconds across the data distribution scenarios, representing better performances between 54 and 57% than the centralized method. Compared to FedAvg and FedProx approaches, FLAD outperforms both for any non-IID scenario, and achieves very similar results to FedProx, which achieved the best IID results.

The convergence dynamics of the FL methods, under different data distribution scenarios, are illustrated in Fig. 3.4. In the IID scenario (Fig. 3.4.a), all methods demonstrate stable MAE convergence with comparable final performances. This result establishes that baseline FL methods are well-suited for the ideal IID data condition. Nevertheless, FLAD's architectural advantage becomes increasingly pronounced in non-IID scenarios. For instance, in Fig. 3.4.b,

FLAD converges slightly slower, but reaches better MAE performances than the baselines. Indeed, FLAD's adaptive client selection mechanism prioritizes underperforming clients and allocates resources proportional to validation loss deficiency scores. Similar results are obtained in Fig. 3.4.c, where FLAD converges to better MAE values than the benchmarks. Finally, in the most challenging scenario (Fig. 3.4.d), FLAD demonstrates the best results, confirming its robustness to aggressive data heterogeneity, unlike the FedAvg and FedProx methods that achieve higher MAE results.

Tableau 3.5 Meta-Model Performance Comparison on Processing 1st Stage Predictions

Meta-Model	**MAE (s)**
Linear Regression	18.146
Ridge Regression	21.351
Lasso Regression	20.790
Random Forest	5.273
Gradient Boosting	3.513
XGBoost	**2.009**

To address domain shift challenges, we evaluate multiple regression algorithms as potential meta-models for refining FLAD predictions, including linear/ridge/Lasso regressions, random forest, gradient boosting, and XGBoost. As shown in Table 3.5, XGBoost achieved the best MAE performance (equals to 2.009 seconds), which is around 43% lower than the closest benchmark.

The practical effectiveness of this meta-model enhancement is then illustrated in Fig. 3.5, which compares the performance of the standalone FLAD baseline against the complete proposed FML-AD pipeline. This comparison, conducted over a 90-second real-time DDoS TCP attack generated on our 5G O-RAN testbed, evaluates both MAE and processing latency throughout the attack lifecycle. As shown, the integrated pipeline consistently outperforms the baseline, validating the benefits of the two-stage refinement approach. Specifically, the temporal analysis reveals four distinct operational phases that characterize the framework's behavior along the

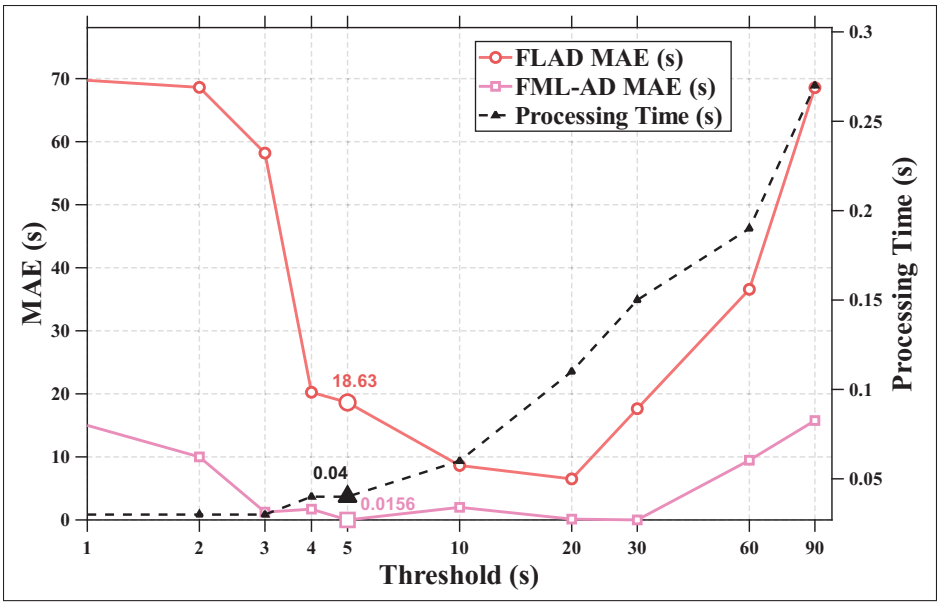


Figure 3.5 Performance comparison between FLAD and FML-AD predictions across temporal thresholds during 120-second DDoS TCP attack on 5G O-RAN testbed

attack's event (threshold reflects the progress of the attack). Indeed, during the *Initial Blind Spot Phase* (threshold in 1-3 seconds), both methods operate with minimal attack data, yet FML-AD achieves substantially better performance than FLAD with an MAE of 9 seconds compared to FLAD's MAE of 67 seconds. This early performance advantage is critical for real-time security operations, where rapid threat assessment during initial attack stages significantly impacts mitigation effectiveness. *The High-Efficacy Mitigation Window* (threshold in 3-5 seconds) emerges as the optimal operational point, where FML-AD achieves an MAE of 0.0156 seconds and only 0.04 seconds processing time. This performance represents a significant improvement over FLAD's MAE of 18.63 seconds, demonstrating FML-AD's ability to extract meaningful patterns from limited data. As the temporal context expands into the *Performance Gap Phase* (threshold in 5-30 seconds), the performance gap increases. FML-AD maintains a superior performance with an MAE often close to 0.1 seconds and consistently below 1.4 seconds MAE, while FLAD shows gradual improvement with optimal performance achieved after 20 seconds of data gathering. Finally, in the *Data Saturation Phase* (threshold above 30 seconds for FML-AD and above 20 seconds for FLAD), both methods demonstrate an MAE degradation against

information overload. However, this degradation is less pronounced for FML-AD, demonstrating its operational stability. Based on these results, the 3-5 threshold range allows FML-AD to achieve optimal MAE results, while costing less than 0.04 seconds in processing time. In comparison, FLAD achieves the best MAE on the threshold value of 20 seconds, incurring a higher processing time of 0.11 seconds.

3.3 Conclusion

In this paper, we propose FML-AD, a novel federated meta-learning framework for predicting the remaining duration of ongoing cyberattacks in 5G O-RAN environments. Building upon previous centralized methods, we introduce two key innovations within FML-AD, namely, a Transformer-based FL architecture resilient to non-IID data distributions through adaptive client prioritization, and a two-stage prediction pipeline (FLAD and XGBoost) deployed within a duration prediction xApp in O-RAN. Through our comprehensive evaluation, we demonstrated that FML-AD achieves substantial performance improvements, outperforming standalone FLAD and maintaining an average MAE of 2.0 seconds across diverse attack scenarios. Deployment on our 5G O-RAN testbed delivered high MAE performances during critical mitigation windows (thresholds) with near real-time processing. Future work will explore adaptive threshold selection for dynamic network conditions, extend the framework to multi-attack scenarios, and investigate enhanced privacy-preserving techniques within the FL paradigm.

CONCLUSION ET RECOMMANDATIONS

Cette thèse a abordé la problématique fondamentale de la prédiction temporelle des cyberattaques dans les environnements 5G et O-RAN, un enjeu critique pour la sécurité des réseaux de nouvelle génération. Face aux limitations des approches traditionnelles de détection et de mitigation, nos travaux ont exploré comment l'intelligence artificielle peut transformer la cybersécurité d'une discipline réactive en une pratique proactive et prédictive.

Dans cette perspective, notre recherche a démontré que la prédiction de la durée des attaques constitue un élément clé pour des stratégies de sécurité adaptatives. En effet, en anticipant le comportement temporel des menaces, il devient possible d'optimiser l'allocation des ressources de sécurité, de planifier les interventions de manière plus efficace et de minimiser l'impact sur la qualité de service perçue par les utilisateurs légitimes.

Pour atteindre cet objectif, le développement d'architectures basées sur les Transformers a représenté une avancée significative pour la modélisation des séquences temporelles d'attaques. La capacité de ces modèles à capturer des dépendances complexes sur de longues périodes a ouvert de nouvelles perspectives pour la compréhension de la dynamique des cybermenaces.

Parallèlement, l'extension de ces approches vers l'apprentissage fédéré a permis d'adresser les défis cruciaux de confidentialité et d'hétérogénéité des environnements réseau distribués. Cette évolution vers des paradigmes d'apprentissage collaboratif préserve la vie privée des données tout en maintenant l'efficacité des modèles de prédiction.

En poursuivant cette logique d'amélioration, l'intégration du méta-apprentissage dans le framework FML-AD a finalement permis de surmonter le défi de l'adaptation au domaine entre les environnements d'entraînement et de déploiement. Cette approche hybride représente une solution prometteuse pour le déploiement opérationnel de systèmes de prédiction dans des infrastructures réseau réelles.

Au-delà de ces contributions techniques, cette thèse a établi un nouveau paradigme pour la recherche en sécurité réseau. Le passage de la détection à la prédiction comportementale ouvre la voie à une cybersécurité plus contextuelle et adaptative, capable de s'ajuster dynamiquement à l'évolution des menaces.

Sur le plan opérationnel, les implications pratiques de ces travaux sont significatives pour les opérateurs de réseaux. La prédiction de durée d'attaque permet non seulement d'améliorer l'efficacité des mesures de sécurité, mais aussi d'optimiser l'utilisation des ressources et de garantir une meilleure expérience utilisateur. L'approche fédérée assure par ailleurs la conformité avec les réglementations croissantes en matière de protection des données.

Cependant, malgré ces avancées, plusieurs défis demeurent et ouvrent des perspectives de recherche futures. L'extension à d'autres types d'attaques, l'optimisation de l'efficacité énergétique, et l'intégration renforcée avec les architectures réseau existantes constituent autant de directions prometteuses. Le développement de mécanismes d'inférence temps-réel et la robustification contre les attaques adversariales représentent également des axes de recherche importants.

En synthèse, cette thèse contribue à l'émergence d'une cybersécurité plus intelligente et adaptative, alignée avec les exigences des réseaux modernes. La prédiction temporelle des cyberattaques s'impose comme une composante essentielle des stratégies de sécurité proactive, permettant de construire des infrastructures réseau plus résilientes face à la complexité croissante des cybermenaces.

Finalement, les approches développées dans cette recherche jettent les bases pour une nouvelle génération de systèmes de sécurité, capables non seulement de détecter les menaces mais aussi d'anticiper leur évolution et de s'adapter dynamiquement pour en minimiser l'impact, contribuant ainsi à la construction d'un écosystème numérique plus sûr et fiable.

BIBLIOGRAPHIE

O-RAN Alliance. (2023). O-RAN Architecture Description.

[Accessed : Oct. 15, 2024]. (2024). Open5GS. Repéré à https://open5gs.org/.

[Accessed : Oct. 15, 2024]. (2024). ZeroMQ. Repéré à https://zeromq.org/.

Abubakar, R., Aldegheishem, A., Majeed, M., Mehmood, A., Alrajeh, N. & Maple, C. (2020). An Effective Mechanism to Mitigate Real-time DDoS Attack Using Dataset. *IEEE Access*. doi : 10.1109/ACCESS.2020.2995820.

Akiba, T., Sano, S., Yanase, T., Ohta, T. & Koyama, M. (2019). Optuna : A Next-generation Hyperparameter Optimization Framework. *Proc. ACM Int. Conf. Knowl. Discov. and Data Mini. (SIGKDD)*.

Alalyan, F., Awad, M., Jaafar, W. & Langar, R. (2025). Secure Distributed Federated Learning for Cyberattacks Detection in B5G Open Radio Access Networks. *IEEE Op. J. of the Commun. Soc.*, 6, 3067-3081. doi : 10.1109/OJCOMS.2024.3523468.

Aleesa, A., Thanoun, M., Mohammed, A. & Sahar, N. (2021). DEEP-INTRUSION DETECTION SYSTEM WITH ENHANCED UNSW-NB15 DATASET BASED ON DEEP LEARNING TECHNIQUES. *J. of Engineer. Sci. and Technol.*, 16, 711-727.

Amrollahi, M. et al. (2020). Enhancing Network Security Via Machine Learning : Opportunities and Challenges. Dans *Handbook of Big Data Privacy* (pp. 165–189). Springer.

Banerjee, J., Maiti, S., Chakraborty, S., Dutta, S., Chakraborty, A. & Banerjee, J. S. (2019). Impact of Machine Learning in Various Network Security Applications. *Proc. Int. Conf. on Comput. Methodolo. and Commun. (ICCMC)*, pp. 276-281. doi : 10.1109/ICCMC.2019.8819811.

Bousalem, B., Silva, V. F., Langar, R. & Cherrier, S. (2022a). DDoS attacks detection and mitigation in 5G and beyond networks : A deep learning-based approach. *Proc. IEEE Glob. Commun. Conf.*, pp. 1259–1264.

Bousalem, B., Sakka, M. A., Silva, V. F., Jaafar, W., Letaifa, A. B. & Langar, R. (2023a). DDoS attacks mitigation in 5G-V2X networks : A reinforcement learning-based approach. *Proc. Int. Conf. Netw. Serv. Manag. (CNSM)*, pp. 1–5.

Bousalem, B., Silva, V. F., Langar, R. & Cherrier, S. (2022b). DDoS Attacks Detection and Mitigation in 5G and Beyond Networks : A Deep Learning-based Approach. *Proc IEEE Glob. Commun. Conf.*, pp. 1259-1264. doi : 10.1109/GLOBECOM48099.2022.10001562.

Bousalem, B., Sakka, M. A., Silva, V. F., Jaafar, W., Ben Letaifa, A. & Langar, R. (2023b). DDoS Attacks Mitigation in 5G-V2X Networks : A Reinforcement Learning-Based Approach. *Proc. Int. Conf. Netw. Serv. Mngt. (CNSM)*, pp. 1-5. doi : 10.23919/CNSM59352.2023.10327917.

Bouzinis, P. S., Radoglou-Grammatikis, P., Makris, I., Lagkas, T., Argyriou, V., Papadopoulos, G. T., Sarigiannidis, P. & Karagiannidis, G. K. (2025). StatAvg : Mitigating Data Heterogeneity in Federated Learning for Intrusion Detection Systems. *IEEE Trans. Netw. Serv. Manage.*, 22(4), 2944-2955. doi : 10.1109/TNSM.2025.3564387.

Chang, R. K. (2002). Defending Against Flooding-Based Distributed Denial-of-Service Attacks : A Tutorial. *IEEE Commun. Mag.*, 40(10), 42–51.

Chen, Y., Wang, S., Li, X. & Zhou, M. (2022). Federated Learning with FedProx for Security Anomaly Detection in MEC Environments. *IEEE Trans. Mob. Comput.*, 21(8), 2876–2890.

Deprez, M. & Robinson, E. C. (2024). Chapter 8 - Feature extraction and selection. Dans Deprez, M. & Robinson, E. C. (Éds.), *Machine Learning for Biomedical Applications* (pp. 175-192). Academic Press. doi : https://doi.org/10.1016/B978-0-12-822904-0.00013-3.

Doriguzzi-Corin, R. & Siracusa, D. (2024). FLAD : Adaptive Federated Learning for DDoS Attack Detection. *Comput. & Secu.*, 137, 103597.

Dzaferagic, M., Missi Xavier, B., Collins, D., D'Onofrio, V., Martinello, M. & Ruffini, M. (2025). ML-Based Handover Prediction Over a Real O-RAN Deployment Using RAN Intelligent Controller. *IEEE Trans. Netw. Serv. Manage.*, 22(1), 635-647. doi : 10.1109/TNSM.2024.3468910.

Fisher, R. A. (1925). *Statistical methods for research workers*. Edinburgh : Oliver and Boyd.

Hamhoum, W., Lakhdar, H. & Cherkaoui, S. (2024). Fortifying Open RAN Security with Zero Trust Architecture and Transformers. *Proc. IEEE Int. Conf. Commun. (ICC)*, 2216-2221. Repéré à https://api.semanticscholar.org/CorpusID:271940793.

Khan, M. & Ghafoor, L. (2024). Adversarial Machine Learning in the Context of Network Security : Challenges and Solutions. *J. of Computatio. Intelli. and Robot.*, 4(1), 51–63.

Li, T., Sahu, A. K., Zaheer, M., Sanjabi, M., Talwalkar, A. & Smith, V. (2020a). Federated Optimization in Heterogeneous Networks. *Proceed. of ML and Syst.*, 2, 429–450.

Li, X., Huang, K., Yang, W., Wang, S. & Zhang, Z. (2020b). Federated Learning for IoT Security : A Comprehensive Analysis. *IEEE IoT J.*

Mavoungou, S., Kaddoum, G., Taha, M. & Matar, G. (2016). Survey on Threats and Attacks on Mobile Networks. *IEEE Access*, 4, 4543-4572. doi : 10.1109/ACCESS.2016.2601009.

McMahan, B., Moore, E., Ramage, D., Hampson, S. & y Arcas, B. A. (2017). Communication-Efficient Learning of Deep Networks from Decentralized Data. *Proc. Int. Conf. Artifi. Intelli. Statist.*, 1273–1282.

Mirkovic, J. & Reiher, P. (2004). Internet Denial of Service : Attack and Defense Mechanisms. *Comput. J.*, 34(4), 52–60.

Moustafa, N. & Slay, J. (2015). UNSW-NB15 : a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set). *Proc. Milit. Commun. Info. Syst. Conf. (MilCIS)*, pp. 1-6. doi : 10.1109/MilCIS.2015.7348942.

Neto, E. C. P., Dadkhah, S., Ferreira, R., Zohourian, A., Lu, R. & Ghorbani, A. A. (2023). CICIoT2023 : A Real-Time Dataset and Benchmark for Large-Scale Attacks in IoT Environment. *Sensors*, 23(13).

O-RAN Alliance. [Accessed : Oct. 15, 2024]. (2024). O-RAN SC Projects. New York, NY, USA : O-RAN Alliance. Repéré à https://docs.o-ran-sc.org/en/latest/projects.html#near-realtime-ran-intelligent-controller-ric.

Odhiambo Omuya, E., Onyango Okeyo, G. & Waema Kimwele, M. (2021). Feature Selection for Classification using Principal Component Analysis and Information Gain. *Expert Syst. with Appl.*, 174, 114765. doi : https://doi.org/10.1016/j.eswa.2021.114765.

Pedregosa, F. et al. (2011). Scikit-learn : Machine Learning in Python. *J. of ML Res.*, 12, 2825–2830.

Phyu, H. P., Stanica, R. & Naboulsi, D. (2023). Multi-Slice Privacy-Aware Traffic Forecasting with FL at RAN Level. *IEEE Trans. Netw. Serv. Manage.*, 20(4), 5038–5052. doi : 10.1109/TNSM.2023.3267725.

Sakka, M. A., Jaafar, W. & Langar, R. (2025, Jun.). Predicting Cyberattack Duration in Next Generation Networks : A Novel Transformer-Based Approach. *Proc. IEEE Int. Conf. Commun.*, pp. 3144-3149. doi : 10.1109/ICC52391.2025.11161837.

Sakka, M. A., Alalyan, F., Jaafar, W. & Langar, R. (2026). DDoS Attack Duration prediction. IEEE Dataport. Repéré à https://dx.doi.org/10.21227/rzj7-9792.

Sedjelmaci, H. & Touati, H. (2024). Two-Layer Federated Learning for 5G Network Slicing Security. *IEEE Trans. Netw. Serv. Manage.*, 21(1), 1178–1189. doi : 10.1109/TNSM.2023.3294568.

SRS. [Accessed : Oct. 15, 2024]. (2024). srsRAN project. Repéré à https://www.srsran.com/5g.

Sun, X., Tang, Z., Du, M., Deng, C., Lin, W., Chen, J., Qi, Q. & Zheng, H. (2022). A Hierarchical Federated Learning-Based Intrusion Detection System for 5G Smart Grids. *Electron.*, 11(16). doi : 10.3390/electronics11162627.

Sun, Y., Lu, J., Ho, D. W. C. & Li, L. (2023). Real-time estimation of DoS duration and frequency for security control. *IEEE Trans. Info. Forens. Secu.*

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L. & Polosukhin, I. (2023). Attention Is All You Need. Repéré à https://arxiv.org/abs/1706.03762.

Vibhute, A. D., Khan, M., Patil, C. H., Gaikwad, S. V., Mane, A. V. & Patel, K. K. (2024). Network anomaly detection and performance evaluation of Convolutional Neural Networks on UNSW-NB15 dataset. *Procedia Comput. Sci.*, 235, 2227-2236. doi : https://doi.org/10.1016/j.procs.2024.04.211. Int. Conf. ML Data Engineer. (ICMLDE 2023).

Wang, H., Zhang, R., Liu, Y. & Chen, K. (2023a). FedProx for Enhanced Security in 5G Network Slicing. *IEEE Int. Conf. Commun.*, pp. 1–6.

Wang, W., Jian, S., Tan, Y., Wu, Q. & Huang, C. (2023b). Robust unsupervised network intrusion detection with self-supervised masked context reconstruction. *Comput. Secur.*, 128(C). doi : 10.1016/j.cose.2023.103131.

Yang, Q., Liu, Y., Chen, T. & Tong, Y. (2019). Federated Machine Learning : Concept and Applications. *ACM Trans. Intelli. Syst. Technol.*